Diplôme de conservateur
de bibliothèque (DCB)

Mémoire d'étude / janvier 2013

Bibliothèque et Open data. Et si on ouvrait les bibliothèques sur l'avenir ?

Aurore CARTIER

Sous la direction de Frédérique Schlosser

Conservatrice, responsable du département des réseaux en charge de la politique *Open data* de la bibliothèque des Champs Libres – Médiathèque des Champs Libres

Remerciements

*Mes remerciements vont en premier lieu à ma directrice de mémoire Mme **Frédérique Schlosser**, responsable du département des réseaux en charge de la politique Open data à la bibliothèque des Champs Libres, pour son soutien, ses relectures, ses conseils avisés et surtout l'esprit de confiance dans lequel elle a permis à ce travail de voir le jour. Je lui en suis très reconnaissante.*

*Dans cet esprit, je tenais également à remercier Mme **Marine Bedel**, directrice de la bibliothèque de Rennes Métropole et de la bibliothèque municipale de Rennes pour avoir proposé ce sujet qui me tenait à cœur depuis déjà un moment.*

*Nombreux sont encore les acteurs du territoire rennais à qui je souhaite exprimer ma gratitude pour le soutien, le temps et l'intérêt qu'ils ont bien voulu accorder à mon travail. Par ordre d'apparition : à la Cantine numérique rennaise, au Collectif Open Data Rennes Métropole et à la communauté d'agglomération Rennes Métropole, respectivement en les personne de **Simon Chignard**, spécialiste de l'Open data et auteur d'un ouvrage récemment paru intitulé « Comprendre l'ouverture des données publiques », de **Léa Lacroix**, co-fondatrice du Collectif et de **Bernadette Kessler**, responsable de l'ouverture des données publiques au sein de la direction générale de la Communication de l'agglomération Rennes Métropole. Chacun des entretiens qu'ils ont eu la gentillesse de m'accorder ont été des atouts précieux pour mon travail.*

*Je tenais, par ailleurs, à adresser des remerciements tout particuliers à **Romain Wenz**, responsable du projet data.bnf.fr à la Bibliothèque nationale de France pour son soutien et l'aide extrêmement précieuse qu'il m'a apporté tout au long de ce travail, et même avant. Il est pour beaucoup dans mon choix pour ce sujet et je l'en remercie sincèrement.*

*Merci également à toutes les personnes que j'ai eu l'occasion de croiser au cours de ce mémoire et avec lesquelles j'ai eu le plaisir d'échanger de façon ponctuelle ou plus durable, de vive voix, par téléphone ou par courriel : **Sarah Toulouse**, responsable du projet de bibliothèque numérique patrimoniale à la bibliothèque des Champs Libres, **Jean-Marie Bourgogne**, responsable Open data à Montpellier, **Maxime Michali**, **Clément Renaud**, **Florian Strzelecki** et **Julien Sempéré**.*

Enfin mes remerciements vont aussi à ma famille et à mes proches : à mes camarades de promotion pour leur soutien dans les moments un peu difficiles, à ma mère pour ses relectures si précieuses, à mes cousines pour leur hospitalité, et enfin à mes colocataires pour leur présence au quotidien, et tellement plus encore.

Si l'Open data est avant tout une affaire d'ouverture et de disponibilité de données, ces qualités s'appliquent tout aussi bien à chacune des personnes que la poursuite de ce travail a mis sur ma route et je ne peux que m'en féliciter.

(cc) BY-NC-ND

Résumé

Depuis le milieu des années 2000, plusieurs gouvernements et plusieurs collectivités françaises et étrangères se sont lancés dans des démarches d'ouvertures des données publiques. En France, la communauté d'agglomération Rennes Métropole et la bibliothèque des Champs Libres y font figures de pionnières. Comme l'explique Sir Tim Berners-Lee, l'*Open data* peut être « source de potentialités infinies pour la sphère publique comme pour la sphère privée, car si l'on partage des données sur le Web, des données publiques, des données scientifiques, des données citoyennes, quelles qu'elles soient, d'autres que nous saurons en tirer des créations merveilleuses que nous ne n'aurions jamais imaginées ». Les enjeux et les attentes de l'*Open data* sont donc multiples, tant dans les domaines économiques et politiques que démocratiques et scientifiques. Pourtant, en France, les données culturelles dont relèvent les données de bibliothèques bénéficient encore de « l'exception culturelle française » établie par le régime dérogatoire à la loi CADA de 1978. En conséquence, les bibliothèques ne sont soumises à aucune obligation légale d'ouverture des données. Pourtant, nombreuses sont les initiatives en bibliothèques qui gravitent autour de l'*Open data* et plus particulièrement des enjeux du *Linked Open data* comme en atteste les projets *data.bnf.fr* et OpenCat de la BnF et les publications de plusieurs bibliothèques de lecture publique comme la médiathèque des Champs Libres. En conclusion de ces observations, le présent mémoire s'attache à mettre en lumière la place et les enjeux des données produites par la bibliothèque et les institutions culturelles, au regard des démarches plus globales d'ouverture des données publiques. Il s'agit bien de comprendre et de montrer ce que les bibliothèques ont à gagner à s'approprier les logiques et les techniques du *Linked Open data,* aujourd'hui mais surtout demain.

Descripteurs :

Administration publique et Internet

Linguistique -- Informatique

Structures de données (informatique)

Exploration de données

(cc) BY-NC-ND

Abstract

Since the mid 2000s, several French and foreign governments, community or institutions launched into the opening of public data. In France, the Rennes Métropole urban community and the Champs Libres library are though of as pioneers. As Sir Tim Berners-Lee explained, Open data can be "a source of infinite potential for the public sector as in the private sphere, because if you share data on the Web, public data, scientific data , citizen data whatsoever, other than we will draw wonderful creations that we would never have imagined. " The challenges and expectations about Open data are multiple, both economic and political, as welle a democratic and scientific. However, in France, the cultural data to which the library data belong still pey on the "French cultural exception" established by a derogation from the CADA law of 1978. As a result, libraries are not forced by the law to open their data. However, several intiatives alradey exist involving library data. Most of those projects, such as data.bnf.fr or OpenCat are more specifically based on Linked Open Data technologies. Following these observations, this paper seeks to highlight the role and challenges of opening the data produced by libraries and other cultural institutions, facing the more global process of public data opening. The point id to understand and to demonstrate what libraries have to win by launching into the logic and the technologies of Linked Open Data, today, but especially tomorrow.

Keywords

Internet in public administration

Computational linguistics

Data structures (Computer science)

Data minin

Droits d'auteurs

Cette création est mise à disposition selon le Contrat : « **Paternité-Pas d'Utilisation Commerciale-Pas de Modification 2.0 France** » disponible en ligne *http://creativecommons.org/licenses/by-nc-nd/3.0/deed.fr* ou par courrier postal à Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.

BY-NC-ND

Sommaire

Sigles et abréviations **9**

Introduction **11**
Annonce du plan **13**

L'*Open data* : aux origines du mouvement **15**
Les motivations de la démarche **15**
L'argument technique : un contexte technologique favorable 15
L'argument démocratique 17
L'argument économique 21
Un *Open data* des territoires **24**
Des initiatives locales … 24
… Aux démarches nationales 26
Un avenir encore incertain **29**
Un cadre juridique en chantier 29
Quel schéma pour l'*Open data* ? 32

Bibliothèques : les bons élèves de l'*Open data* **35**
L'expérience de la bibliothèque Champs Libres à Rennes **35**
L'ouverture des données : une demande de la collectivité 35
Stimuler la réutilisation : le recours à des tiers médiateurs 37
Quelles réutilisations ? Quels réutilisateurs ? 39
Des projets qui fleurissent… **40**
La Bibliothèque Nationale Universitaire (BNU) de Strasbourg 41
Le réseau des médiathèques de Toulouse 41
La bibliothèque de Saint-Etienne 42
La Bibliothèque départementale de prêt (BDP) de Saône-et-Loire 43
Le projet Data.bnf.fr 43
A l'étranger 46
Des prédispositions et un savoir faire à partager **49**
Partager des données normées et structurées 49
Des données faciles à ouvrir 51
Les bibliothèques médiateur de l'*Open data* 51

***Open data* et *Linked Open data* : quelles promesses pour les bibliothèques ?** .. **53**
Promouvoir la réutilisation des données publiques culturelles **53**
Quelles sont les données publiques culturelles ? 54
Revenir sur l'exception culturelle française 55
Quelles conséquences économiques pour le secteur culturel ? 57
Améliorer l'offre de services **59**
Améliorer la recherche des utilisateurs 59
Améliorer son image et sa notoriété 61
Enrichir les catalogues **61**
Mutualiser des acquis, agréger des ressources 61
Pérenniser l'information 63
Le projet OpenCat de la BnF : un catalogue ouvert 64
Vers une révision des pratiques et des outils professionnels **65**
Faire évoluer les normes bibliographiques 65
Vers une rénovation du marché logiciel 65

Conclusion **68**

(cc) BY-NC-ND

Sources **73**

Bibliographie **75**

Généralités 75
Bibliothèques 76
Collectivité 77
Economie - Entreprise 78
Aspects juridiques 79
Aspects techniques – Web de données 80
International 81
Culture 81
Webographie 82

Table des annexes **85**

Glossaire **103**

Index **105**

Table des matières **109**

(cc) BY-NC-ND

Sigles et abréviations

AACR *Anglo American Cataloguing Rule*

ABES Agence Bibliographique de l'Enseignement Supérieur

API *Application Programming Interface*

APIE Agence du patrimoine immatériel de l'État

ARK *Archival Resource Key*

BnF Bibliothèque nationale de France

BNU Bibliothèque Nationale Universitaire

CADA Commission d'Accès aux Documents Administratifs

CC-BY-SA *Creative Commons Attribution Share Alike*

CCFR Catalogue Collectif de Frabne

CC0 *Creative Commons Zero*

CENL *Conference of European National Library*

CISI Comité interministériel pour la société de l'information

CMU Couverture Maladie Universelle

CNIL Commission Nationale Informatique et Liberté

CNN Conseil National du Numérique

CNRS Centre Nationale de la Recherche Scientifique

COEPIA Conseil d'Orientation de l'Édition Publique et de l'Information Administrative

CSV *Comma-Separated Values*

DADVSI Droit d'Auteur et aux Droits Voisins dans la Société de l'Information

DCAT *Data Catalog Vocabulary*

EPIC Etablissement Public Industriel et Commercial

FING Fondation Internet Nouvelle Génération

FRAP Fondation pour la recherche sur les administrations et les politiques publiques

FRBR *Functional Requirements of Bibliographic Records*

HTML *HyperText Mark-Up Language*

IFLA International Federation of Library Associations and Institutions

IGN Institut National de l'Information Géographique et Forestière

INIST L'Institut de l'Information Scientifique et Technique

ISBD *International standart for bibliographic description*

JDN Journal du Net

(cc) BY-NC-ND

LCEN Loi pour la Confiance dans l'Economie Numérique

MARC *Machine Redeable Cataloguing*

OCLC *Online Computer Library Center*

ODBL Open Database Licence

ODC-BY *Open Data Commons Attribution*

OKF *Open Knowledge Foundation*

OOP Organisations d'Observation des Parlements

PAGSI Le programme d'action gouvernemental pour la société de l'information

PRISME Pôle de ressources et d'information sur le monde de l'entreprise

RDA *Resources : Description et Access*

RDF *Resource Description Framework*

RMN Réunion des Musées Nationaux

SIGB Système Intégré de Gestion de Bibliothèque

SKOS *Simple Knowledge Organization System*

SNCF Société Nationale des Chemins de Fer Français

SUDOC Système Universitaire de Documentation

TIC Technologies de l'information et de la communication

URI *Uniform Resource Identifier*

XML *Extensible Mark-Up Language*

W3C *World Wide Web Consortium*

(cc) BY-NC-ND

INTRODUCTION

La nomination, le 15 mai 2012 de Jean-Marc Ayrault à la tête du gouvernement français a permis de rassurer les partisans français du mouvement d'ouverture des données publiques, entrepris sous la présidence Sarkozy, sur l'avenir de l'*Open data*. Ancien maire de Nantes et président de la communauté d'agglomération Nantes Métropole, Jean-Marc Ayrault s'était déjà prononcé en faveur de l'*Open data* avec l'inauguration le 21 novembre 2011 du portail d'accès public *data.nantes.fr*[1]. Dès lors, l'annonce, en apparence contradictoire, de l'abrogation de la mission Etalab et du départ de Séverin Naudet ont suscité certaines incompréhensions et inquiétudes auprès du public averti comme en témoignent quelques titres phares : « Etalab dissout, vers la fin de l'*Open data* à la française ?[2] » dans le Journal du Net (JDN), « Et si c'était la faim de l'Open data ?[3] » de LiberTic ou encore « Vol noir de corbeaux sur l'*Open data* français[4] » d'OWNI pour ne citer que quelques exemples. Dans ce contexte, comment ne pas s'interroger sur la pertinence de ce mémoire ? Pourquoi parler d'ouverture des données de bibliothèques, au moment où les démarches concernant celles contraintes à l'ouverture par la loi, semblent vivre leurs dernières heures ?

En premier lieu parce que le décret n°2012-1198[5] ne prévoit finalement que le rattachement de la mission Etalab au nouveau secrétariat général pour la modernisation de l'action publique[6]. Par ailleurs, parce qu'indépendamment des aspects légaux, les bibliothèques ont beaucoup plus à gagner de l'*Open data* qu'il n'y paraît, même s'il est vrai qu'avec des titres tel que « *Open data* – Où l'on apprend que « Le journal de Mickey » est la star des bibliothèques[7] » publié dans de la presse nationale comme le blog du Monde, il est parfois difficile de rendre compte du sérieux de la démarche *Open data* en bibliothèque et de la plus value scientifique qu'elle peut apporter pour le monde de la recherche, de l'information et de la documentation. Cette brève revue de presse doit donc être relativisée mais chacun de ces titres restent néanmoins révélateur de la méconnaissance et des incertitudes qui continuent d'entourer l'*Open data* et montrent à quel point le mouvement peine encore à trouver ses marques.

Comment expliquer ce phénomène ?

---

[1] http://www.entreprenantes.com/magazine/192-nantes-ouverture-des-donnees-le-point-sur-le-lancement.html (consulté le 18/12/12)

[2] http://www.journaldunet.com/ebusiness/le-net/suppression-etalab-1012.shtml (consulté le 18 décembre 2012)

[3] http://owni.fr/2012/11/06/et-si-cetait-la-faim-de-l-open-data/ (consulté le 18 décembre 2012)

[4] http://owni.fr/2012/11/02/vol-noir-de-corbeaux-sur-lopen-data-francais/ (consulté le 18 décembre 2012)

[5] Décret n° 2012-1198 du 30 octobre 2012 portant création du secrétariat général pour la modernisation de l'action publique, JORF n°0254 du 31 octobre 2012. http://www.legifrance.gouv.fr/affichTexte.do;jsessionid=?cidTexte=JORFTEXT000026557680&dateTexte=&oldAction=rechJO&categorieLien=id

[6] http://fr.scribd.com/doc/111681963/Opendata-%E2%80%93-le-Gouvernement-confirme-la-poursuite-des-demarches-d%E2%80%99ouverture-des-donnees-publiques. (consulté le 18/12/12) : La tête de la mission Etalab a finalement été confiée à Henri Verdier, ancien président du Pôle de compétitivité Cap Digital. Pour faire taire les inquiétudes, le chef du gouvernement en a profité pour rappeler qu'il souhaitait que la France « *passe à la vitesse supérieure, un an après la mise en ligne du portail data.gouv.fr* ». et que selon lui, Henri Verdier était « *tout à fait habilité pour conduire cette politique (...). C'est une mission exigeante, ambitieuse, mais elle est nécessaire* ».

[7] http://bigbrowser.blog.lemonde.fr/2012/12/13/open-data-ou-lon-apprend-que-le-journal-de-mickey-est-la-star-des-bibliotheques/

(cc) BY-NC-ND

En premier lieu, il semble important d'insister sur la confusion qui règne autour du terme « *Open data* », pourtant strictement équivalent à celui « d'ouverture des données publiques », plus souvent employé par les institutions françaises. Mais au-delà, d'un simple problème de traduction, l'*Open data* fait appel à des concepts dont les réalités et les enjeux n'ont que rarement été définis avec clarté. Dans le monde de la documentation et particulièrement en bibliothèque, l'amalgame est particulièrement fréquent, entre les concepts *d'Open access*, *d'Open source*, *d'Open content*, et *d'Open data*. Tous participent bien d'une idée d'ouverture et reposent en général sur des logiques juridiques communes, néanmoins la nature et la vocation des informations traitées nécessitent quelques nuances. Ainsi, la confusion la plus fréquente est faite entre « *Open access* » et « *Open data* », or il convient de bien les différencier : l'*Open access* concerne le contenu rédactionnel ou iconographique comme des documents numérisés, des photographies, des dossiers pédagogiques etc.-, qui ne peut s'appréhender que sous la forme d'un bloc, tandis que l'*Open data* propose des données brutes et/ou chiffrées comme des statistiques, des données de fréquentation, de budget etc. La distinction est parfois ténue, notamment dans le secteur culturel où il n'est pas rare de voir des projets mêler les deux sans distinction. A titre d'exemple, une bibliothèque numérique pourra proposer des documents numérisés issus du domaine public en *Open access* et les métadonnées descriptives qui leur sont associées en *Open data.*

Avant de poursuivre, attardons nous un peu sur la sémantique. Comprendre l'*Open data* nécessite de revenir sur trois notions clés dont il convient de bien délimiter le sens et la portée : « données », « publiques » et « ouvertes ». Si l'on s'appuie sur les définitions proposées par l'*Open Knowledge Foundation* ou l'*Open Government Working Group*[8], par « données » doivent être entendues des informations de première main, intègres, actualisées, dans l'idéal brutes[9], accessibles au format numérique, lisibles et exploitables par des machines. Pour ce qui est de l'adjectif « publiques », il permet de préciser et de délimiter le champs d'action politique de l'*Open data*. Bien que l'ensemble de la société, y compris les acteurs issus du secteur privé, puisse prendre part à l'ouverture des données, cette dernière n'a d'obligation légale que pour une partie du secteur public.

Ce périmètre est à l'origine défini par la loi CADA[10] de 1978 reprise par la circulaire du 14 février 1994 sur la diffusion des données publiques comme « les données collectées ou produites, dans le cadre de sa mission, par un service public sur des fonds publics » entendu par « […] l'Etat, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission [de service public] ». Enfin l'adjectif « ouvertes » vaut à la fois pour les caractéristiques techniques, juridiques et économiques de ces données. Elles doivent être publiées sous des formats non propriétaires (HTML, CSV, XML, RDF etc.), être idéalement gratuites et libres de droit, c'est-à-dire sous

---

[8] Fondé en 2007, ce groupe de travail rassemble une trentaine d'avocats et de juristes s'attelant à définir un ensemble de principes attachés aux données d'un gouvernement ouvert.

[9] Sir Tim Berners-Lee, « We want raw data now ! ».
http://www.ted.com/talks/tim_berners_lee_on_the_next_web.html

[10] Loi n°78-753 dite loi CADA du 17 juillet 1978 portant diverses mesures d'amélioration des relations entre l'administration et le public et diverses disposition d'ordre administratif, social et fiscal : elle pose le droit d'accès des citoyens à la plupart des documents administratifs, définis comme les documents produits ou reçus dans le cadre de leur mission de service public par l'Etat, les collectivités ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission.
http://www.legifrance.gouv.fr/affichTexte.do;jsessionid=40877996F2423CEEBB3898084FD7B38D.tpdjo07v_1?cidTexte=JORFTEXT000000339241&categorieLien=id

(cc) BY-NC-ND

des licences autorisant leur réutilisation à des fins commerciales ou non, avec pour unique réserve une mention de paternité (BY). Ce concept d' « ouverture » ou d' « *open* » a fait l'objet d'une définition approfondie[11] par l'*Open Knowledge Foundation* insistant particulièrement sur la non-discrimination des accès, l'absence de restriction technologique et l'ensemble des dispositifs juridiques permettant une réutilisation la plus large possible.

Le terme « *Open data »* apparaît, quant à lui, au milieu des années 1990 dans une publication du Comité sur les données géophysiques et environnementales du Conseil national de la Recherche aux États-Unis cherchant à démontrer et à promouvoir l'intérêt à constituer des bases de connaissances ouvertes dans le domaine scientifique[12]. Cette anecdote est intéressante car paradoxalement, avec le secteur culturel, l'enseignement et la recherche restent, encore aujourd'hui, les grands absents de l'*Open data* français. En conséquence, les données de bibliothèques ne sont pas à proprement parler des « données publiques ». Deux raisons majeures expliquent cet état de fait : en premier lieu l'important dispositif juridique sur la protection du droit d'auteur et la propriété intellectuelle qui régit ces domaines ; en second lieu, « l'exception culturelle » française. S'il n'existe pas encore d'obligation juridique quant à l'ouverture de ces données, ce n'est pas pour autant que les institutions productrices ou les utilisateurs ne gagneraient pas à ce qu'elles le soient, d'autant qu'une large partie des données ou métadonnées produites par les institutions culturelles ou scientifiques françaises ne sont ni sensibles ni soumises à droits d'auteur car tombées dans le domaine public.

Ce débat autour de l'ouverture des données scientifiques et culturelles rejoint, par ailleurs, d'autres problématiques rencontrées par les bibliothèques dans le secteur de la culture, de l'enseignement supérieur et de la recherche, à savoir le problème du financement de l'édition scientifique électronique et les réflexions autour de l'*Open Access Gold*[13]. Elles incitent de plus en plus d'institutions scientifiques et culturelles à envisager, de leur propre initiative, de nouveaux modèles techniques et économiques de diffusion de l'information et des connaissances. Leur intérêt pour l'*Open data* participent de ces mêmes réflexions. Soutenues dans cette démarche par le Ministère de la culture[14], plusieurs institutions culturelles se sont déjà lancées dans l'*Open data* et les bibliothèques font bien souvent figure de pionnières.

ANNONCE DU PLAN

En partant de réflexions inspirées par l'exemple de la médiathèque des Champs-Libres à Rennes, ce présent mémoire s'efforcera de comprendre les enjeux et de mettre en évidence l'intérêt que peuvent trouver les bibliothèques à

---

[11] *Open Knowledge Definition*, http://opendefinition.org/okd/.

[12] National Research Council, Committee on Geophysical and Environmental Data, *On full and open exchange of scientific data*, 1995, 21 p. : « International programs for global change research and environmental monitoring crucially depend on the principle of full and open exchange (i.e., data and information are made available without restriction, on a non-discriminatory basis, for no more than the cost of reproduction and distribution ».

[13] Bernard Teissier, « Se réapproprier et réorganiser les circuits de diffusion : l'opinion d'un chercheur sur les licences nationales » in *Les licences nationales*, Arabesque, avril-mai-juin 2012, p.11

[14] En janvier 2010, Frédéric Mitterrand a crée au sein de son ministère un département des programmes numériques, chargé d'encadrer et de promouvoir la réutilisation des données publiques culturelles, et, notamment, de mettre en œuvre, de façon échelonnée, la politique incitative prônée par le rapport Ory-Lavollée de 2009. Le rapport *Culture & Médias 2020* confirme cette engagement du Ministère en faveur d'un Open data culturel. Cf. p.45 « S'appuyer sur la réutilisation des données publiques culturelles ».

(cc) BY-NC-ND

s'impliquer dans des démarches d'*Open data*, à l'heure où l'avenir du mouvement, dans d'autres secteurs, continue de poser question. Nous commencerons par revenir sur les origines du mouvement par une présentation de ses grands enjeux et des premières expériences françaises et étrangères d'ouverture des données publiques. Cette première partie, d'apparence éloignée du cœur du sujet, nous semblait un préalable doublement nécessaire, à la fois en raison du caractère encore très méconnu de l'*Open data*, et mais aussi pour replacer les initiatives bibliothéconomiques dans leur contexte général. Dans un second temps, nous livrerons un état des lieux de l'*Open data* en bibliothèque avec un focus particulier sur l'expérience des Champs-Libres afin d'en mettre en évidence les freins mais surtout les perspectives qu'il pourrait offrir aux bibliothèques à travers ce mouvement.

(cc) BY-NC-ND

L'*OPEN DATA* : AUX ORIGINES DU MOUVEMENT

Comme son nom l'indique l'*Open data* n'est pas né en France. Bien que celle-ci France ait une longue tradition d'accès aux données publiques, c'est du côté des pays anglo-saxons que l'on trouve les racines de l'*Open data*. Néanmoins, il est intéressant de constater la rapidité avec laquelle le concept a su s'exporter. Par effet domino, l'*Open data* touche de plus en plus de pays, reproduisant en général le même schéma pour agiter les mêmes enjeux.

LES MOTIVATIONS DE LA DEMARCHE

L'*Open data* est un mouvement complexe, à la croisée de notions et d'enjeux multiples, à la fois techniques, juridiques, économiques ou encore démocratiques. Tous ne sont pas nécessairement convergents, malgré une démarche commune. On peut « entrer en *Open data* » pour des motifs divers, c'est ce qui fait la richesse du mouvement mais également ce qui le rend si difficile à appréhender.

L'argument technique : un contexte technologique favorable

Les enjeux premiers de l'*Open data* ne sont pas *stricto sensu* informatiques ou techniques, mais son développement reste étroitement lié au contexte d'évolution des nouvelles technologies de l'information et de la communication. Difficile, pour ne pas dire impossible, de concevoir l'*Open data* indépendamment de l'explosion, au cours des dernières décennies, d'Internet, de la téléphonie mobile et de l'*Open source*. De nouveaux usages sont apparus quant à la consommation d'informations liées au territoire, accompagnés d'une volonté participative des citoyens que l'on retrouve depuis maintenant un moment dans le domaine du web (web 2.0, wiki, forum, logiciels libres etc.). C'est l'analyse proposée par le Conseil du numérique dans son rapport de 2009 sur l'*Open data* :

> Les évolutions technologiques de ces quinze dernières années devraient permettre, en théorie, une mise à disposition du plus grand nombre, gratuite, de toutes les données publiques. En conséquence, elles ont considérablement accru dans le même temps les attentes, voire les exigences de la société civile. En particulier, celle de pouvoir diffuser et réutiliser ces informations[15].

Sans revenir sur l'historique du Web, il convient néanmoins de rappeler les changements de paradigmes sociaux-économiques qu'il a engendrés comme la capacité de toucher un public international instantanément et sans surcoût de distribution. En offrant la possibilité de publier un service à partir d'un outil voire d'un pays autre que celui d'origine de l'information, il délie cette dernière des règles de la territorialité. Celle-ci est désormais accessible partout, tout le temps, depuis son ordinateur, son salon, son téléphone, dans la rue, et à toute heure du jour et de la nuit.

[15] Avis n° 12 du Conseil national du numérique relatif à l'ouverture des données publiques (« Open data »), 5 juin 2012, 23 p.

(cc) BY-NC-ND

En France, une première prise en compte de ce contexte technologique a été formulée à la fin des années 1990 par Lionel Jospin dans son discours de Hourtin. Il y fait le lien entre transparence, communication et nouvelles technologies de l'Internet :

> Pour répondre à l'exigence démocratique de transparence de l'Etat, un accès plus aisé à l'information publique est par ailleurs indispensable. Depuis près de 20 ans, l'accès aux documents administratifs est devenu une véritable liberté publique ; aujourd'hui, la technologie facilite les conditions de leur diffusion. Les données publiques essentielles doivent désormais pouvoir être accessibles à tous gratuitement sur internet. […] J'entends privilégier une conception ambitieuse du droit de l'information du citoyen : la diffusion internationale de nos documents publics doit à cet effet être favorisé[16].

Cette volonté politique prend forme à travers le lancement du Programme d'action du gouvernement pour la Société de l'information (PAGSI) confié au Comité interministériel pour la société de l'information (CISI). Rendu public en janvier 1998, il vise à favoriser le développement de l'industrie et du marché de l'information grâce à l'utilisation des nouvelles technologies de la communication. Pour la première fois, l'apport des technologies y est pensé de concert pour l'ensemble des ministères. Le plan définit 6 axes pour « l'entrée de la France dans la société de l'information » comprenant déjà les grands objectifs que s'efforcent de poursuivre aujourd'hui l'*Open data* : renforcer la présence de la culture française et francophone sur le web notamment par la numérisation et la diffusion du patrimoine, promouvoir un État plus transparent et plus efficace via une réorganisation et une meilleure diffusion en ligne des informations publiques, favoriser l'appropriation par le secteur privé des apports des TIC sur le plan commercial, la création d'emploi et la compétitivité internationale.

Concrètement, le plan prévoyait la migration des services Minitel de l'Etat vers l'Internet, la diffusion gratuite des données publiques essentielles comme les « grands textes de notre droit, l'information administrative du public, les principaux documents publics, des données culturelles essentielles et le renvoi à une personnalité indépendante l'élaboration d'une doctrine claire et les conditions de la tarification des données publiques ». Autant d'éléments auxquels se confrontent encore aujourd'hui les démarches *Open data*. Par ailleurs, le PAGSI évoquait déjà la nécessité d'adapter le cadre législatif aux nouveaux médias annonçant quelques années plus tard la loi de 2004 pour la confiance dans l'économie numérique[17] (LCEN) puis la loi de 2006 relative au droit d'auteur et aux droits voisins dans la société de l'information (DADVSI).

La profusion soudaine de données accessibles via l'internet peut néanmoins engendrer de réelles difficultés pour l'utilisateur dans sa recherche de l'information ou du service adéquat. L'explosion de l'offre et des acteurs présente un risque de « noyade » pour l'internaute dans une masse d'informations sans intérêt ou non pertinentes et le risque d'un gommage des distinctions jusqu'ici assez claires entre information publique officielle, information privée, information

---

[16] Lionel Jospin, Discours de Hourtin du 25 août 1997, « Préparer l'entrée de la France dans la société de l'information », Université de la Communication.

[17] La loi pour la confiance dans l'économie numérique, n° 2004-575 du 21 juin 2004, abrégée sous le sigle LCEN, est une loi française sur le droit de l'Internet, transposant la directive européenne 2000/31/CE du 8 juin 2000 sur le commerce électronique et certaines dispositions de la directive du 12 juillet 2002 sur la protection de la vie privée dans le secteur des communications électroniques.

(cc) BY-NC-ND

spécialisée et information grand public. Comme dans toute démarche éditoriale, l'*Open data* nécessite de réfléchir au public auquel s'adressent les données. Une démarche souvent négligée du fait de la facilité et de la rapidité à publier sur internet. L'introduction du web dans les pratiques de diffusion pose pourtant plus que jamais la question de l'adéquation entre la nature de l'information et sa réception par le public. La volonté présente dans l'*Open data* de regrouper les informations au sein de plateformes publiques spécialisées, clairement identifiées et labellisées participe d'un effort intéressant de clarification et de simplification de la recherche pour l'internaute.

Cependant, parallèlement à cet effort de transparence et de médiation, le développement de l'internet suscite également des craintes quant à la nature des informations partagées. La sécurité des usagers et la protection de leurs données personnelles sont des points de préoccupation récurrents des internautes et des institutions productrices. Le concept du « *cloud computing* » et l'internationalisation des supports de l'information entrainent des conflits entre différents systèmes juridiques, que ce soit en matière de protection des données personnelles, ou en matière de droit d'auteur, ou de responsabilité. Il est tout aussi important pour la diplomatie française d'être en mesure de défendre sa conception du partage des données et de parvenir à trouver un terrain d'entente entre ses positions internes et les réglementations en cours à l'échelle européenne et mondiale.

La loi relative à l'informatique, aux fichiers et aux libertés du 6 janvier 1978, dite loi CNIL[18], constitue un premier effort de prise en compte juridique de l'impact des technologies de l'informatique puis de l'internet sur la société d'information. Elle prévoit un certain nombre de garde-fous pour protéger l'individu des dangers liés à la multiplication de fichiers contenant des informations nominatives. Les responsables privés ou publics de données nominatives sont soumis à des obligations sévères concernant la déclaration des fichiers, la collecte, la conservation et la sécurité des informations, leur communication et leur commercialisation dans le respect des droits de la personne.

L'argument démocratique

Le lancement en décembre 2006 du site *Wikileaks*, représenté depuis 2007 par le très controversé Julian Assange, marque un tournant dans les relations qu'entretiennent pouvoirs publiques et Internet. La révélation en 2010 de milliers de documents militaires américains secrets afférents à d'éventuels crimes de guerre lors du conflit afghan parfait la notoriété de ce site dont les actions prennent une résonnance mondial. En divulguant de manière anonyme, non identifiable et sécurisée, des documents sensibles qui auraient été volontairement cachés par les pouvoirs politiques et militaires, le site pointe du doigt le manque de transparence des gouvernements et de leurs actions. Mais il pose plus généralement la question de la gestion de l'information publique tant sur le plan pratique – comment stocker, sécuriser et diffuser l'information, sur quels supports, par quels vecteurs ? – que sur le plan éthique et politique – qui maitrise réellement l'information, les

[18] Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, dite loi CNIL, réformée par la loi du 6 août 2004, transposant, de façon libre, la directive européenne du 24 octobre 1995 sur la protection des données à caractère personnel (dir. 95/46/CE). http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006068624&dateTexte=vig (consulté le 04/12/2012).

(cc) BY-NC-ND

lois sur les secrets d'Etat, les délais légaux de communication, la protection des individus sont elles encore pertinentes, efficaces, souhaitables et surtout adaptées aux technologies actuelles de création et de diffusion de l'information ?

Les réponses à ces questions, pour autant qu'elles puissent exister avec certitude, dépassent largement le cadre de cette étude, mais elles constituent un arrière plan qu'il convient d'avoir à l'esprit pour comprendre l'explosion des débats politiques, juridiques et éthiques qui animent aujourd'hui nos sociétés de l'information. L'ampleur des débats et l'envolée des mouvements relatifs à l'ouverture des données publiques, constatés à l'échelle mondiale au cours des trois dernières années, s'imposent comme des suggestions de réponse à ces interrogations.

L'engagement des gouvernements pour la transparence

En France comme à l'étranger, la transparence est une des attentes fortes des mouvements citoyens. C'est aussi une démarche de qualité de plus en plus revendiquée par les pouvoirs publics, à l'instar des modèles de gouvernance scandinaves[19]. Très liée aux débats politiques anglo-saxons autour de la notion d' « *accountability*[20] », la transparence s'impose comme une valeur centrale des démocraties modernes. Preuve en est le Mémorandum sur la transparence et l'ouverture du gouvernement prononcé à l'intention des Ministres et des directeurs d'agence, par le président Barack Obama au lendemain de son investiture, le 21 janvier 2009 :

> Je m'engage à promouvoir un niveau d'ouverture sans précédent au gouvernement. Nous travaillerons ensemble à raffermir la confiance publique et à établir un système fondé sur la transparence, la participation publique et la collaboration. L'ouverture renforcera notre démocratie et favorisera l'efficience et l'efficacité de notre gouvernement[21].

Ce discours se concrétise rapidement par la publication d'une *Directive pour un gouvernement ouvert*[22] rappelant les trois grands piliers de la coopération entre l'État et les citoyens : la transparence du gouvernement, la participation et la collaboration. Le 5 mars 2009, Obama crée un poste de responsable fédéral des systèmes d'information de l'administration confié à Vivek Kundra[23]. Ce poste, créé sur mesure, s'accompagne de la mission de mettre en place des technologies permettant de concrétiser les promesses politiques du président. Cela se traduit par le lancement, du projet de portail fédéral *data.gov* mis en ligne dans sa première

---

[19] Cet article du journaliste Rémi Nielson paru en juin 2010 dans le Monde diplomatique propose une analyse intéressante du modèle du gouvernance scandinave, en mettant en lumière les bienfaits de la transparence mais également ses limites.
http://www.monde-diplomatique.fr/2010/06/NILSEN/19207

[20] « être redevable de… » : exprime l'obligation des pouvoirs publics de rendre des comptes à leurs administrés.

[21] Barack Obama, *Memorandum on Transparency and Open governement* : « My administration is comitted to creating an unprecedented level of openness in Government. We will work together to ensure the public trust and establish a system of transparency, public participation, and collaboration. Openness will strengthen our democracy and promote eficiency and effectiveness in Government ».
http://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment (consulté le 27 novembre 2012)

[22] *Open Government Directive*
http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf
(consulté le 04/07/2012)

[23] Ancien responsable des systèmes d'information du district de Columbia de la ville de Washington, il est également à l'origine, en 2008, du concours d'applications *Apps for Democracy*, qui avait rencontré un succès inattendu. http://www.appsfordemocracy.org/

(cc) BY-NC-ND

version en ligne le 21 mai 2009, soit seulement 120 jours après la signature du Mémorandum. Le site ne comporte alors que 47 jeux de données, récoltées auprès des DSI[24] de toutes les agences fédérales américaines et pouvant être facilement mises en ligne. En juillet 2011, le portail hébergeait plus de 3000 jeux de données alimentés par près de 400 fonctionnaires.

L'argument démocratique de l'*Open data* est également repris en Grande-Bretagne lors des élections parlementaires de mai 2010, en réaction au scandale des notes de frais des parlementaires qui avait éclaté à l'été 2009. Transparence et ouverture des données publiques fleurissent dans tous les programmes. Bien qu'antérieur à son élection, le portail *data.gov.uk* mis au point par Tim Berners-Lee et Nigel Shadbolt en septembre 2009, est repris à son compte par le premier ministre David Cameron, fraichement élu en 2010. Ce portail était initialement conçu comme un soutien à l'innovation et à la création de nouveaux services aux citoyens, mais le Premier Ministre l'infléchit vers un outil de transparence de l'action de l'Etat. Il nomme un *Public Sector Transparency Board* incluant les co-fondateurs historiques et affirme sa volonté de construire, par ce biais, « le gouvernement le plus transparent du monde ».

L'initiative britannique doit beaucoup à la conviction et aux efforts de Tim Berners-Lee en faveur des « *open* », « *raw* » and « *linked* » *data* ainsi qu'aux travaux de l'*Open Knowledge Foundation*. Grâce aux initiatives de cette dernière plusieurs avancées en faveur de la transparence démocratique ont déjà été faites comme en atteste la publication récente du 29 août 2012, de la *Déclaration pour l'Ouverture et la Transparence Parlementaire*[25]. Soutenue par 76 organisations[26] issus de 53 pays, cette déclaration se veut un appel des Organisations d'Observation des Parlements (OOP) issues de la société civile aux parlements nationaux et aux organismes législatifs locaux, régionaux et transnationaux pour renforcer l'ouverture des institutions et l'engagement citoyen dans le travail parlementaire. Elle promeut *l'Open data* comme un vecteur électronique d'accès à l'information et à la transparence parlementaire.

Il convient de rester prudent quant à cet argument démocratique car l'ouverture des données publiques reste la manifestation d'un choix sur lequel il convient de garder l'esprit critique. Pour cela, Sandra Moatti[27] suggère de consulter la « liste des données qui ne seront jamais publiées[28] », dressée par le *Think trank*[29] de la FRAP[30]. On y trouve encore le salaire des hauts fonctionnaires, les détenteurs de la dette française (70% étrangers), le taux d'absentéisme des

---

[24] DSI : Direction des systèmes informatiques

[25] http://openingparliament.s3.amazonaws.com/docs/declaration/1.0/francais.pdf (consulté le 22/11/2012)

[26] Trois organisations françaises ont participé à cette déclaration : ClubJade, Transparence international France et l'association Regards Citoyens. La liste complète des organisations est disponible à l'adresse suivante : www.openingparliament.org/organizations

[27] Sandra Moatti, *op. cit.*

[28] http://www.ifrap.org/Les-donnees-publiques-qui-ne-sont-abusivement-jamais-publiees,12411.html (consulté le 09/09/2012)

[29] Un *Think tank* ou laboratoire d'idée est une institution de droit privé à but non lucratif et généralement politiquement indépendante regroupant des experts qui produisent des données et émettent des propositions dans le domaine des politiques publiques. http://fr.wikipedia.org/wiki/Think_tank (consulté le 09/09/2012).

[30] L'IFRAP est la Fondation pour la recherche sur les administrations et les politiques publiques. Fondée en 1985 par Henri Zimmern, elle se donne pour but « d'effectuer des études et des recherches scientifiques sur l'efficacité des politiques publiques, notamment celles visant la recherche du plein emploi et le développement économique, de faire connaître le fruit de ces études à l'opinion publique, de proposer des mesures d'amélioration et de mener toutes les actions en vue de la mise en œuvre par le Gouvernement et le Parlement des mesures proposées ». Elle est reconnue d'utilité publique par décret ministériel depuis novembre 2009.
http://www.ifrap.org/La-Fondation-iFRAP,3.html (consulté le 09/09/2012)

(cc) BY-NC-ND

enseignants et des autres fonctionnaires, le coût de la CMU de base ou encore le montant de l'argent du paritarisme qui va chaque année aux syndicats, pour ne citer que quelques exemples. De ce côté là, les initiatives anglo-saxones semblent avoir poussé la logique plus loin puisque le gouvernement britannique a décidé de publier indifféremment les marchés publics de plus de 10 000 livres à côté des noms et des salaires annuels de tous les hauts fonctionnaires ou encore l'ensemble des crimes, délits et incivilités géolocalisés sur le territoire britannique via le site *Police.uk*.

Un débat a également été lancé aux Etats-Unis par Lawrence Lessig fondateur du Centre pour l'internet et la Société à l'Université de Stanford et créateur des licences *Creative Commons* sur les vertus mais aussi les limites de la transparence[31]. Daniel Kaplan, délégué général de la Fing invite à rester vigilant quant aux risques et aux méfaits d'une « transparence totale ».

> [La transparence] peut également produire des conséquences problématiques. On peut imaginer qu'au lieu de s'engager dans des discussions ouvertes et constructives sur des sujets d'intérêt général, les groupes d'intérêt s'en servent pour contester la moindre décision, la moindre ligne de dépense publique. Au lieu d'aboutir à des décisions plus sages et plus largement soutenues, la transparence totale dresse les intérêts particuliers les uns contre les autres, inhibe la prise de décision et réduit la confiance. La nouvelle visibilité des données invite même à les trafiquer à la source, pour faire en sorte que même des analyses indépendantes, fondées sur des moyens informatiques massifs, produisent les conclusions qu'attendent les producteurs de données. L'ouverture des données publiques réduirait ainsi la fiabilité des données publiques[32].

Il illustre son propos avec le cas du Bangalore où la mise en accès de l'information sur la propriété foncière aurait principalement permis aux personnes les plus aisées et à quelques entreprises d'évincer les populations pauvres de leurs terres et de les marginaliser.

Pour une modernisation des modèles de gouvernances

Derrière l'enjeu de transparence se cache un objectif plus profond et plus structurel de la démarche *Open data* : la rénovation des modèles de gouvernances, dont la transparence n'est qu'une manifestation. L'*Open data* devient un moyen de faire connaître les méthodes de décision et de gestion des politiques publiques. En décembre 2007, un groupe de travail[33] composé d'une trentaine de partisans du principe de « gouvernement ouvert » coordonnés par Tim O'Reilly et Carl Malamud s'est retrouvé en Californie à Sebastopol pour définir les conditions d'un gouvernement ouvert et démontrer en quoi l'*Open data* gouvernemental est essentiel à la démocratie. Internet étant considéré comme le nouvel espace public du monde moderne, il est de l'intérêt des gouvernements de saisir cette opportunité pour mieux cerner les besoins des citoyens et les inciter à participer plus activement à l'action publique. En reprenant l'idée du « Web 2.0 », Tim O'Reilly,

---

[31] Hubert Guillaud, La transparence a t-elle des limites ?, Internet Actu.net, Fing, 20 novembre 2009 http://www.internetactu.net/2009/10/20/la-transparence-a-t-elle-des-limites/ (consulté le 10/09/2012)
[32] Daniel Kaplan, L'ouverture des données publiques et après ?, OWNI, 18 janvier 2011 http://owni.fr/2011/01/18/louverture-des-donnees-publiques-et-apres/ (consulté le 09/09/2012)
[33] Composition et agenda du *Open governement working group,* sponsorisé par la Sunlight Foundation, Google et Yahoo. https://public.resource.org/open_government_meeting.html (consulté le 03/12/12)

(cc) BY-NC-ND

suggère la création d'une plateforme d'innovation ouverte qui permettrait à chaque citoyen de contribuer à résoudre des problèmes collectifs en faisant remonter l'information et l'expertise disséminée au sein de la société.

Cet appel au *Crowdsourcing* peut se révéler particulièrement utile lorsque des données publiques s'avèrent inexistantes, absentes, ou bien simplement incomplètes. Des associations citoyennes n'hésitent déjà pas à les produire elles-mêmes en croisant différentes sources ou en sollicitant la participation collective du public comme l'a fait l'Association Regards Citoyens pour créer une base de données sur les personnes auditionnées à l'Assemblée nationale dans le cadre d'une enquête sur le poids des *Lobbies*[34]. La foule devient la source et conforte l'idée que si les acteurs publics ne prennent pas l'initiative d'ouvrir leurs données d'autres décideront de le faire à leur place, comme en témoigne *Wikileaks*. En engageant une démarche *Open data*, l'Etat et les collectivités, voire les entreprises, acceptent de montrer leur cuisine, mais restent ainsi en mesure de choisir où, quand, quoi et de quelle manière elles le feront.

L'argument économique

Lorsque le 12 janvier 2011, Francis Maude annonce la création de la *UK Public data corporation* chargée d'arbitrer les mises en ligne des données publiques publiées via la plateforme *data.gov.uk*, il insiste sur l'enjeu économique de l'*Open data*, jusqu'ici relégué au second plan.

> Les données publiques soutiennent une part croissante de l'économie. Les technologies actuelles permettent à chacun d'utiliser et de réutiliser cette information de maintes façons nouvelles et innovantes. Le rôle du gouvernement est d'aider à maximiser les bénéfices que la société retire de ses développements. Aujourd'hui, beaucoup d'agences font face à un conflit entre les revenus issus de la vente de données disponibles et la possibilité de rendre ces données disponibles gratuitement pour être exploitées à des fins économiques ou sociales[35].

Comme le souligne le rapport Mandelkern[36], l'information constitue une denrée valorisable. En conséquence les données publiques sont également susceptibles, par leur diffusion, de générer des profits pour l'Etat, les collectivités, les établissements producteurs et l'ensemble des acteurs économiques qu'ils soient publics ou privés. Si l'on se place dans une logique commerciale, l'enjeu des données publiques est double : elles sont susceptibles de rapporter des profits pour les institutions publiques qui les vendent et d'en générer pour les acteurs publics ou privés qui les achètent. Or, c'est justement l'ensemble de ce système que se propose de remettre en cause l'*Open data*, car à l'inverse, une autre logique consiste à considérer que les données produites par les institutions le sont au titre de leurs missions de service public. Elles ont de fait été financées par l'impôt, et donc déjà payées. Tout l'enjeu de l'*Open data* est de montrer qu'un autre système,

---

[34] Les résultats de l'enquête sont communiqués en *Open data* via le site de l'association sous le lien suivant : http://www.regardscitoyens.org/transparence-france/etude-lobbying/téléchargement/ (mars 2011- consulté le 06/09/2012)

[35] The Cabinet Office, 12 janvier 2011, « *Public data Corporation to free up public data and drive innovation* » http://www.cabinetoffice.gov.uk/news/public-data-corporation-free-public-data-and-drive-innovation

[36] Dieudonné Mandelkern, Bertrand Marais (Du), *Diffusion des données publiques et révolution numérique : rapport de l'atelier*, « Des moyens nouveaux au service de la diffusion des données publiques », Paris, la Documentation française, 1999, 123 p.

(cc) BY-NC-ND

fondé sur les valeurs de transparence, démocratie et service public peut également être viable économiquement.

Ce nouveau modèle économique s'appuie sur plusieurs fondements de la théorie économique néo-classique comme le concept d'externalités positives[37] ou celui de concurrence pure et parfaite[38].

> En augmentant la quantité de données mises à disposition sur le marché, on se rapproche autant que possible du marché parfait, ce qui génère de la croissance, permet l'apparition de nouvelles activités économiques et au bout du compte des services de meilleurs qualité, de plus grande efficacité pour le public[39].

En tant qu'objet commercial, les données publiques possèdent des caractéristiques propres qui influent sur la nature de leur marché. En premier lieu, leur consommation ou leur usage par un agent n'en prive pas un autre car l'information fait partie des biens publics ou collectifs. En conséquence, leur circulation et leur réutilisation sont fluides et non exclusives ce qui rend le contrôle juridique ou technique de leur utilisation secondaire sur les marchés difficile. Par ailleurs, l'information joue un rôle non négligeable dans l'abaissement des coûts de transaction ce qui la rend utile mais non indispensable. Son élasticité-prix est donc très forte. Si les données sont vendues trop chères, les acteurs du marché s'en passent. En revanche, si leur marge de revenus fixes est limitée par cette élasticité, la qualité du bien public et les externalités positives qu'elles peuvent dégager sont bien supérieures et devraient, dans la logique *Open data*, suffire à promouvoir leur libre et gratuite mise à disposition.

En corroboration de ces principes, plusieurs études se sont attachées à conduire un raisonnement par l'absurde pour démontrer le coût qui pouvait résulter d'un défaut d'information. Le *US Geological survey*, équivalent de l'IGN français au début des années 1980 l'a appris à ses dépends. Les pertes économiques occasionnées par le défaut d'information lié à la fermeture en 1883, pour des motifs budgétaires, de la moitié des stations d'annonce de tempêtes autour des Grands Lacs aux USA en est un bon exemple[40]. Ces conclusions ont ainsi incité plusieurs pays, depuis la fin des années 1980, à ériger la diffusion des données publiques en vecteur d'une stratégie dynamique de développement de l'internet[41] et de soutien à l'économie[42].

En s'appuyant sur ce concept d'externalités, l'enjeu de l'ouverture des données publiques est de permettre à d'autres acteurs, publics ou privés, du marché national ou international, de réutiliser ces données à des fins commerciales ou non,

---

[37] Dans la théorie néo-classique, les externalités positives (ou économies externes) désignent les situations où un acteur est favorisé par l'action de tiers sans qu'il ait à payer.

[38] Les conditions de la concurrence pure et parfaite ont été explicitées en 1921 par l'économiste Frank Knight. Parmi elles se trouve la transparence de l'information qui suppose une information gratuite et immédiate de tous les agents sur tous les autres et sur tous les bien échangés.

[39] Entretien avec Richard Sterling, The Cabinet Office, dans Valérie Peugeot, *Les enjeux publics, économiques et citoyens de l'ouverture des données. L'expérience britannique*, 2010, p. 285.

[40] « The value of weather information services for nineteenh century Great Lake » shipping, E. Craft, American Economic Review, December 1998, p.1059

[41] Internet est le premier système qui offre, pour un prix marginal de diffusion tendant vers 0 et pour un investissement initial relativement faible, un véritable face à face personnalisé entre offreur et consommateur. Il offre ainsi l'avantage de permettre une diffusion géographiquement ciblée et plus adaptée aux besoins des utilisateurs.

[42] Rapport de B. Oudet, président de *Internet Society-France* pour le Commissariat général au Plan, Internet et administration à l'étranger, une source d'idées, septembre 1998
www.leibniz.imag.fr/PLAN

(cc) BY-NC-ND

comme source d'innovation. Dans son ouvrage *Comprendre l'ouverture des données publiques*, Simon Chignard rappelle que depuis 2003, dans le cadre de l'évaluation de la directive PSI, la Commission européenne s'est attachée à mesurer et estimer le marché de l'information publique. Plusieurs études ont participé de cette initiative comme l'étude MEPSIR[43] de 2006 qui annonçait des retombées de près de 30 milliards d'euros grâce à l'*Open data*. S'il est toujours difficile de mesurer le montant réel de ces retombées, il est certain qu'elles restent loin des chiffres annoncées. De plus l'enquête *Open Government Data Benchmark Study 2011*, réalisée aux Etats-Unis par Socrata révèle que seuls 3% des développeurs interrogés ont pour motivation de produire un service un jour rentable. Leur première motivation[44] étant que leur application puisse être utile au quotidien pour ceux qui l'utilisent. L'aspect lucratif ne vient, au mieux, que dans un second temps.

> Les horaires de bus (pour prendre un exemple d'une donnée qui est souvent réutilisée par les développeurs d'applications mobiles) ne sont quasiment jamais vendus. Ils sont proposés gratuitement au format papier, qu'il s'agisse des guides horaires ou même de ceux affichés dans les stations. De même pour les données statistiques de l'INSEE référencées sur le portail data.gouv.fr. Certaines données ouvertes localement, notamment géographiques, étaient parfois tarifées, mais en pratique elles étaient rarement vendues. C'est d'ailleurs le coût de la non-gratuité qui a joué en faveur de leur ouverture. Il ne s'agit pas pour autant de dire que ces données n'ont pas de valeur, mais plutôt que leur valeur n'est pas monétaire[45].

Du côté des entreprises privées, l'*Open data* est encore un mouvement à sens unique. Selon l'étude Bluenove[46], si certaines utilisent bien les données issues de l'*Open data*[47], peu nombreuses sont celles à pratiquer elles-mêmes une politique d'ouverture, exception faite de certaines entreprises à vocation de service public ou fonctionnant sur une base de capitaux publics comme la SNCF, Suez Environnement ou encore La Poste. Les entreprises attendent plus de l'*Open data* qu'elles ne sont prêtes à donner. Sans revenir en détail sur les cas de réutilisation des données publiques par le secteur privé[48], les premiers résultats de Bluenove permettent de mieux cerner les cas de participation de réutilisation des données publiques par les entreprises : aide à la publication de données, développement d'applications, génération de trafic internet afin de miser sur la publicité, explicitation de données grâce à des outils de visualisation, amélioration du son image, coup marketing, pour ne citer que les cas les plus fréquents.

S'il semble assez évident que le véritable enjeu économique de l'*Open data* se trouve pourtant du côté du secteur privé et particulièrement dans l'industrie et le commerce du Web, de ses services, produits et applications, ce n'est pas encore

---

[43] *Measuring European Public Sector Information Re-Use*.
[44] Pour plus de 40% des répondants.
[45] http://donneesouvertes.info/2012/03/12/ou-est-le-marche-de-lopen-data/ (consulté le 04/12/12)
[46] Les Transports du futur, *Livre blanc : Open data: quels enjeux, quelles opportunités pour les entreprises ?*, Bluenove, novembre 2011.
http://fr.slideshare.net/transportsdufutur/livre-blancbluenove?ref=http://lecercle.lesechos.fr/entreprises-marches/high-tech-medias/internet/221140403/livre-blanc-open-data-quels-enjeux-et-opport (consulté le 04/12/12)
[47] Comme l'explique Romain Wenz, responsable du projet *data.bnf.fr,* plusieurs entreprises privées issues du secteur du web ont utilisé le *dump* des données de *data.bnf.fr* pour tester et finaliser les paramétrages de leurs propres produits à partir d'un corpus de données homogènes, structurées et libres de droit donc plus facilement réutilisables que d'anciennes données clients.
[48] Cf. analyses et résultats exposés dans Bluenove, *op. cit.*

(cc) BY-NC-ND

dans ce secteur que la réutilisation des données publiques se montre la plus dynamique. Comme le souligne Simon Chignard, c'est encore du côté des collectifs citoyens et des communautés *Open source* que les initiatives sont les plus nombreuses. L'*Open data* reste étroitement liée à la notion de territoire et de localité.

UN *OPEN DATA* DES TERRITOIRES

Si les enjeux de l'*Open data* sont aujourd'hui surtout associés aux grands projets de plateformes étatiques, l'existence du mouvement est souvent bien antérieur à leur création. Aux Etats-Unis comme en France, le schéma de l'*Open data* est celui d'un mouvement ascendant, né de l'échelon local pour remonter vers des démarches nationales.

Des initiatives locales …

Aux Etats-Unis comme en France, les premières initiatives d'ouverture des données publiques remontent, au milieu des années 2000 et sont le fait de municipalités comme Washington, Chicago, San Francisco ou encore New-York. Un point de départ est donné en 2005 par le lancement des API du service *Google Maps* qui favorise le lancement de nombreuses applications et *mashups* s'appuyant sur des données de géolocalisation comme le service *ChicagoCrime.org* du journaliste et programmateur Adrian Holovaty qui sur une carte les données relatives aux crimes et délits perpétrés dans la ville, récupérées auprès de la municipalité. A travers son blog, la journaliste américaine Emilie Badger fournit un aperçu intéressant des meilleures initiatives citadines américaines en matière d'*Open data*[49]. Un peu plus tardif en France, l'*Open data* semble néanmoins répondre au même schéma de progression. Les premières expériences d'ouverture des données publiques ont été le fait de villes ou de communautés d'agglomération. Simon Chignard attribue notamment ce phénomène au simple fait que depuis les lois de décentralisation de 1982-1983, une grosse majorité des données publiques sont produites à l'échelle territoriale. Par ailleurs, selon les conclusions du Conseil national du numérique, c'est à l'échelle locale que les politiques d'*Open data* sont les plus pertinentes et les plus dynamiques, car elles touchent au plus près les besoins et les attentes des citoyens.

> Cette dimension participative [de l'Open data] est particulièrement pertinente au niveau locale : citoyens et associations mais aussi entrepreneurs pourraient développer des services innovants et contribuer à améliorer la qualité de vie de la population[50].

La ville de Rennes et la communauté d'agglomération de Rennes Métropole font figurent de pionnière dans ce domaine. Depuis mars 2010, plus de 188 jeux de données issus de plus de 1500 organismes publics ou associatifs[51] de la ville et l'agglomération ont été mise à disposition via la plateforme *data.rennes-metropole.fr* pour permettre la création d'une soixantaine de services disponibles sur Internet via fixe ou mobile. A l'image des choix de l'agglomération en matière

---

[49] http://www.fhimt.com/2012/03/12/le-top-de-lopen-data-municipale-aux-etats-unis/ (consulté le 04/12/12)
[50] Conseil national du numérique (CNNum), Avis n°12 relatif à l'ouverture des données publiques (« Open data »), 5 juin 2012, 23 p.
[51] http://www.data.rennes-metropole.fr/notre-demarche/ (consulté le 03/09/2012)

(cc) BY-NC-ND

de politique d'investissements, ils concernent majoritairement le réseau des transports (STAR-Kéolis), l'accessibilité, l'environnement, la culture ou un ensemble d'informations pratiques géolocalisées Conformément à la logique *Open data* et au respect des lois CNIL, ces données ne sont pas nominatives et sont en strict rapport avec le territoire rennais. Cette initiative permet à Rennes Métropole d'exploiter à moindre coût des données jusqu'à présent réservées à un usage interne. Elle peut ainsi faire jouir les habitants et les usagers de la ville d'une offre unique en France en terme de nombre d'applications et de services en lien avec leur territoire, sur des thématiques diverses. Telle est, du moins, l'objectif annoncé de cette plateforme :

> Ce patrimoine immatériel reste largement ignoré du grand public alors que cette somme d'informations est de nature à améliorer le quotidien des citoyens, à accroitre la transparence, à créer de la valeur d'usage. […] Enfin c'est participer au développement de l'attractivité de la métropole rennaise et poursuivre la construction d'une ville novatrice et audacieuse où s'invente la société et les usages de demain[52].

La ville de Paris a également souscrit à l'ouverture des données publiques depuis juin 2010 par décision du Conseil de Paris. Cette démarche s'est faite en concertation entre les services de la direction des affaires juridiques de la ville et diverses associations citoyennes. Tandis que Rennes Métropole a choisi de développer sa propre licence[53], la ville de Paris a pris le parti de la licence DDbL (*Open database License*) développée par l'*Open Knowledge Foundation*. Comme la licence développée par Rennes Métropole, elle autorise la réutilisation des données à des fins commerciales, à la condition que les réutilisateurs acceptent ensuite de partager leur création dans les mêmes conditions, selon des principes similaires aux dispositions de la licence *Creative Commons* CC-BY-SA. Pour ceux qui ne souhaiteraient pas partager à l'identique, la ville propose une plateforme payante non-exhaustive. Comme à Rennes, on y trouve surtout des données relatives à la citoyenneté, à l'urbanisme, aux transports, aux déplacements, aux services, à l'environnement et quelques jeux de données culturelles comme les notices des œuvres de fonds des musées, la liste des ouvrages disponibles et les statistiques de consultation des prêts dans les bibliothèques du réseau.

D'autres villes ont également rejoint la démarche comme Le Mans dont la plateforme propose déjà de nombreuses données cartographiques sous licence ODBL ou encore la ville et la communauté d'agglomération de Montpellier pour ne citer que quelques exemples. Puis d'autres initiatives ont par la suite dépassé le cadre communal ou intercommunal pour se positionner au niveau du département ou de la région. Certains départements comme la Loire ou Paris ont fait le choix de mettre en ligne une partie de leurs archives et un grand nombre de collectivités territoriales, principalement des régions, ont publié, en priorité, des informations géographiques. Avec son portail *data.visitprovence.com*, le Conseil général des Bouches du Rhône a ainsi regroupé et publié un jeu de 49 données touristiques

---

[52] http://www.data.rennes-metropole.fr/notre-demarche/ (consulté le 03/09/2012)

[53] La ville de Rennes et l'agglomération Rennes Métropole ont souhaité diffuser leurs données selon leur propre licence. Elle s'appuie sur les termes définis par la loi n°78-753 du 17 juillet 1978 relative à la réutilisation d'informations publiques mais autorise également : « les utilisateurs à partager, modifier et utiliser librement les données publiques disponibles sur le site www.data.rennes-metropole.fr, sous réserve que la modification ne constitue pas une altération des données ou une dénaturation de leur sens ». Direction Générale de l'Information et de la Communication, *Licence de réutilisation des informations publiques Rennes métropole en accès libre*, janvier 2011, 2 p. http://www.data.rennes-metropole.fr/fileadmin/licence/licence_janvier2011.pdf

(cc) BY-NC-ND

géolocalisées. Ce sont en général des données « métiers » de type cartographique.

Comme le souligne le rapport des Ponts[54] ces démarches répondent plus souvent aux besoins de partager des données entre entités publiques, dans l'espoir de réaliser des économies d'échelle, qu'à un souci d'ouverture vers le secteur privé.

De fait, la nature des données, la finalité de l'ouverture et le choix du public ciblé peuvent varier énormément d'une collectivité ou d'une initiative à l'autre. Leurs statuts sont également hétérogènes. Certaines plateformes ont été développées à l'initiative de services d'Etat, d'autres des régions et des départements, parfois sous forme d'association, d'EPIC, voire même sous statut privé comme c'est le cas dans le Pas-de-Calais. Comme en témoigne la carte de *l'Open data* développée et maintenue à jour par l'association LiberTIC, l'*Open data* français reste un mouvement non seulement inachevé mais surtout éparse et hétérogène. Les initiatives ont tendance à fleurir sans cohérence ou hiérarchie territoriale, avec un risque important d'enchevêtrement administratif. Pressées par des enjeux de pouvoir politique, de *marketing* et de communication, les collectivités précipitent souvent la publication de leurs données sans concertation préalable avec d'autres acteurs du territoire. Il s'agit là d'un sujet à part entière, mais il y a réellement une réflexion à mener sur l'échelle des initiatives d'*Open data* au regard du partage des compétences induit par les lois de décentralisation de 1982-1983, de 2003 et la réforme actuelle des collectivités territoriales.

… Aux démarches nationales

Comme le souligne le CNN, face à cette prolifération relativement rapide mais aussi à l'éclatement du mouvement, il est nécessaire de réfléchir à une « gouvernance des données publiques[55] ». Une suggestion consiste à faire évoluer la CADA en une « Haute autorité des données publiques ». Le point essentiel concerne l'aspect opérationnel, en insistant sur le fait qu'« il est absolument indispensable d'animer et d'accompagner les projets, de travailler à s'assurer à la réutilisation des données dans une stratégie globale de modernisation des administrations[56]». C'est effectivement une dynamique que l'on retrouve actuellement à travers deux mouvement nationaux d'encouragement à la réutilisation des données, chacun répondant néanmoins à des logiques bien différentes : d'une part la mission Etalab pour un *Open data* gouvernemental, d'autre part le collectif Open data France, pour un *Open data* territorial concerté.

La mission Etalab

Parallèlement à cet « Open data des territoires », des démarches ont commencé à voir le jour à l'échelle gouvernementale pour mettre en valeur, à l'échelle nationale, les données produites par les ministères, les grands établissements publics et les administrations centrales. En s'inspirant des plateformes américaine *data.gov* et anglaise *data.gov.uk*, l'État français lance

---

[54] Romain Lacombe, Pierre-Henri Bertin, François Vauglin, Alice Vieillefosse, *Pour une politique ambitieuse des données publiques. Les données publiques au service de l'innovation et de la transparence,* Rapport à la délégation aux usages de l'Internet dans le cadre du portail Proxima Mobile, remis au Ministre de l'Industrie, de l'énergie et de l'économie numérique, le 13 juillet 2011.

[55] Conseil national du numérique (CNNum), Avis n°12 relatif à l'ouverture des données publiques (« Open data »), 5 juin 2012, 23 p.

[56] Conseil national du numérique (CNNum), Avis n°12 relatif à l'ouverture des données publiques (« Open data »), 5 juin 2012, 23 p.

(cc) BY-NC-ND

également la sienne avec le portail *data.gouv.fr*. L'idée remonte à 2008 lorsque Éric Besson, secrétaire d'État chargé de la prospective, de l'évaluation des politiques publiques et du développement de l'économie numérique propose, dans le cadre du Plan numérique 2012, la création d'un portail unique de données publiques. En février 2010, le député Frank Riester réunit un groupe d'expert du numérique chargé de faire des propositions à Nathalie Kosciusko-Morizet, alors secrétaire d'état alors chargée de la prospective et du développement de l'économie numérique. Dès lors, la création de ce portail figure parmi les principales propositions d'encouragement à la création de services innovants. Son principe est validé dans le cadre de la Révision générale des politiques publiques en juin 2010 et fait l'objet d'une communication en Conseil des ministres le 24 novembre 2010.

Le 21 février 2011, le Premier ministre François Fillon, lance finalement la mission Etalab[57]. Rattachée au Secrétariat général du Gouvernement, elle est confiée à Séverin Naudet pour « coordonner l'action des administrations de l'Etat et apporter son soutien aux établissements publics administratifs pour faciliter la réutilisation la plus large possible de leurs informations publiques[58] ». En avril suivant, le Président de la République, Nicolas Sarkozy, en définit les missions à l'occasion du discours d'installation du Conseil national du Numérique :

> C'est pour un Etat plus transparent, une démocratie exemplaire, que j'ai demandé au gouvernement de créer un portail de données publiques en ligne. […] La question de l'accès aux données de base est une question absolument centrale.

Le 26 mai suivant, une circulaire[59] du premier ministre François Fillon, précise la démarche à suivre et conduit finalement l'ouverture le 5 décembre 2011 du portail *data.gouv.fr*[60] en proposant dès son lancement plus de 350 000 données issues des administrations centrales et déconcentrées mais aussi de quelques collectivités territoriales volontaires.

Le collectif Open data France

Dans une stratégie de coopération nationale des données publiques, la plateforme étatique *data.gouv.fr* a fait savoir qu'elle s'efforcerait de « rendre accessible les données de chaque collectivité qui le souhaite ». En encourageant la mutualisation d'outils techniques, le portail du gouvernement se propose comme une solution d'accompagnement à l'ouverture des données publiques des collectivités, qui restent néanmoins nombreuses à considérer avec prudence voire méfiance cette proposition. Beaucoup l'estiment, en effet, contraire à la logique de décentralisation mais s'accordent en revanche sur la nécessité d'une coopération, à l'échelle nationale, entre initiatives territoriales. Lors d'une présentation faite à l'occasion de la « semaine de l'*Open data* » à Nantes en mai 2012, Jean-Marie Bourgogne, responsable du programme Montpellier territoire Numérique, s'est

---

[57] Décret n°2011-194 du 21 février 2011 portant création d'une mission « Etalab » chargée de créer un portail unique destiné à mettre à disposition, de manière libre et gratuite, l'ensemble des informations publiques de l'Etat pour faciliter la coordination entre les administrations.

[58] http://fr.slideshare.net/Etalab/communiqu-de-presse-cration-de-la-mission-etalab-charge-de-crer (consulté le 09/12/12)

[59] Circulaire du 26 mai 2011 relative à la création d'un portail unique des informations publiques de l'Etat « data.gouv.fr » par la mission Etalab et l'application des dispositions régissant le droit de réutilisation des informations publiques.

[60] http://www.data.gouv.fr/

(cc) BY-NC-ND

attaché à relever les entraves qu'il reste aujourd'hui à surmonter pour optimiser l'*Open data* des territoires français. Parmi les difficultés rencontrées, il évoque la dispersion des acteurs, l'individualité des actions, la duplication des tâches comme la veille notamment, le manque de co-production et le peu d'actions collectives entreprises, même à l'échelle d'un territoire.

En réponse à ce besoin de coordination, plusieurs représentants de collectivités impliquées dans l'*Open data* ont décidé de se regrouper. Le 1er février 2012, l'association nantaise LiberTIC conviait à Paris, dans les locaux mis à disposition par la FING, les pilotes des différents projets *Open data*, identifiés à ce jour, pour réfléchir à la création d'une instance de réflexion supérieure commune. Sans prétendre à l'exhaustivité, on y retrouve aussi bien des institutions publiques que des collectivités ou de collectifs citoyens[61]. Comme l'explique Claire Gallon, fondatrice de l'association LiberTIC :

> L'idée d'une harmonisation remonte à 2010. A l'époque nous étions moins nombreux, beaucoup de projets étaient en construction et peu de collectivités pouvaient s'appuyer sur des équipes dédiées disposant d'un peu de temps pour réfléchir. […]. L'objectif était d'envisager des harmonisations et mutualisations dans leurs démarches afin de favoriser le développement du mouvement en France. L'approche consistait à offrir un cadre de réflexion et d'échange pour que les participants puissent travailler sur leurs problématiques actuelles et proposer des solutions collectives[62].

Ces réflexions ont rapidement conduit à la création d'un Collectif baptisé « *Open data* France ». Ce dernier s'est donné pour mission de mener des actions communes sur plusieurs chantiers : juridique, communication, veille, outils collaboratifs, normalisation, référentiel des bonnes pratiques, indicateurs, gouvernance, mise en place d'un Observatoire de l'*Open data*[63] etc. Un des gros travaux du moment concerne la constitution d'un « *Dataset* des *Datasets* », c'est-à-dire un catalogue des jeux de données permettant aux utilisateurs de prendre la mesure, à travers un seul fichier, de l'étendue des jeux de données libérées. Comme le souligne le Collectif :

> Si chacun des projets Open data utilise le même formalisme (colonnes et types d'informations) pour créer son catalogue des jeux des données, il sera d'autant plus facile de créer un méta-catalogue à l'échelle de tous les projets Open data[64].

Ce travail devrait permettre de faciliter l'identification des jeux intéressants pour les réutilisateurs, le repérage des jeux de données identiques ou proches dans chacun des projets ou encore des jeux de données qui mériteraient d'être normalisés pour faciliter leur agrégation et leur croisement grâce à la mise en relation des métadonnées. Comme le suggère David Eaves :

---

[61] *Villes ou Communautés urbaines* : Brest Métropole, Brocal, Grand, Toulouse, La Rochelle, Montpellier, Nantes et Nantes Métropole, Paris, Rennes Métropole ; *Départements* : Gironde, Loir-et-Cher, Loire-Atlantique, Maine et Loire, Saône et Loire ; *Régions* : Aquitaine, Provence-Alpes-Côte d'Azur ; *État* : Mission Etalab ; *Organisme* : LiberTIC, FING, CNIL, VVL, Pilote41.

[62] LiberTIC, Compte rendu de la première rencontre inter-administrations ouvertes, 13 février 2012. http://libertic.wordpress.com/2012/02/13/706/ (consulté le 04/09/2012).

[63] Ibid. : Claire Gallon (LiberTIC) propose un site commun de l'Open data sur l'exemple de Proyecto Aporta en Espagne. Ce site met à disposition une veille, des supports pédagogiques et est un lieu de partage d'outils qui a vocation à faciliter la mutualisation et le développement des initiatives d'ouverture des administrations publiques.

[64] http://opendatafrance.net/les-ateliers/atelier-3-dataset-des-datasets/ (consulté le 04/09/12).

(cc) BY-NC-ND

> Arrêtons donc de faire la comptabilité des portails *Open data*, commençons plutôt à dresser la liste des jeux de données qui adhèrent à des schémas communs. Étant donné que l'*Open data* est regardé de plus en plus favorablement par les gouvernements, *la création de ces schémas est actuellement, à mon sens, le défi majeur du mouvement Open data*. […] Nous arrivons à un point où le nombre de portails *Open data* devient de moins en moins pertinent. Ouvrir davantage de portails ne permettra pas à *l'Open data* de se déployer davantage. Ce qui va nous permettre de nous déployer sera d'établir des structures de données communes leur permettant de fonctionner *transversalement aux juridictions*[65].

La mise en œuvre de la directive 2003/98/CE dans les États-membres demande, par ailleurs, de disposer de normes de données pour structurer les informations publiques. Ces données de structuration sont en réalité des métadonnées pour lesquelles, le gouvernement britannique a déjà procédé à la définition d'un *e-Governement Metadata Standard (eGMS)*. En France, un travail de normalisation similaire a commencé à voir le jour à l'initiative de quelques acteurs[66], avec la mise en ligne d'un « document de normalisation des *dataset* [67] ». Le schéma de description proposé s'attache à promouvoir une publication du catalogue selon les standards du web de données. Le catalogue permettrait d'utiliser les données selon le format DCAT (*Data Catalog Vocabulary*) du W3C permettant une interopérabilité et la création future de liens sémantiques entre les données de chaque catalogue pour constituer un méta-catalogue.

UN AVENIR ENCORE INCERTAIN

Un cadre juridique en chantier

Derrière ces enjeux forts de transparence démocratique et de croissance économique, l'*Open data* s'inscrit nécessairement dans un cadre juridique sans lequel, la dynamique d'ouverture des données publiques, n'aurait ni existence légale ni force de persuasion. Ce cadre juridique est à l'image du mouvement à savoir complexe, mouvant, en construction.

D'une logique de demande à une logique d'offre

Depuis les fondements posés par la loi CADA du 17 juillet 1978 en matière d'accès aux documents administratifs, l'émergence d'une société de l'information et le développement de l'Internet et des nouvelles technologies de l'informatique, ont contribué à faire évoluer les problématiques relatives à la diffusion de ces documents. L'ère du numérique appelle à une nouvelle médiation entre les administrations et les demandeurs. Comme le rappelle le Conseil national du numérique, le discours du premier Ministre Lionel Jospin à Hourtin le 25 août 1997[68] marque la première reconnaissance, en France, de la dimension économique

---

[65] David Eaves, Un État des lieux de l'Open data, OWNI, 2 novembre 2011. http://owni.fr/2011/11/02/etat-des-lieux-open-data-eaves/ (consulté le 04/09/2012)

[66] Conseil général de la Gironde et de la Région Aquitaine, de la Ville de Montpellier, de la Métropole Nantaise et de la Ville de Paris.

[67] https://docs.google.com/spreadsheet/ccc?key=0Ajn0buNVT5BRdDQ5eWdBYzRuNFhIRDdQZkVhb0Q2T3c&hl=fr#gid=2. Chaque nature de champs est accompagnée de sa déclinaison DCAT. Le schéma de données DCAT est basé sur le schéma de données RDF, élaboré à l'université DERI de Galway et aujourd'hui maintenu par le W3C.

[68] Lionel Jospin, « Préparer l'entrée de la France dans la société de l'information », Discours du 25 août 1997 à Hourtin.

(cc) BY-NC-ND

de certaines données publiques et la nécessité de les diffuser gratuitement. Peu suivi d'effet, il faut attendre l'ordonnance du 6 juin 2005[69] pour compléter le dispositif créé en 1978. En transposant la directive européenne du 17 novembre 2003[70], elle pose les principes d'une libre réutilisation des informations publiques, définies comme « les informations contenues dans les documents administratifs ». La directive ne revient pas sur les exceptions définies par la loi CADA, dues à la concurrence avec d'autres textes de lois relatifs au secret défense, au secret statistique[71], à la protection des données publiques nominatives (loi CNIL[72]), au secret bancaire ou encore au droit d'auteur et à la propriété intellectuelle. Elle permet en revanche le passage de la notion de « documents administratifs » à celle d' « information publique ». Pour passer à une logique d'*Open data*, le CNN encourage le basculement total vers la notion de « données publiques ». Cet affinage nécessite un effort plus important des administrations dans la communication de leurs données[73].

Par ailleurs, comme le fait remarquer Bernadette Kessler, responsable du service innovation numérique à la direction de la communication de Rennes Métropole, le code des marchés publics reste largement inadapté à la logique *Open data* et entrave les collectivités dans le soutien qu'elles pourraient apporter aux applications et aux initiatives économiques locales. Selon elle, « il faut s'interroger sur la philosophie juridique de l'*Open data* [74] » et penser à un cadre juridique plus en adéquation avec les ambitions politiques et économiques que l'on souhaite donner à la démarche. Elle est rejointe en ce sens pas le CNN depuis juin 2012 s'efforce de réfléchir à une réforme du droit relatif à la mise à disposition des données publiques :

> Le lancement du portail data.gouv.fr et les initiatives des collectivités territoriales ont impulsé une dynamique importante dont il faut se féliciter. Cependant le cadre législatif sous-jacent, et notamment la loi du 17 juillet 1978 (dite loi CADA) demeurent encore par certains aspects restrictifs, ambigu, et inadapté aux caractéristiques et aux enjeux de la société du numérique[75].

L'enjeu est de passer d'une logique de demande, qui prévaut encore aujourd'hui, à une logique d'offre. Un travail qui pourrait s'inscrire dans le cadre de l'élaboration et de la transposition en droit français d'une nouvelle directive européenne sur la réutilisation des informations du secteur public.

---

[69] Ordonnance 2005-650 du 6 juin 2005 relative à la liberté d'accès aux documents administratifs et à la réutilisation des informations publiques.
http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000629684

[70] Directive 2003/98/CE du 17 novembre 2003 concernant la réutilisation des informations du secteur public.

[71] Loi n° 51-711 du 7 juin 1951 sur l'obligation, la coordination et le secret en matière de statistiques.
http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000888573 (consulté le 28/10/2012)

[72] Loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, dite Loi CNIL : la loi ne traite pas explicitement de la commercialisation des données publiques.
http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000886460&dateTexte=20120901 (consulté le 28/10/2012)

[73] C'est notamment le cas pour les archives : un dossier papier dont les bornes chronologiques ou certains documents ne rentrent pas dans le cadre légal de communication doit faire l'objet d'une demande de dérogation. La procédure peut conduire au retrait du dossier des pièces compromettantes ou à l'inverse à la communication de seulement quelques pièces du dossier mais jamais à la notion de contenu informationnel. L'informatique permettrait de franchir le cap à condition de permettre l'occultation pour des données ou au contraire l'extraction des éléments pertinents.

[74] Entretien avec Bernadette Kessler, responsable de l'action I'Open data au sein de la direction générale de la communication de la communauté d'agglomération Rennes Métropole, Rennes, le 24 octobre 2012.

[75] Avis n° 12 du Conseil national du numérique relatif à l'ouverture des données publiques (« Open data »), 5 juin 2012, p.4.

(cc) BY-NC-ND

Aujourd'hui la publication, le plus souvent sous la forme d'une mise en ligne, n'est requise à titre obligatoire que pour certains documents, généralement législatifs ou réglementaires[76]. Pour les autres, les procédures régissant le droit d'accès peuvent prendre jusqu'à quatre mois en cas de refus ou d'absence de réponse nécessitant la saisine de la CADA[77]. Par ailleurs, cette logique de demande suppose que le demandeur connaisse l'existence des données. Derrière l'Open data, il y a donc un vrai travail de signalement qu'il convient de penser et de réaliser, bien au-delà de la simple mise en ligne. Le passage à une logique d'offre, implique de prendre les devants et systématiser la mise à disposition des données publiques pour que l'offre crée la demande et suscite l'innovation, non l'inverse.

Les licences Open data

L'*Open data* est une des solutions proposées pour répondre à ces attentes. Les licences *Open data* laissent en effet un grande marge de manœuvre et d'initiatives aux utilisateurs finaux pour laisser libre cours à leur créativité voire pour certains leur soif d'entreprenariat. Il n'existe pas une, mais des « licences ouvertes » conçus en fonction de contenus et de supports spécifiques, par ailleurs réutilisables et adaptables les besoins particuliers. Sans prétendre à l'exhaustivité, certaines licences types se retrouvent fréquemment utilisées dans le domaine de l'Open data :

L'*Open database License*[78] (ODBL) et l'*Open data Commons*[79] (ODC-By), toutes deux de l'*Open Knowledge Foundation*. La première permet à chacun d'exploiter publiquement, commercialement ou non, des bases de données à condition néanmoins de maintenir la licence sur la base de données et les modifications qui pourraient y être apportées et d'en indiquer l'usage ; la seconde autorise l'utilisation, la copie, la redistribution, la modification, la réalisation de travaux dérivés de la base de données, sous réserve d'indiquer le nom de l'auteur de la base de données originale. Les licences de types *Creative Commons*[80], à l'exception de la CC0, ne sont pas vraiment adaptées à des bases de données, car elles sont uniquement centrées sur le droit d'auteur, et négligent de fait le droit *sui generis* des bases de données. En revanche, elles sont fréquemmen utiliser pour des documents en *Open access*. Enfin l'*Open Licence* ou Licence Ouverte avec sa transcription française, présentée le 18 octobre 2011 par Etalab[81] est le dernier type de licences le plus fréquemment utilisé par les institutions françaises. Elle respecte le droit commun en appliquant les vertus de l'*Open data* et s'avère compatible avec les licences internationales les plus utilisées (ODBL et ODC). Elle est facile à comprendre et à appliquer, disponible en français et en anglais. Elle autorise la reproduction, copie, publication et transmission de « l'Information » ; sa diffusion et sa redistribution, son adaptation, modification, et extraction, notamment pour créer des « Informations dérivées ». L'information peut être exploitée à titre commercial en la combinant avec d'autres informations, ou en l'incluant dans un autre produit ou une autre application.

[76] On pense notamment à la base Légifrance.
[77] Commission d'accès aux documents administratifs
[78] http://opendatacommons.org/licenses/odbl/
[79] http://opendatacommons.org/
[80] http://creativecommons.org/about/cc0
[81] http://www.etalab.gouv.fr/pages/licence-ouverte-open-licence-5899923.html

(cc) BY-NC-ND

Bien qu'il soit tout à fait possible de modifier ces licences, il est en général préférable d'utiliser des licences *Open data* sans les modifier car elles fonctionnent selon les mêmes principes et dans une logique d'interopérabilité et de compatibilité. La licence ouverte d'Etalab est ainsi compatible avec les *Open Government Licence* (OGL) du Royaume-Uni, les *Creative Commons Attribution 2.0* (CC-BY 2.0) de *Creative Commons* et l'*Open data Commons Attribution* (ODC-BY) de l'*Open Knowledge Foundation*. Toutes prévoient une mention d'origine : nom du ou des producteurs, lien hypertexte vers le site ou catalogue source et date de dernière mise à jour. Ces indications peuvent être mises en annexe de la base de données ou du fichier créé, à la suite des données, dans un cadre attaché aux données ou dans un fichier joint. A l'échelle française, outre la licence Étalab, on peut également mentionner d'autres licences attachées à une démarche d'ouverture des données publiques ou privées comme la licence APIE[82] (Agence du patrimoine immatériel de l'État), la licence IP[83] (informations publiques) et beaucoup d'autres licences spéciales rédigées par les producteurs eux-mêmes, généralement basées sur des licences qui existent déjà comme c'est le cas pour celles de Rennes Métropole[84]. Cette dernière ne confère « en aucun cas des droits de propriétés intellectuels sur ces données ». Seul le droit moral, inaliénable et imprescriptible selon l'article L.121-1 du Code de la propriété intellectuel, reste attaché aux données. Le producteur conserve la propriété intellectuelle de ses données et la collectivité précise bien décliner toutes responsabilité économique ou juridique :

> Dans l'hypothèse où des évolutions, liées notamment au changement de format, à la modification de structure des bases de données, ou de fichier et/ou d'évolution des modalités de mise à disposition des données, induiraient pour le Licencié des adaptations techniques de ses équipements, le surcoût qui en découlerait ne pourrait en aucun cas être mis à la charge de l'Administration. Le Licencié ne peut prétendre à aucune indemnité ni compensation à ce titre[85].

Quel schéma pour l'*Open data* ?

D'après le site de la Fondation pour la Recherche sur les Administrations Publiques et les Politiques Publiques (FRAP), les données publiques représenteraient « un marché de 820 millions d'euros mais qui pourrait atteindre 1,3 milliard d'euros », ce qui « représente une source supplémentaire de revenus que les ministères désireraient ardemment capter dans un contexte de contrainte budgétaire particulièrement forte.[86] ». Sur cette estimation un arrêté du 4 août 2010 prévoit ainsi un intéressement de 20 % du chiffre d'affaires pour la réutilisation commerciale de la base des données comptables des collectivités du ministère des Finances. Néanmoins, comme le souligne Simon Chignard ou Bernadette Kessler, à Rennes Métropole, les retombées économiques ne sont pas à la hauteur des attentes. De fait, beaucoup d'initiatives de réutilisation restent encore très liées à la

---

[82] http://www.economie.gouv.fr/apie/donnees-et-images-modeles-licences-type-0
[83] http://www.rip.justice.fr/category/licences/licence-ip
[84] http://www.data.rennes-metropole.fr/notre-demarche/licence-d-utilisation/
[85] Direction Générale de l'Information et de la Communication, *Licence de réutilisation des informations publiques Rennes métropole en accès libre*, janvier 2011, 2 p. http://www.data.rennes-metropole.fr/fileadmin/licence/licence_janvier2011.pdf
[86] La Gazette des communes, « Données publiques : une nouvelle économie des services publics en débat », Réutilisation des données publiques : des promesses vertigineuses. http://www.lagazettedescommunes.com/48192/donnees-publiques-une-nouvelle-economie-des-services-publics-en-debat/

(cc) BY-NC-ND

communauté de l'*Open source* ou à une logique citoyenne et l'*Open data* cherche encore son public commercial.

Par ailleurs, comme dans toute phase de construction, il reste difficile de définir avec précisions les contours du mouvement. Encore jeune, il suscite autant d'espoirs que d'interrogations, ce qui en fait sa force et sa faiblesse. On « entre en Open data » par conviction, par curiosité, par intérêt, dans un but précis ou au contraire en quête de quelques solutions ou de quelques propositions. Dans bien des cas, il est intéressant de constater qu'au sein d'une même institution ou d'une même association, les motivations peuvent être très différentes voire contradictoires. Comme l'explique Clément Renaud[87] du Collectif lyonnais Opendata69, au sein de l'association « tout le monde n'a pas les mêmes attentes vis-à-vis du mouvement ». Tandis que certains membres sont très investis dans un travail main dans la main avec l'agglomération du Grand Lyon pour l'ouverture des données publiques, d'autres, comme lui, sont plus convaincus de l'intérêt économique de l'*Open data* notamment pour le secteur privé et particulièrement dans le secteur du Web, lui-même étant moins persuadé de l'impact économique local comme semble également le constater d'autres collectifs régionaux.

Comme l'a démontré la FING trois avenirs s'offrent aujourd'hui à l'*Open data* : un avenir radieux où une masse de données contribuerait à promouvoir la transparence démocratique et une croissance économique appuyée sur des services aux citoyens ; un avenir sombre où les questions éthiques et juridiques paralyseraient une masse de données aux usages potentiellement malveillants ; et enfin celui d'une indifférence voire d'une déception face à une démarche qui n'a pas réussi à tenir ses promesses. À la croisée de ces trois avenirs possibles, on retrouve le problème crucial de la médiation qu'appelle aujourd'hui l'*Open data*. Comme tout mouvement, pour perdurer l'*Open data* doit faire écho auprès d'un public et nécessite pour cela d'être animé. Comme l'illustre l'article de David Kaplan « L'ouverture des données publiques et après ?[88] » ou encore celui de Sabine Blanc[89] à partir de l'expérience rennaise, le problème de l'animation suscite toujours autant d' interrogations parmi les acteurs de l'*Open data* dont les rôles n'ont jamais encore été clairement définis.

---

[87] Entretien avec Clément Renaud, le 12 septembre 2012 à Lyon. Journaliste et développeur, Clément Renaud s'est spécialisé, après un passage en Chine, dans le monde du Web et de la donnée. Membre d'un groupement d'entrepreneurs baptisés « Make Sens, challenging poeple for social business », il a déjà organisé dans ce cadre un *workshop* sur l'Open data et programme un nouvel évènement de ce type avec le Collectif Opendata69.

[88] http://owni.fr/2011/01/18/louverture-des-donnees-publiques-et-apres/ (consulté le 26/11/12)

[89] Sabine Blanc, « Open data péRennes », 21 février 2012.
http://owni.fr/2012/02/21/open-data-perennes/ (consulté le 13 décembre 2012)

(cc) BY-NC-ND

BIBLIOTHEQUES : LES BONS ELEVES DE L'*OPEN DATA*

Au regard des grands enjeux qui ont motivé jusqu'ici l'ouverture des données publiques, la place des données culturelles dans cette démarche ne semble pas nécessairement évidente. Une brève requête par catégories dans les bases ouvertes de données publiques permet rapidement de confirmer la faiblesse du champ culturel au sein de ces initiatives. À Rennes par exemple, les résultats ne font ressortir que trois jeux de données culturelles : un émanant des services de la ville sur la localisation des panneaux d'affichage de l'agglomération et les deux autres, produits par la bibliothèque à partir de ses statistiques de fréquentation par jour et par heure. Ainsi, si l'on peut regretter cette faible représentativité des données culturelles, on peut en revanche se réjouir de la présence des données de bibliothèques parmi elles. A partir du cas de Rennes Métropole et de quelques autres exemples nationaux er internationaux, nous essaierons de montrer en quoi les bibliothèques peuvent prétendre aujourd'hui au titre de « bons élèves de l'*Open data* » dans le domaine culturel mais pas uniquement.

L'EXPERIENCE DE LA BIBLIOTHEQUE CHAMPS LIBRES A RENNES

Comme nous l'avons évoqué, la communauté urbaine de Rennes Métropole a été à l'origine d'une des premières démarches françaises d'ouverture des données publiques. Conséquemment, la médiathèque des Champs-Libres fait également figure de pionnière dans ce domaine. Cet exemple permet d'illustrer une des démarches possibles d'*Open data* en bibliothèques. Il permet surtout de mettre en évidence les atouts mais aussi les limites de la stratégie adoptée.

L'ouverture des données : une demande de la collectivité

A Rennes, l'initiative d'ouvrir les données de la bibliothèque est venue conjointement de la direction de la communication de Rennes Métropole, en la personne de Bernadette Kessler et de la direction des affaires culturelles, qui souhaitaient que celle-ci puisse trouver sa place au sen du projet *Open data* de la collectivité.

Quelles données ouvrir ?

L'enjeu pour les services de la bibliothèque a d'abord été de trouver des jeux de données à proposer qui ne soient ni contraints par la protection du droit d'auteur et de la propriété intellectuelle, ni soumis aux restrictions imposées par la CNIL qui puissent présenter un intérêt quelconque pour le public et surtout qui soient facilement techniquement exportables et communicables pour la bibliothèque. Comme le rappelait Samuel Goeta de la « Netscouade » lors de la semaine de l'Open Data en mai 2012[90], les données culturelles regroupent essentiellement trois

[90] Samuel Goeta, Compte-rendu de l'atelier sur les données culturelles, Semaine de l'Open Data 21-26 mai 2012, juin 2012. http://www.opendataweek.org/wp-content/uploads/2012/06/CR-AtelierB-2505.pdf

(cc) BY-NC-ND

types de contenus : les données de fonctionnement d'une institution culturelle, les œuvres et productions artistiques et enfin le catalogue et les données de catalogage de l'institution culturelle, plus généralement qualifiées de métadonnées. Ces dernières constituent la valeur-ajoutée des institutions culturelles puisqu'elles viennent qualifier et enrichir le contenu des œuvres présentes dans les institutions culturelles : musées, bibliothèques, archives etc. Il convient de les distinguer de l'œuvre elle-même car en tant qu'éléments d'enrichissement, elles sont produites par l'institution elle-même et présentent donc l'avantage de ne peut être soumises au droit d'auteur. En tant qu'informations libres de droit et plus-values scientifiques, ces données constituent donc pour les institutions culturelles, des candidats idéales à une politique d'Open Data.

Un choix contraint …

A Rennes, une première réflexion a fait ressortir plusieurs jeux de données, présentant des enjeux, des niveaux de complexité, des temps de travail et des intérêts très variés pour le public comme des données de prêt, des données budgétaires, les données de fréquentation et ou encore un projet encore inachevé de bibliothèque numérique patrimoniale en RDF. Quelques suggestions sont également remontées du public lui même. Frédérique Schlosser, coordinatrice de l'action *Open data* au sein de la bibliothèque s'est avouée parfois surprise par ces suggestions : « un journaliste est venu me voir un jour, pour me demander si nous possédions des données sur les dons d'ouvrages par la bibliothèque aux associations tandis qu'un chercheur espérait des statistiques sur les ouvrages le plus volés ». Mais en réalité, le véritable intérêt de la DGC[91] se portait sur les données bibliographiques de la bibliothèque. Malheureusement, comme l'explique Frédérique Schlosser, « le système de gestion intégré de la bibliothèque ne nous permettait pas d'exporter de manière structurée les données du catalogue mais seulement une liste d'ouvrages transposée dans un tableur excel[92] ». L'opération étant par ailleurs extrêmement laborieuse, une mise à jour régulière n'était pas envisageable, ce qui diminuait l'intérêt de ces données pour le lecteur. De fait, il est important de rappeler que si l'*Open data* n'est pas à proprement une démarche technique, comme bien souvent elle se heurte aux contraintes et aux rythmes d'évolution imposés par celle-ci.

Le choix s'est finalement porté sur les statistiques de fréquentation de la bibliothèque. Ce n'était certes pas le plus stimulant mais comme l'explique Frédérique Schlosser, il résultait d'un compromis entre offre de services et attentes potentielles du public, dont elle perçoit cependant bien les limites. Ces données, collectées par les portiques antivol du rez-de-chaussée et de la salle Vie du Citoyen de la bibliothèque présentent l'avantage d'être facilement exportables, relativement simples à comprendre et régulièrement actualisées. Grâce au travail d'épuration, d'affinement et de mises en forme des données réalisées avec l'aide de Sarah Toulouse et de Léa Lacroix, celles-ci ont pu être mises en ligne assez rapidement au format CSV.

[91] Direction générale de la communication/
[92] Entretien avec Frédérique Schlosser.

(cc) BY-NC-ND

… et attentiste

Contrairement à l'idée véhiculée par l'appel à la donnée brute (*Raw data*) lancé par Tim Berners-Lee à la TED en 2009[93], une donnée n'est jamais vraiment brute. Elle possède toujours un format, un auteur, un contexte et une force signifiante induite par sa simple publication. Toute mise en ligne de données nécessite donc une réflexion préalable et un travail de reprise formelle et de documentation par les établissements, sur les informations mises en ligne. Or comme le souligne Frédérique Schlosser, « tout cela prend du temps et avant de nous investir d'avantage dans la démarche, nous voudrions être sûr que cela en vaut la peine[94] ». Or créer une dynamique autour de l'*Open data* au sein d'établissements qui n'ont ni de temps ni d'effectifs suffisants à dédier au projet relève d'un tour de force qui frôle parfois le cercle vicieux. Les retour d'investissements en *Open data* se calculent souvent d'après les statistiques de réutilisation des données mises en ligne et les développements d'applications qu'elles ont suscités, or la quantité et la qualité de ces réutilisations sont largement facteurs de celles des données elles-mêmes. Le risque est donc de voir la démarche s'essouffler avant même d'avoir commencé. Pour éviter la quadrature du cercle, l'enjeu est donc de trouver une forme de « médiation » capable de créer et de stimuler des liens entre producteurs et consommateurs de données et d'insuffler une dynamique de plus long terme.

Stimuler la réutilisation : le recours à des tiers médiateurs

Le relatif succès des opérations d'ouverture de données, entreprises par les Champs-Libres au sein de la métropole rennaise s'explique en grande partie par les relations qu'ont réussi à nouer les services de la bibliothèque et d'autres acteurs du territoire autour de cette question de l'*Open data*. Comme dans d'autres collectivités, c'est du milieu associatif qu'est venu ce souci de médiation. Deux acteurs ont joué un rôle particulièrement important dans la promotion des données ouvertes par la collectivité de manière générale, et par la bibliothèque plus particulièrement : le « Collectif *Open data* Rennes Métropole » mené par Léa Lacroix[95] et la « Cantine Numérique Rennaise », sous l'impulsion de Simon Chignard.

Le collectif Open data Rennes Métropole

A la différence d'autres villes comme Montpellier, où la démarche citoyenne a précédé l'initiative institutionnelle, à Rennes, le Collectif Open data Rennes Métropole s'est constitué postérieurement à l'ouverture des premiers jeux de données de la collectivité. Fondé en janvier 2012 sous l'impulsion de Léa Lacroix, alors stagiaire au sein de la direction des services de la communication de l'agglomération rennaise, il réunit pour l'instant un noyau dur d'une dizaine de personnes animant des ateliers proposés à titre gratuit aux citoyens ou des forums de dialogue et d'échanges entre institutions productrices de données et développeurs-réutilisateurs.

---

[93] http://www.ted.com/talks/lang/fr/tim_berners_lee_on_the_next_web.html
[94] Frédérique Schlosser, entretien du 20 juillet 2012 aux Champs-Libres.
[95] http://opendatarennes.wordpress.com/

(cc) BY-NC-ND

Les réunions ont lieu environ une fois par mois au sein de la Cantine Numérique Rennaise, autour d'ateliers spécifiquement dédiés à l'*Open data* ou plus généralement sur d'autres sujets relatifs à la culture ou à l'actualité civique et citoyenne, comme les élections par exemple. Ces ateliers sont l'occasion de rencontres et d'échanges sur des projets en cours, de nouvelles perspectives et un moyen de faire appel aux bonnes volontés. Du fait de sa postériorité à la démarche d'ouverture, le collectif se positionne, dès le départ, comme une instance d'animation de l'*Open data* à l'échelle de la collectivité et non comme une institution citoyenne combative et militante. Comme l'explique Léa Lacroix :

> Dans la mesure où l'initiative d'ouverture des données émanait d'une démarche volontaire de la collectivité, nous avons eu le sentiment que les choses nous tombaient directement dans le bec et nous nous sommes même octroyé la liberté de critiquer les premières données mises en ligne[96].

Le Collectif s'est surtout formé dans l'idée de promouvoir les données auprès des développeurs et de faciliter les rencontres et les échanges entre ces derniers et les producteurs de données. Il compte aujourd'hui, de manière plus ou moins formelle, une dizaine de personnes toutes déjà bien implantées dans le domaine de l'*Open data*. Ses activités sont étroitement liées à celles de la Cantine Numérique Rennaise.

La Cantine Numérique Rennaise

Comme on peut le lire sur le site[97], cette association de type loi de 1901, dont on retrouve le concept dans d'autres villes comme Paris ou encore Toulouse, a pour objectif de favoriser le développement économique et social par la création d'un "territoire d'innovation numérique". Elle vise à favoriser l'activité, l'invention et l'échange autour des pratiques, services et usages innovants utilisant le numérique. Elle se présente comme :

> Une communauté ouverte, et active incarnée par un lieu physique, une dimension numérique, et des acteurs individuels ou collectifs du territoire local : collectivités, organismes, entreprises, associations, collectifs et publics[98].

Assez proche de l'esprit d'ouverture de l'*Open data*, ce lieu présente également l'avantage de fournir un cadre neutre de rencontre entre la Collectivité, le Collectif, divers acteurs privés ou publics de l'activité numérique rennaise et les institutions productrices comme la bibliothèque, avec qui les contacts ont été facilités par la proximité géographique, puisque la cantine numérique et la bibliothèque des Champs-Libres partagent le même bâtiment. Cette neutralité du lieu est d'autant plus importante que quelques développeurs ont pu manifester une certaine méfiance vis-à-vis de la Collectivité en ne souhaitant pas, par exemple, que leurs applications soient recensées ou même disponibles sur la plateforme de la Collectivité. L'intérêt de ces associations est de parvenir à créer des tiers-lieux de confiance au sein desquels les personnes intéressées peuvent échanger sur les

[96] Léa Lacroix, entretien du 20 juillet 2012 à la Cantine Numérique Rennaise, Rennes.
[97] http://www.lacantine-rennes.net/
[98] http://lacantine-rennes.net/wp-content/uploads/statuts-rev2012.pdf (consulté le 28/08/2012)

(cc) BY-NC-ND

données, faire état de leur réflexion ou de leurs incompréhensions ou même venir dialoguer avec des acteurs de la fonction publique comme des officiers d'Etat Civil comme cela a été le cas à Rennes à l'occasion de l'ouverture des fichiers sur les prénoms.

Une collaboration avec la bibliothèque

En collaboration avec ces partenaires, la médiathèque des Champs-Libres a pris part à l'animation d'ateliers autour de la réutilisation des données publiques comme l'atelier pratique du 18 juin 2012 sur la réutilisation des données monté par le Collectif avec le soutien de la Cantine Numérique Rennaise (CNR), ou encore, quelques semaines plus tard, celui du 30 juin sur les outils techniques de « visualisation des données »[99]. Elle a aussi contribué à l'organisation d'un débat à la CNR sur « Quelles données ouvrir en bibliothèques ? ». Ces rencontres ont été l'occasion pour les producteurs de données et les ré-utilisateurs, qu'ils soient simples consommateurs, informaticiens ou programmateurs, d'échanger sur leurs besoins, leurs attentes et les éventuelles difficultés rencontrées (formats fermés ou propriétaires, complexité des données, mise à jours, légende etc.) pour essayer de remédier à des situations de blocages techniques qui peuvent constituer un frein supplémentaire à la progression de l'*Open data* à l'échelle de la collectivité.

Quelles réutilisations ? Quels réutilisateurs ?

Le collectif Open data Rennes Métropole et la CNR ont donc joué un rôle important dans la réutilisation des données de la bibliothèque. Par le biais du Collectif, Léa Lacroix a lancé l'« Opération #biblioviz, visualiser les données de la bibliothèque Rennes Métropole[100] ». La première démarche fut de récupérer les tableaux de données et d'en proposer une version simplifiée plus accessible aux utilisateurs les moins aguerris. Les fichiers ont été enrichis d'un travail de description et d'analyse de contenu qui facilite la prise en main des données. A partir de là, le collectif se propose de recenser toutes les initiatives de réutilisation des données des bibliothèques tout en rappelant qu'il ne s'agit là ni d'un concours ni d'un appel à projet. L'opération #biblioviz a déjà donné lieu au recensement de cinq initiatives de réutilisation: quatre d'ordre de l'infographique et la dernière de type application. Elles s'inspirent fortement du site de « *datavisualisation* » de la fréquentation des lieux culturels de Londres « *When should I visit ?* », qui indique par jour de la semaine le taux de fréquentation du lieu[101].

Ces outils permettent de mieux comprendre les cycles d'activité de la bibliothèque et de mettre en relation la fréquentation des lieux avec ses rythmes d'activités (fermeture annuelle, horaire d'été etc.) et ceux de ses lecteurs (période de partiels, rendus de thèses et mémoires, rentrée scolaire etc.). Ces outils donnent au lecteur des clés pour optimiser sa fréquentation et son utilisation des services de la bibliothèque mais offre aussi à la bibliothèque un moyen d'externaliser l'analyse et l'exploitation de certaines de ses données statistiques. Par sa politique d'*Open data* la bibliothèque des Champs-Libres contribue ainsi non seulement à la mise en

---

[99] http://opendatarennes.wordpress.com/2012/06/19/compte-rendu-des-precedents-ateliers/
[100] http://opendatarennes.wordpress.com/operation-biblioviz/
[101] Comme le font remarquer Léa Lacroix et Florian Strzelecki dans leur compte-rendu de la rencontre Compte-rendu de la rencontre Culture & Patrimoine du 19 avril 2012 à la Cantine numérique rennaise : le site anglais est un bon exemple de visualisation des données, mêmes si les chiffres ne sont pas fournis par les institutions publiques directement mais issues du réseau social *FourSquare*. Il s'agit donc de données publiques mais issues du *crowdsourcing*.

(cc) BY-NC-ND

place d'une offre de service innovante mais elle gagne aussi de nouveaux outils lui permettant de mieux appréhender le profil de ses lecteurs, leurs pratiques et d'améliorer en retour son offre de services. Loin d'être un investissement et un service aussi unilatéral qu'il y paraît, l'*Open data* possède une dimension « donnant-donnant ».

Il convient tout de même de noter que ce bilan d'activité est quelque peu biaisé par le fait que quatre de ces initiatives sont du même auteur, à savoir Léa Lacroix alias « Auregann ». Les idées ne manquent pas pourtant, comme en témoigne les propositions de Florian Strzelecki alias « Exirel » : Jeune informaticien, issu du milieu du logiciel libre, il trouve, dans ces données ouvertes, une opportunité de conjuguer ses compétences professionnelles avec ses goûts personnels tout en les mettant au service d'un idéal citoyen :

> Mon parcours personnel doit sans doute entrer en ligne de compte dans ces idées, mais je ne saurais dire comment, tout est un peu lié. J'aime le logiciel libre, l'ouverture et les communautés qui se créent autour d'outils libres, je suis membre du « Parti Pirate » et convaincu que l'open-data est importante pour notre société, et enfin j'aime lire et écrire, j'aime les livres, le numérique, et les bibliothèques Il y a aussi un côté "défi" à trouver une représentation avec aussi peu de données. D'autant plus lorsque la question qui est soulevée fréquemment par les producteurs de données est "à quoi ça sert ?", et à laquelle il faut répondre par des exemples concrets[102].

Avec quelques personnalités gravitant autour de l'association, il suggère d'utiliser les chiffres de fréquentation pour réaliser une longue frise décorative pour la bibliothèque qui permettrait de retracer les flux de visiteurs ou bien encore de créer une mélodie à partir de ces statistiques en s'inspirant, pour la composition, de sons familiers de la bibliothèque : bruit de pages qui se tournent, livres qui se ferment, bruit de pas, chuchotements etc. Ce type de suggestion permet d'entrevoir la dimension à la fois marketing et culturelle – notamment en tant que support de création artistique - que peut représenter l'*Open data* pour des données d'institutions comme les bibliothèques dont les enjeux citoyens et économiques sont certainement moins évidents qu'ailleurs. Cependant, comme l'explique Florian Strzelecki, par manque de temps et manque de compétence en matière musicale, le projet reste encore à l'état d'idée et attend les bonnes volontés pour être développé.

DES PROJETS QUI FLEURISSENT…

A la suite ou parallèlement à la démarche entreprise par la bibliothèque des Champs-Libres, d'autres établissements français ou étrangers ont également décidé de se lancer dans des projets *Open data*. En raison des données issues de la numérisation, la frontière est parfois ténue, en bibliothèque, entre des démarches d'*Open access* et d'*Open data* à proprement parler, d'autant que toutes deux participent d'une dynamique commune. D'ailleurs il n'est pas rare que les deux démarches convergent au sein d'un même projet puisqu'associées aux documents numérisés se trouvent, en général, des métadonnées qui peuvent être également

[102] Florian Strzelecki, développeur Web chez Envivio à Rennes et membre du Collectif Open data Rennes Métropole. http://exirel.me/. (Échange du 27/08/2012)

(cc) BY-NC-ND

librement téléchargeables et réutilisables. Encore une fois, l'évolution extrêmement rapide de ces initiatives ne permet pas de prétendre à l'exhaustivité. En revanche quelques établissements reviennent régulièrement parmi les exemples cités.

La Bibliothèque Nationale Universitaire (BNU) de Strasbourg

Lors de sa séance du 27 janvier 2012, le conseil d'administration de la Bibliothèque nationale universitaire (BNU) de Strasbourg a décidé d'autoriser la libre réutilisation des fichiers images produits par la BNU et de leurs métadonnées associées, en les plaçant respectivement sous la Licence Ouverte d'Etalab et en CC0. Grâce à cette initiative la BNU se place dans l'esprit de l'*Open data* pour permettre la réutilisation par tous de ses contenus à des fins dépassant largement les enjeux bibliothéconomiques. Interviewé par Rémi Mathis, Frédéric Blin, le directeur de la conservation et du patrimoine de la BNU, membre du comité de direction de l'IFLA, justifie sa démarche et sa prise de position sur le sujet en invoquant sa mission de service public :

> La BNU a trois missions principales : favoriser la recherche, valoriser nos fonds, les conserver. Notre position de départ est qu'il nous faut lever tous les freins ou complications inutiles à l'accomplissement de ces trois missions. […] Outre ces éléments de contexte, nous estimons, d'une part, que les données bibliographiques produites par un établissement public sont un bien public, et d'autre part, que l'acte de numériser un document appartenant au domaine public ne crée pas de droit d'auteur spécifique lié à la version numérique de ce document. Par conséquent, il est de notre mission de service public de rendre ces données et fichiers numériques librement accessibles et réutilisables par les citoyens[103].

Les données bibliographiques comprenant les métadonnées des documents numériques, produites par la BNU ainsi que les fichiers numériques issus d'une numérisation, par la bibliothèque, d'œuvres du domaine public conservées dans ses collections, ont à ce titre été placées sous Licence Ouverte ou autre licence compatible. Plusieurs enjeux sont associés à l'ouverture de ces données : faciliter le travail des chercheurs et les inciter à découvrir et exploiter les fonds. Pour Frédéric Blin, encourager les auteurs et les éditeurs à recourir à des documents de la BNU pour illustrer leurs productions fait partie d'une stratégie de valorisation des fonds et des collections.

Le réseau des médiathèques de Toulouse

Le réseau des médiathèques de Toulouse a également choisi de mettre un pied dans l'*Open data*. Outre la localisation des bibliothèques, médiathèques et bibliobus de la ville accessibles via la plateforme de l'agglomération *ToulouseMetropole.data*[104], la bibliothèque municipale de Toulouse diffuse actuellement sous licence ouverte et via le site de partage de photos *Flickr*, le fonds photographique d'un photographe local du 19e siècle et ancien conservateur

[103] http://alatoisondor.wordpress.com/2012/03/21/il-est-de-notre-mission-de-service-public/ (consulté le 26/11/2012)
[104] http://data.grandtoulouse.fr/les-donnees/-/opendata/card/12402-mediatheques-bibliotheques-et-bibliobus

(cc) BY-NC-ND

du Muséum d'Histoire naturelle de Toulouse, Eugène Trutat[105]. Ces documents n'étaient consultables jusqu'alors que sur certains postes multimédias de la Bibliothèque d'Etude et du Patrimoine. Le portail a ouvert le 26 Juin 2008, avec 310 photos, décrites, nommées et géotaggées. Il est destiné à être enrichi au fur et à mesure des clichés des 5000 photographies sur plaques de verre, des années 1870-1920, concernant Toulouse, la région et les Pyrénées.

Les photos en basse définition sont utilisables sans restriction, tandis que l'usage des photos en haute définition se fait selon la grille tarifaire de la bibliothèque. La démarche a également une dimension participative : enrichir la collection à partir de contributions extérieures permettant de géolocaliser ou de dater certains clichés, autant de métadonnées d'indexation, également rendu disponible sous licence *Open data*. Si ces opérations d'enrichissement on pu donner lieu à quelques commentaires fantaisistes, Jocelyne Deschaux se félicite de l'intérêt que suscitent les photos auprès d'internautes français ou étrangers, et d'une notoriété renforcée pour l'établissement.

> Nous avons le bonheur de recevoir beaucoup de commentaires sur les photos, preuve de l'intérêt qu'elles suscitent, y compris à l'étranger. En ce sens, Flickr nous a aidés à trouver un nouveau public[106].

D'après les premières statistiques, au 1er janvier 2010, le fonds Trutat a enregistré 569 000 vues, soit une augmentation de 341 000 en un an, 78% des photos étaient partagées, 61% ajoutées en favoris, 27% commentées. Sur l'année 2009, ce fonds a fait l'objet de 934 affichages quotidiens, en moyenne. Désormais 31% des internautes qui se connectent à la bibliothèque numérique de Toulouse arrivent via Flickr..

La bibliothèque de Saint-Etienne

Pour Thierry Veyron conservateur du patrimoine en charge de la numérisation à la bibliothèque municipale de Saint-Etienne, dans la Loire, un des enjeux de la numérisation et de l'Open data est de permettre une « parfaite analogie entre le service public du patrimoine en salle et celui sur le net[107] ». C'est la raison pour laquelle la bibliothèque a décidé de mettre en ligne les textes, images et films issus de ses propres numérisations et qui participent de l'originalité, comme le fonds forézien et le fonds industriels de l'Ecole des Mines touchant tous eux au patrimoine local. Aucune mention légale n'accompagne ces documents à l'exception d'une estampille de provenance « Ville de Saint-Etienne ». Tous ces documents sont donc considérés *de facto* comme appartenant au domaine public au même titre que les originaux. Ils sont donc librement téléchargeables et ouverts à une réutilisation souhaitée la plus large possible. À titre d'exemple Thierry Veyron explique que « Si un éditeur nous le demande, nous lui fournirons gratuitement le fichier à haute définition[108] ».

---

[105] Eugène Trutat (1840-1910) a longuement sillonné les Pyrénées et photographié autant les paysages, les villages, que les habitants, livrant ainsi des témoignages vivant sur la vie locale.
[106] http://www.lagazettedescommunes.com/48352/l'exception-culturelle-se-glisse-dans-lere-du-numerique/ (consulté le 26/11/2012)
[107] http://www.lagazettedescommunes.com/48352/l'exception-culturelle-se-glisse-dans-lere-du-numerique/ (consulté le 12/12/12)
[108] Id.

(cc) BY-NC-ND

La Bibliothèque départementale de prêt (BDP) de Saône-et-Loire

La Saône-et-Loire a également rejoint l'*Open data* avec l'ouverture, depuis septembre 201,1 du site *opendata71.fr*. La bibliothèque départementale de prêt y participe en fournissant des données et des contenus de tous type. En ce sens, elle constitue lepremier réseau rural à prendre part au mouvement : Comme l'explique, Jérôme Triaud, conservateur des bibliothèques et directeur de la lecture publique de Saône-et-Loire, ce choix lui paraissant tout aussi stratégique que naturel :

> J'ai pensé que la bibliothèque départementale ne pouvait pas rester à la traîne et qu'elle devait s'engager pleinement dans ce projet départemental. Mettre à disposition les données brutes que nous produisons est le prolongement direct de notre activité de bibliothécaire. Nous collectons des documents, organisons les savoirs, réalisons leur médiation : cela est connu. Mais les bibliothécaires sont de plus en plus conscients qu'ils sont des producteurs de contenus…Et des données nous en produisons au quotidien : les métadonnées de nos catalogues pour commencer, mais aussi nos bibliographies commentées – les fameux coups de cœur – nos dossiers documentaires, nos statistiques etc…Bref, il m'est apparu que nous ne pouvions pas rester en retrait[109].

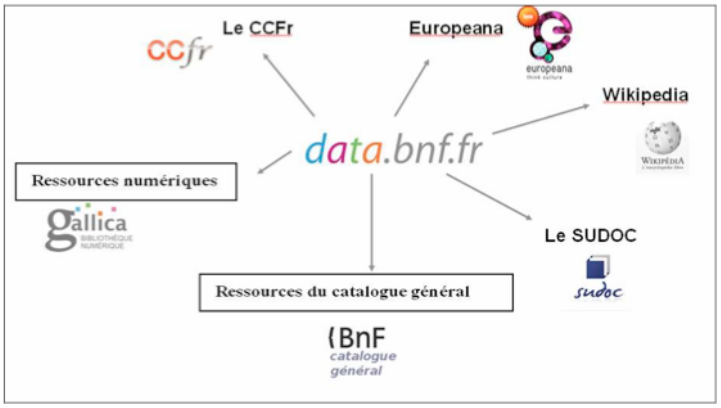
Les données versées sont majoritairement issues du catalogue donc de nature bibliographique mais la BDP a également mis à disposition l'ensemble des statistiques produite à partir de l'activité de son réseau de 130 bibliothèques comme de la cartographie documentaire, des états des prêts, des taux d'inscription, la répartition des inscrits, mais aussi les activités liées à l'animation culturelle ou la formation professionnelle.

Le projet Data.bnf.fr

La bibliothèque nationale de France est devenu un acteur incontournable de l'*Open data* culturel en France grâce à son projet *data.bnf.fr*. Comme on le trouve présenté sur le site de la BnF, ce projet s'inscrit dans le cadre du schéma numérique de la Bibliothèque nationale de France, qui a pour but de rendre les données de la bibliothèque plus visibles sur le Web. En ce sens, le projet répond parfaitement aux objectifs fixés en 2009 par la lettre de mission remise à Bruno Racine par la Ministre de la culture et de la communication, Christine Albanel : il s'agit d'établir un bilan de l'action, des moyens et des besoins des bibliothèques dans le domaine du numérique, étudier les conditions d'une plus grande harmonisation des politiques nationales et d'une meilleure mise en réseau et favoriser le développement de coopérations entre secteurs public et privé. L'idée de *data.bnf.fr* est de rassembler autour de ses pages auteurs, œuvres et thèmes, les ressources de la Bibliothèque nationale de France aujourd'hui dispersées entre différentes bases, ainsi que des ressources extérieures. Mise en ligne en juillet 2011 dans une première version, le site comporte désormais plus de 200 000 pages auteurs, œuvres et thèmes. Ces pages regroupent plus de trois millions de références des catalogues de la BnF (BnF catalogue général, BnF archives et manuscrits) et plus de 140 000 liens vers Gallica.

[109] http://alatoisondor.wordpress.com/category/droit-et-bibliotheques/ (consulté le 20/12/12)

(cc) BY-NC-ND



Source 1 : http://data.bnf.fr/docs/databnf-presentation.pdf

Le projet s'est accompagné d'une longue réflexion et d'un travail important sur les évolutions des modèles bibliographiques, en particulier le modèle FRBR[110]. Les données présentes dans *data.bnf.fr* y sont exprimées selon cette logique de description arborescente. Les pages sont rassemblées autour des entités œuvres, auteurs et thèmes, décrites à des niveaux différents : œuvre, manifestation, expression et *item*. Les auteurs sont reliés aux documents selon leur type de contribution et à différents niveaux : auteur d'une œuvre, traducteur, illustrateur ou préfacier d'une édition, possesseur d'un manuscrit par exemple. Ainsi depuis la page de cicéron, http://data.bnf.fr/11885977/ciceron/ ,on retrouve aussi les ouvrages Grecs qu'il a traduits . Le corpus initial de *data.bnf.fr* porte sur des œuvres et des auteurs majeurs pour lesquels il existe de nombreuses notices dans les différentes bases de la BnF. Il s'élargit progressivement à des œuvres (musicales, anonyme etc.) et des auteurs nouveaux (juristes, compositeurs etc.). Cette extension permet de mettre en valeur les auteurs et œuvres moins connus et moins présents sur le Web. En effet, sur le Web la politique documentaire n'a rien d'absolu, est doit être adaptée en fonction de ce qui existe déjà. En l'occurrence, les ressources disponibles sur les auteurs littéraires majeurs sont déjà abondantes, et c'est dans les domaines plus rares et techniques que la bibliothèque apporte le plus de valeur[111]. L'internaute peut ainsi accéder directement aux documents des catalogues de la BnF ainsi qu'aux documents numérisés de Gallica sans avoir besoin de connaître auparavant les services de la BnF et sans passer par un portail spécifique.

L'application du modèle FRBR a pris corps à travers une expression des données au format RDF[112]. Cette modélisation permet leur exposition selon les standards du web de données[113] afin de garantir l'interopérabilité entre les bases internes et avec l'extérieur, en utilisant des données structurées exploitables par les machines, de façon renouvelée. Le modèle repose sur les adresses des pages

---

[110] *Functional Requirements of Bibliographic Records* - Spécifications fonctionnelles des notices bibliographiques

[111] Ce constat se rapproche de l'idée de « longue traine », utilisée par Chris Anderson en référence à l'essai de Clay Shirky3 qui remarquait qu'un nombre relativement petit de weblogs ont de nombreux liens web pointant vers eux, alors que la longue queue composée de millions de weblogs n'ont que peu de liens qui pointent sur eux.

[112] *Ressource Description Framework* (RDF) : Le RDF n'est ni un format de fichier, ni un schéma XML, ni un langage informatique mais un modèle ou un cadre, c'est-à-dire une organisation théorique et logique de l'information. Ce modèle de graphe, mis au point entre 1999 et 2004, recouvre à la fois une architecture et plusieurs syntaxes, dont l'XML, pour publier des données sur le Web.

[113] Tim Berners-Lee et le W3C établissent quatre règles basées sur les technologies du Web pour y publier des données structurées dans un cadre d'interopérabilité commun : le protocole HTTP, l'identification par les URL et les URI, le principe d'hypertextualité et le langage HTML. Ces quatre éléments forment les pierres angulaires du cadre d'interopérabilité au cœur des enjeux et des technologies du Web de données.

(cc) BY-NC-ND

HTML générées automatiquement à partir des données existantes, accessibles dans les différentes bases de la BnF (BnF catalogue général, BnF archives et manuscrits, Gallica, etc.). Le projet utilise les outils du Web sémantique (distances de Levenstein, Jaro-Winkler etc.) pour l'alignement automatique et le regroupement par œuvres. L'articulation des différents silos s'appuie sur des notices d'autorité qui disposent d'identifiants pérennes : les identifiants ARK[114]. Parmi les autres notices d'autorité qui fondent les pages du site se trouvent les autorités personne et organisation pour les pages auteurs, les autorités titres pour les pages œuvres et les autorités RAMEAU pour les pages thématiques. Les notices d'autorité-sujet du référentiel RAMEAU[115] ont été converties dans le langage RDF SKOS (*Simple Kowledge Organisation System*), dans le cadre du projet européen TELplus. Ce référentiel est désormais maintenu à jour sur *data.bnf.fr* avec la base de données totale et actualisée de la BnF, et fait partie des jeux de données les plus réutilisées du projet.

Les données présentes dans *data.bnf.fr* sont régies par une licence ouverte. Seules les données au format RDF, c'est-à-dire l'ensemble des données d'autorités et des métadonnées descriptives produites par la BnF, c'est-à-dire celles sur lesquelles elles possèdent un droit de propriété intellectuelle[116], sont placées sous Licence ouverte[117]. Ces métadonnées constituent réellement la valeur ajoutée scientifique de la BnF. Elles relèvent de la « licence ouverte » de l'Etat, élaborée par la mission Etalab pour le portail data.gouv.fr. Cette licence est compatible avec les licences Creative Commons :By utilisées pour l'*Open data* au niveau internationale. La licence de l'Etat est utilisée pour la plateforme *data.gouv.fr* où l'on retrouve également mentionnées les données issues de *data.bnf.fr*. La seule obligation du réutilisateur à l'égard de ses données est la mention de la paternité de « l'information », c'est-à-dire faire apparaître la source – au moins le nom du « producteur » - et la date de la dernière mise à jour des données utilisées. Cette mention est une simple reconnaissance de la paternité mais ne peut en aucun cas être considérée comme une caution pour le réutilisateur. Cette mention permet à l'administration de rendre compte de la réutilisation des données de manière objective en retrouvant les tiers qui les utilisent, et cette mention participe à l'effort de publicité et de transparence en aidant les autres utilisateurs à retrouver les données directement à la source. La BnF ne s'engage en rien sur la qualité des données ou le maintien d'une fourniture continue, vis-à-vis du réutilisateur. Elle s'engage en revanche à maintenir un régime ouvert pour les données disponibles, ce qui permet aux utilisateurs de s'en servir avec confiance.

Bien que régie par le droit français, la licence ouverte de la BnF a été conçue pour être compatible avec toute licence libre étrangère qui exige *a minima* la mention de paternité. Les données RDF de *data.bnf.fr* peuvent donc être récupérées et réutilisées librement et gratuitement, selon les besoins, intégralement

---

[114] *Archival Ressource Key* : Système d'identifiants mis en place par la *California Digital Library* (CDL), et qui a vocation à identifier des objets de manière pérenne.

[115] RAMEAU : Répertoire d'Autorité-Matière Encyclopédique et Alphabétique Unifié.

[116] Bibliothèque nationale de France, Licence d'utilisation commerciale, « Art. 3 – Droits de propriété intellectuelle de la BnF : 3.1 - La BnF est titulaire du droit *sui generis* du producteur de la base de données au titre des investissements substantiels, tant quantitatifs que qualitatifs, qu'elle a engagés pour la constitution, la vérification et la présentation du (des) document(s) au sens du code de la propriété intellectuelle ; 3.2 - La BnF est titulaire de droits sur les marques de commerce, de fabrique ou de services apposées sur le(s) Document(s) et notamment sur celles contenues dans les mentions de source et les marquages qui sont associés au(x) Document(s) ; 3.3 - La BnF est titulaire de droits de propriété intellectuelle sur les métadonnées et les données bibliographiques descriptives et d'autorité qu'elle élabore ». http://www.bnf.fr/documents/autorisation_usage_fr.pdf (consulté le 21/10/2012)

[117] http://data.bnf.fr/licence (consulté le 20/10/2012)

(cc) BY-NC-ND

ou partiellement, notamment par des développeurs et des professionnels des bibliothèques en France comme à l'étranger, par le secteur public, comme le secteur privé. Pour les données de Gallica (bilbiothèque numérique), qui sont d'une autre nature, présentes dans un autre format que RDF, la réutilisation est libre et gratuite dans le respect de la législation en vigueur et notamment du maintien de la mention de source "Bibliothèque nationale de France" auprès des données. L'utilisateur peut les adapter ou les modifier, à condition d'en informer clairement les tiers et de ne pas en dénaturer le sens. En revanche la réutilisation commerciale de ces contenus, à savoir l'acquisition des données en vue de l'élaboration d'un produit ou d'un service destiné à être mis à disposition de tiers à titre onéreux, ou à titre gratuit mais avec une finalité commerciale, est payante et fait l'objet d'une licence d'utilisation commerciale de documents hors droits de la BnF applicable aux images, textes et ouvrages. Il est par ailleurs intéressant de constater qu'une part importante des réutilisations des données, notamment RDF est le fait d'entreprises privées de développeurs qui ont besoin, pour créer, tester et faire la démonstration de leurs produits, de disposer de corpus relativement massifs de données structurées et non-sensibles.

A l'étranger

L'*Open data* est une démarche qui l'on retrouve également pratiquée par d'autres grandes institutions culturelles et bibliothèques étrangères. Le monde anglo-saxon s'impose encore une fois comme précurseur dans ce domaine. En novembre 2010, *la British Library* annonçait la libération sous CC0 de plus de trois millions de notices bibliographiques. Depuis fin 2010, la *British Library* met à disposition plus de 3 millions de notices[118]. Cette décision concerne de nombreux ouvrages publiés depuis 1950 au Royaume-Uni et en Irlande. Plusieurs bibliothèques universitaires américaines ont également rejoint cette démarche. L'Université de Yale aux Etats-Unis a annoncé au printemps 2011qu'elle plaçait 250 000 images numérisées issues de ses collections dans le domaine public et qu'elle les libérait de tout droit de réutilisation[119]. Plus récemment, c'est au tour de la Harvard Library d'annoncer, le 25 avril 2012, la mise à disposition de 12 millions de notices bibliographiques en CC0. Enfin à Washington, la *Library of Congress* a lancé un *National Jukebox* facilitant la réutilisation et le partage d'enregistrements sonores numérisés par ses soins[120]. Toujours dans le monde ango-saxon, les institutions culturelles australiennes sont assez actives dans le domaine de l'*Open data*. En effet, rejointe en cela pas d'autres établissements de type archives, musées, la Bibliothèque nationale d'Australie participe à part entière au portail data.gov.au[121]. La majorité de ses données bibliographiques y sont accessibles sous licence CC-BY.

Les institutions culturelles scandinaves sont également des acteurs très impliqués dans ce mouvement et en particulier les institutions suédoises. L'initiative majeure se trouve du côté des musées avec la création d'une base nationale de données, en *Linked Open data,* la *Swedish Open Cultural Heritage*, de 3,4 millions d'objets, issus de musées nationaux et régionaux, accessible sur le

---

[118] http://www.numerama.com/magazine/17420-3-millions-de-notices-bibliographiques-liberees-par-la-british-library.html (consulté le 13/12/12)

[119] Lionel Maurel, Et si la culture s'ouvrait à l'Open data, 25 mai 2011. http://owni.fr/2011/05/25/et-si-la-culture-souvrait-a-lopendata/

[120] http://www.loc.gov/jukebox/ (consulté le 13/12/12)

[121] http://data.gov.au/

(cc) BY-NC-ND

portail Ckan[122]. Ce portail permet également l'accès à une autre base également en *Linked Open data* de 70 000 objets issus des collections du Musée d'Amsterdam. Enfin, outre le musée nordique (*Nordiska Museet*) qui a placé plus de 1000 photographies liées à l'histoire de la Suède dans le domaine public, la Bibliothèque royale de Suède a également fait le choix de publier l'intégralité du catalogue collectif suédois en CC0.

Cette participation des institutions culturelles, et plus particulièrement des bibliothèques, au mouvement d'ouverture des données publiques est une orientation qui semble également se confirmer par les réflexions menées au sein de plusieurs institutions ou organismes de coopération culturelle internationale. En janvier 2011, le groupe de réflexion sur la numérisation du patrimoine européen, autrement connu comme le « Comité des sages[123] » avait déjà tenu à souligner l'intérêt mais aussi certains freins constatés dans la valorisation des données culturelles européennes, notamment quant à l'usage fait des métadonnées :

> 4.1.6 Les institutions culturelles apportent des quantités considérables d'informations complémentaires aux éléments numérisés (des métadonnées), décrivant par exemple l'auteur, la provenance ou l'âge de l'ouvrage, donnant des indications contextuelles ou techniques sur les formats utilisés ou les caractéristiques permettant des outils de recherche à venir localiser l'élément. Ces métadonnées sont essentielles pour fournir aux utilisateurs un arrière plan utile à son travail et aussi pour permettre à des outils de recherche comme *Europeana* de localiser les éléments numérisés pertinents dans le contexte de recherches spécifiques.
>
> 4.1.7 Dans certains cas, les institutions culturelles facturent ou imposent d'autres conditions pour la réutilisation de ces métadonnées et elles ont tendance à être particulièrement méfiantes quant à la réutilisation commerciale des données. Cet usage commercial est largement défini, et peut, par exemple, inclure l'indexation par des outils de recherche commerciaux. Les conditions imposées pour cette utilisation des métadonnées peuvent en retour poser des problèmes pour des services comme *Europeana* qui souhaite ouvrir les métadonnées qu'elles moissonnent afin de leur conférer l'exposition la plus possible[124].

Dans la continuité de ces réflexions, la Conférence des bibliothèques nationales européennes[125] (CENL), organisée en septembre 2011 par la Bibliothèque nationale du Danemark, s'est massivement prononcée en faveur de

---

[122] http://datahub.io/dataset/swedish-open-cultural-heritage

[123] Le "Comité des Sages" a été établi en avril 2010 par Neelie Kroes, vice-présidente de la Commission européenne chargée de la stratégie numérique, et par Androulla Vassiliou, Commissaire chargée de l'éducation et de la culture.

[124] Elisabeth Niggemann, Jacques Decker (de), Maurice Lévy, *The new renaissance,* Report of the « Comité des sages », reflection group on bringing Europe's cultural heritage online, 10 janvier 2011, p. 13. : « 4.1.6 Cultural institutions add considerable amounts of information to digitised objects (metadata), describing for example the author, the provenance and age of the work, giving contextual information, as well as technical information on the formats used and characteristics allowing search engines to locate the object. This metadata is essential to provide the user with a useful background to the work, and also to allow search-engines such as Europeana to locate the digitised objects relevant in the context of a specific search. 4.1.7 In some cases cultural institutions charge for or impose other conditions for the re-use of this metadata and they tend to be particularly wary of the commercial use of the data. This commercial use is broadly defined, and can for example include the indexing by commercial search engines. The conditions imposed on the use of the metadata can in turn pose problems for services such as Europeana, which want to open up the metadata they harvest in order to get the broadest possible exposure. »

[125] Fondé en 1987, la CENL promeut le rôle des bibliothèques nationales en Europe en particulier et s'efforce de renforcer leur responsabilité quant à la préservation des patrimoines culturels nationaux, l'accès à la connaissance et le soutien à la recherche dans ce domaine. Elle se compose traditionnellement de 49 représentants de 46 bibliothèques nationales.

(cc) BY-NC-ND

l'*Open data*[126]. Cette conférence, rassemblant les représentants de 46 bibliothèques nationales européennes, a été l'occasion pour ces établissements, de réaffirmer leur volonté de publier leurs données en *Open data*. La première conséquence serait la publication sous licence ouverte via la bibliothèque européenne TEL[127] des métadonnées versées par les bibliothèques et autres institutions culturelles dans *Europeana*. Elisabeth Niggemann, ancienne présidente du CENL, aujourd'hui remplacée par Bruno Racine expliquait à ce propos que :

> Mettre à disposition des données sous licence ouverte est la clé pour placer nos institutions culturelles comme nos bibliothèques nationales au cœur de l'innovation dans le domaine des applications numériques. C'est le seul moyen pour que la société puisse tirer des bénéfices sociaux et économiques des données que nous avons créées pour enregistrer la production éditoriale des 500 dernières années. La meilleure comparaison est celle entre la bouteille et la conduite d'eau. Plutôt que de mettre en bouteille et d'étiqueter, nous ouvrons le robinet des données de telle sorte que chacun est un accès libre et gratuit et puisse les utiliser à toutes fins utiles[128].

Suite à cette décision, la fondation *Europeana* s'est dotée le 22 septembre 2011 d'un *Data Exchange Agreement*[129], par lequel elle impose à ses fournisseurs de contenus de placer leurs métadonnées sous licence CC0 afin d'en autoriser la libre réutilisation. La concrétisation de cette décision permettrait un accès permanent et un droit de réutilisation pour quiconque et à toute fin possible, des fichiers de métadonnées décrivant les millions d'ouvrages, textes et images, publiés en Europe.

Enfin, il faut également mentionner les travaux menés par le groupe de travail de l'Université de Stanford[130] en juin 2011 pour son rôle dans la promotion de l'utilisation de licences ouvertes auprès de la plupart des bibliothèques nationales évoquées précédemment. Ainsi que les travaux du International Linked Open Data in Libraries, Archives, and Museums Summit[131] ("LOD-LAM") qui s'efforce de promouvoir l'utilisation du *Linked Open data* auprès des grands institutions culturelles de chaque pays, en insistant sur les bénéfices à tirer d'une politique mutualiser entre archives, bibliothèques et musées dans ce domaine. Il s'attache à identifier les outils et les techniques de publication et de travail en *Linked Open data*, étudie les questions de droit et de licence afférent à la publication des métadonnées culturelles et met au point des glossaires et des référentiels de termes relatifs au *Linked Open data* à destination des institutions culturelles.

---

[126] Communiqué de presse, « Europe's national librarians support Open data licencing », 14 ocotobre 2011. http://pro.europeana.eu/web/guest/press-release;jsessionid=56384F9778465230E6132904ED349253?p_p_id=itemsrssportlet_WAR_europeanaportlet_INSTANCE_Gf2y&p_p_lifecycle=1&p_p_state=normal&p_p_mode=view&_itemsrssportlet_WAR_europeanaportlet_INSTANCE_Gf2y_itemId=861950&_itemsrssportlet_WAR_europeanaportlet_INSTANCE_Gf2y_javax.portlet.action=setItemId

[127] TEL : *The European Library*. TEL est un service du CENL.

[128] Elisabeth Niggemann: « Providing data under an open licence is key to putting cultural institutions like our national libraries at the heart of innovations in digital applications. Only that way can society derive full social and economic benefit from the data that we've created to record Europe's published output over the past 500 years. The best analogy is between bottled water and a water main. Rather than bottling it and branding it, we're putting data on tap, so that everyone has free and open access, and can use it for whatever purpose they need ». Cf. Elisabeth Freyre, « Relations étrangères » *Politique(s) et bibliothèques*, Bulletin des bibliothèques de France, t. 56, n°2, 2011.

[129] http://creativecommons.org/weblog/entry/29133 (consulté 15 décembre 2012)

[130] Stanford University. Be part of the Web. Report of the Standfor University workshop, 27 juin-1rst July 2011. http://lib.stanford.edu/files/Stanford_Linked_Data_Workshop_Report_FINAL.pdf

[131] http://lodlam.net/about/

(cc) BY-NC-ND

DES PREDISPOSITIONS ET UN SAVOIR FAIRE A PARTAGER

Comme nous l'avons vu, les bibliothèques s'intéressent de près à l'*Open data* et un certain nombre de projets associés à cette démarche ont déjà vu le jour. Il est vrai que le choix des données, leurs modes de publication et d'exploitation auraient parfois mérité une réflexion et un travail plus approfondi de la part des établissements. Néanmoins, force est de constater, la rapidité des bibliothèques à se positionner sur le terrain de l'*Open data*. Indépendamment de l'intérêt qu'elles peuvent trouver à rejoindre cette démarche, une partie des missions traditionnelles et des pratiques professionnelles des bibliothèques permettent d'expliquer leur réactivité face aux initiatives d'*Open data*.

Partager des données normées et structurées

Coopérer en bibliothèque n'est pas un phénomène nouveau mais un élément structurant de la logique de réseau qui régit leur fonctionnement.

Une dynamique de longue haleine

Cette coopération est d'autant plus intéressante qu'historiquement elle se construit autour de la notion de partage de données, avec l'apparition des premiers catalogues collectifs dès le Moyen Âge[132]. Comme le déclarait en 1909 le bibliothécaire américain C. H. Gould aux membres de la *Library Association*, « il appartient au XXe siècle d'apporter de l'ordre et de la méthode dans les relations entre les bibliothèques plutôt qu'à l'intérieur de celles-ci ». Derrière une démarche humaniste et encyclopédique de catalogage des savoirs, les enjeux de la coopération en bibliothèques sont nombreux : gain de productivité, mutualisation des coûts, économies d'échelle, enrichissement des contenus, partage des savoirs et des savoir-faire etc. La liste est longue et l'introduction de l'informatique en bibliothèque, loin de remettre en cause ces principes, n'a fait que les accentuer en ouvrant de nouvelles perspectives de partage et de communication.

La constitution de réservoirs communs

L'informatisation et l'offre de services bibliographiques ont infléchi les méthodes de travail et d'indexation en faveur d'un catalogage dérivé d'un réservoir bibliographique unique qui permet d'assurer la cohérence des catalogues locaux avec les catalogues nationaux. Les bibliothèques nationales ou réseaux de bibliothèques anglo-saxons comme la *Library of Congress* ou le *Research libraries information network* américain, ont joué un rôle important et précurseur dans la mise en place de ces réservoirs, dont certains ont rapidement acquis une audience mondiale comme l'OCLC (*Online Computer Library Center*), ouvert en 1971 aux USA avec une base de données de 100 millions de notices. En France et en Europe, ces outils ont également connu un développement considérable. C'est notamment la cas avec le déploiement en 2000 du Système interuniversitaire de Documentation (SUDOC) par l'Agence Bibliographique de l'Enseignement Supérieur (ABES) dans les bibliothèques d'enseignement supérieur mais pas uniquement. Grâce à l'acquisition en masse des notices auprès des réservoirs

[132] Il faut néanmoins attendre les assemblées révolutionnaires pour que la première loi de bibliothèque, le premier code de catalogage et la première grande tentative de catalogue collectif national voient le jour, en France.

(cc) BY-NC-ND

bibliographiques internationaux, le SUDOC compte actuellement plus de 7 millions de notices. L'autre grand réservoir national est le Catalogue collectif de France (CCFr) dont la Bibliothèque nationale de France est maîtresse d'ouvrage. Incluant le SUDOC, il comporte également un volet patrimonial et lecture publique important regroupant ainsi autour de 30 millions de notices de revues et d'ouvrages français et étrangers.

Des règles et normes communes de catalogage

L'importance de cette coopération et son développement fulgurant tout au long du XXe ont très rapidement mis à jour la nécessité d'optimiser au mieux le partage de ces informations. Pour répondre à cette urgence, les bibliothécaires ont commencé à réfléchir aux moyens d'établir une cohérence dans le partage de l'information bibliographique, à l'échelle mondiale. Tel a été l'objet de la Conférence internationale des experts en catalogage réunie en 1969 par l'IFLA[133] à Copenhague qui adopta une résolution demandant la création de normes pour harmoniser la forme et le contenu des descriptions bibliographiques établies dans les différents pays. De cette résolution naquit le concept de la description bibliographique internationale normalisée. L'ISBD(G)[134], publié en 1977, définit le schéma général de la description bibliographique internationale normalisée auquel tous les ISBD ont fait référence. Ce schéma précise les éléments requis pour une description bibliographique établie par une agence bibliographique nationale.

Parallèlement à cette démarche, la bibliothèque du Congrès s'est mise à élaborer un format de catalogage informatique, le MARC[135] reprenant à l'identique les règles de catalogage anglo-américaines (AACR[136]), elles mêmes adaptées des ISBD. A chaque zone ISBD correspond ainsi un champ MARC. Ce format servit rapidement de modèle à une quantité d'autres formats, tous parents, définis au rythme de l'informatisation des catalogues de différentes bibliothèques ou de différents réseaux. Tandis qu'au Etats-Unis, le MARC 21 est aujourd'hui le format le plus utilisé en France, la majorité des grands établissements fonctionnent sous UNIMARC à l'exception de la Bibliothèque nationale de France qui travaille en INTERMARC et de quelques autres établissements que leurs SIGB[137] incitent à travailler sous format anglo-saxon. Cet effort de normalisation des pratiques et des outils ne s'arrêtent pas aux formats MARC, il se poursuit au début des années 1990 avec la mise en chantier du rapport sur les FRBR[138], qui conduit à une révision profonde des codes de catalogage à partir des années 2000. Au delà des données bibliographiques, la normalisation s'étend aux notices d'autorités (indexation matière RAMEAU, données d'auteurs, de lieux etc.). Ces différents modèles normatifs gagnent encore en intérêt en raison du mouvement d'insertion des métadonnées documentaires dans un système de *Linked data* vers lequel s'oriente de plus en plus d'institutions culturelles. Nous y reviendrons.

---

[133] *International Federation of Library Associations and institutions* (IFLA)
[134] *International standart for bibliographic description* (ISBD)
[135] *Machine Redeable Cataloguing*(MARC)
[136] *Anglo American Cataloguing Rules* (AACR)
[137] Système intégré de gestion de bibliothèque
[138] *Functional Requirements of Bibliographic Records* - Spécifications fonctionnelles des notices bibliographiques

(cc) BY-NC-ND

Des données faciles à ouvrir

Ainsi, les données culturelles produites par les institutions et particulièrement les bibliothèques possèdent un double avantage dans un contexte politique d'*Open data* : elles sont non seulement riches et nombreuses grâce aux efforts publics consentis par l'Etat et les collectivités pour la numérisation et la mise en valeur du patrimoine national mais elles sont surtout normées et structurées grâce au savoir faire développé par ces établissements en matière d'indexation et de catalogage. De plus, pour une partie d'entre elles, comme les métadonnées descriptives, se sont des informations libres de droits car produites par les bibliothèques et établissements culturels eux-mêmes. Or ces données, qui constituent le cœur des catalogues de bibliothèques et jusqu'à peu, le cœur de métier des bibliothèques, sont une vraie valeur ajoutée pour le public. À la différence du manque à gagner que représenterait pour les musées de France, la disparition des bénéfices engendrés par les reproductions tarifées d'œuvres de la RMN[139], il n'y a pas réellement de marché de la donnée en bibliothèques, en dehors de l'accès aux grands réservoirs de notices de la BnF et du SUDOC. Du point de vue économique, l'ouverture des données de bibliothèques devient, de fait, plus facilement envisageable que celles d'autres secteurs culturels.

Par ailleurs, si du côté des musées, la visualisation et l'accès aux œuvres (souvent encore protégées par le droit de la propriété intellectuelle) constituent la vraie plus value pour le public, en bibliothèque, les données de descriptions bibliographiques et de disponibilité des œuvres restent le premier service attendu par le public et les autres bibliothèques. En conséquence, les métadonnées descriptives sont non seulement des données faciles à ouvrir mais en plus leur ouverture présente un intérêt de service public. En outre, plus sans doute que dans les musées, une part importante des collections de bibliothèques appartiennent désormais au domaine public. C'est particulièrement vrai pour les collections patrimoniales qui, à la différence des données et des documents produits ou conservés par les services d'archives par exemple, ne sont soumises à aucun texte législatif ou réglementaire relatif aux droits, délais et conditions légales d'accès à l'information qu'ils détiennent[140]. Enfin, à l'exception des fichiers lecteurs et de certaines archives du Web conservées au titre du dépôt légal[141], les données produites ou collectées par les bibliothèques ne possèdent pas de dimension sensible ou personnelle qui justifierait un travail, souvent long et laborieux d'épuration des données et de mise en conformité CNIL.

Les bibliothèques médiateur de l'*Open data*

Parmi les grandes questions qui animent actuellement la communauté professionnelle, celle de l'avenir des bibliothèques est devenu une thématique récurrente. Force est bien de constater que les nouvelles technologies de l'informatique et du numérique ont introduit quelques changements dans les pratiques professionnelles traditionnelles. Le *challenge* pour les bibliothèques est

---

[139] Réunion des musées nationaux (RMN).

[140] http://www.archivesdefrance.culture.gouv.fr/archives-publiques/lois/ : Les dispositions législatives qui s'appliquent aux archives constituent le livre II du code du patrimoine, promulgué en 2004 et modifié par la loi sur les archives du 15 juillet 2008. Elles reprennent également des dispositions du code général des collectivités territoriales et du code de commerce relatives à certains types d'archives.

[141] http://www.numerama.com/magazine/21011-la-bnf-va-archiver-le-web-francais-au-titre-du-depot-legal.html (consulté le 18/12/12)

(cc) BY-NC-ND

non seulement de comprendre ce changement, d'en tenir compte mais aussi de s'y adapter, voire idéalement d'en tirer parti. Avec l'augmentation du nombre de ressources disponibles sur Internet, la diminution progressive des fonctions de catalogage, et les chutes de fréquentation constatées dans de nombreux établissements, les bibliothèques ont tout intérêt à développer de nouveaux services si elles veulent continuer à justifier leur utilité. En conséquence, les fonctions d'accompagnement et de formation[142] aux usagers des bibliothèques sont plus que jamais remises au cœur de leurs missions de service public. Dans ce contexte, il est de l'intérêt des bibliothèques de se positionner comme « médiateur » de ce nouveau type de service documentaire que constitue l'*Open data*.

Une première phase de médiation implique une action de formation auprès des professionnels des bibliothèques portant sur les enjeux, les principes et les fonctionnements de l'*Open data* et du *Linked data*. Cet effort de formation et de dissémination professionnelle par le biais de colloques, de publications et de journées d'études en plus des ateliers et des stages organisés par les institutions traditionnelles de formation est essentiel pour cultiver ce savoir faire particulier des bibliothécaires à la gestion, au tri et à la recherche d'informations. De fait, il n'est pas rare de voir des bibliothécaires sollicités à titre de consultants sur des projets d'ouverture de données autres que celles des bibliothèques. Par ailleurs, plusieurs établissements se sont récemment lancés dans la constitution de services tournés vers le monde de l'entreprise et la recherche d'emploi. C'est notamment le cas de l'atelier PRISME[143] à la BnF ou encore du Pôle « Changer sa vie professionnelle, évoluer, valider[144] » de la Cité des Métiers d'Universcience. En raison des objectifs économiques de stimulation de la croissance économique et de l'emploi associé à l'ouverture des données publiques, il semble particulièrement intéressant et pertinent pour les bibliothèques d'intégrer l'*Open data* à leurs plans de formation.

---

[142] Sur le sujet se reporter aux articles d'Elisabeth Ravoux, *La formation des usagers des bibliothèques publiques aux outils informatiques est une nécessité*, 1er septembre 2011, de Claire Mouraby, « Apprendre à transmettre : Tous laissent entendre qu'ils sont peu préparés à cette mission », publié au BBF n°11 juin 2005 et le mémoire d'Enssib de Laurence Rey, *Nouvelles compétences en bibliothèque : profils de poste et plans de formation des personnels au regard de l'évolution des services,* 2010.

[143] http://www.bnf.fr/fr/collections_et_services/droit_economie_politique/a.monde_entreprise.html

[144] http://www.cite-sciences.fr/fr/cite-des-metiers/contenu/c/1239022828079/le-pole-changer-sa-vie-professionnelle-evoluer-valider/

(cc) BY-NC-ND

OPEN DATA ET *LINKED OPEN DATA* : QUELLES PROMESSES POUR LES BIBLIOTHEQUES ?

Comme nous l'avons vu, les bibliothèques possèdent un nombre suffisant d'atouts pour être en mesure de rejoindre rapidement des logique *Open data* voire d'insuffler des projets innovants. Pourtant si, comme le rappelle le CNN, les données culturelles ont « un potentiel de réutilisation particulièrement fort », elles restent toujours peu accessibles et sous exploitées. Une des raisons premières est qu'en tant qu'institutions culturelles, elles ne sont soumises à aucune obligation légale d'ouverture de données publiques. En conséquence pourquoi ouvrir les données culturelles ? Et surtout de quelle manière ?

Dans le domaine culturel plus qu'ailleurs, l'*Open data* n'a de véritable intérêt que conçu en lien avec les technologies du *Linked data*. Il existe, en effet, une forme de continuité technique entre le web de données et l'ouverture des données publiques. Ces deux mouvements s'entretiennent l'un l'autre autour du besoin d'interopérabilité. Le manque d'ouverture technique ou juridique des données se présente comme une entrave au développement et à la qualité du web sémantique, tandis que l'absence de connexions suffisantes entre les données publiques ouvertes limite leur visibilité, leur diffusion et donc leur potentielle réutilisation, alors que c'est précisément le but recherché. En conséquence, le *Linked Open Data* apparaît comme le format idéal pour une meilleure exploitation et une meilleure diffusion des données publiques culturelles.

PROMOUVOIR LA REUTILISATION DES DONNEES PUBLIQUES CULTURELLES

L'absence d'impulsion législative en faveur de l'ouverture n'incitent pas à renforcer leur présence sur les plateformes *Open data,* qui reste encore fort ténue. Pourtant, l'ouverture des données culturelles fait écho à la mission de diffusion, valorisation et démocratisation du Ministère de la culture telle que la définit le décret de 2007[145], relatif aux attributions du Ministère de la culture et de la communication :

> Article 2 : Le ministre de la culture et de la communication, porte-parole du Gouvernement, a pour mission de rendre accessibles au plus grand nombre les œuvres capitales de l'humanité, et d'abord de la France. […] Il veille au développement des industries culturelles. Il contribue au développement des nouvelles technologies de diffusion de la création et du patrimoine culturels. Il met en œuvre, conjointement avec les autres ministres intéressés, les actions de l'Etat destinées à assurer le rayonnement dans le monde de la culture et de la création artistique françaises et de la francophonie. […]

[145] Décret n° 2007-994 du 25 mai 2007 relatif aux attributions du ministre de la culture et de la communication, porte-parole du Gouvernement, JORF n°121 du 26 mai 2007 page 9881, texte n° 22. http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000461196&dateTexte=&categorieLien=id (consulté le 28/10/2012)

(cc) BY-NC-ND

Ouvrir ses données c'est, dans un sens, contribuer à la démocratisation de la culture[146], à la valorisation du patrimoine national et du domaine public et c'est aussi donner une image plus ouverte des institutions culturelles françaises encore trop souvent considérées comme jalouses de leur prestige et de leur collections.

Quelles sont les données publiques culturelles ?

Une grande partie du patrimoine culturel français est aujourd'hui conservée et gérée par des établissements publics tels que les musées, les bibliothèques, les instituts et centres culturels nationaux ou encore les archives qui détiennent et produisent d'importantes quantités de documents et de données issus de leur activité propre (horaires, programmation, budgets, statistiques etc.) mais surtout de la numérisation (textes et images) et du référencement (bases de données, métadonnées, notices d'œuvres etc.). Avant de poursuivre, il semble important de définir dès maintenant le périmètre défini par le terme « données publiques culturelles ». Une première catégorie de données concerne les reproductions d'œuvres numérisées ou sous format numérique natif, elles-mêmes tombées dans le domaine public. Si l'on s'appuie sur la définition donnée par le groupe de travail sur le patrimoine numérisé[147], elles regroupent l'ensemble des livres et imprimés, des documents photographiques, audiovisuels, sonores, des données topographiques, des documents iconographiques divers. Cet ensemble comprend aussi les documents d'archives, qui ne sont pas réductibles aux catégories précédentes, car ils sont pour la plupart versés par d'autres services publics, et sont pour certains, gérés conjointement avec des collectivités territoriales. Pour ces documents, le Ministère est à la fois un gestionnaire, responsable du versement, de la conservation et de la communication, et un producteur de normes. Une seconde catégorie concerne les données relatives à ces œuvres c'est-à-dire les catalogues, les bases de données descriptives ou tout autre type de métadonnées attachées à la description, qualification ou enrichissement de l'œuvre.

Cependant définir le caractère public d'une donnée culturelle peut se révéler chose complexe, notamment dans le secteur de la documentation. Avant même de s'attarder sur le régime dérogatoire que leur octroie la loi de 1978, d'autres ambiguïtés persistent autour de la complexité du droit d'auteur et de la propriété intellectuelle. Un premier point concerne les œuvres numérisées. Selon le Ministère de la culture et de la communication, les opérations purement mécaniques de numérisation de documents « ne confèrent à la bibliothèque aucun droit de propriété littéraire et artistique sur les œuvres ainsi reproduites[148] ». Pourtant il est encore fréquent que des établissements culturels comme la BnF, les archives nationales, la RMN et des services départementaux d'archives soumettent à autorisation préalable voire parfois à redevance (la réutilisation de données ou de

[146] C'est un argument revendiqué par exemple par l'Association des archivistes français (AAF) qui rappelle dans son communiqué du 6 juillet que « Grâce à l'investissement financier et humain des conseils généraux et de l'État, une soixantaine de départements proposent aujourd'hui en ligne chacun en moyenne 1,5 million de pages numérisées et plus de 90% d'entre eux le font gratuitement, dans une perspective de démocratisation culturelle ». http://www.archivistes.org/La-reutilisation-des-donnees (consulté le 12/12/12)

[147] « Partager notre patrimoine culturel » - *Propositions pour une charte de la diffusion et de la réutilisation des données publiques culturelles numériques*, Rapport Boudy-Ory-Lavollé, remis en mai 2010 à la Ministre de la culture et de la communication, La documentation française, 2009, 65 p. : La mission confiée au groupe de travail portait sur la réutilisation, c'est-à-dire l'usage par des tiers d'informations numérisées produites ou détenues par des services administratifs culturels dans le cadre d'une mission de service public. Le rapport présent les orientations du groupe pour la création d'une charte en faveur de la mise à disposition des contenus culturels numériques.

[148] http://www.culture.gouv.fr/culture/mrt/numerisation/fr/dll/juridi.html#06

(cc) BY-NC-ND

documents numérisés par leur soin alors même que les originaux sont tombés dans le domaine public. Pour les défenseurs de l'*Open data*, ces pratiques, permises par la loi, portent atteinte au domaine public en faisant renaître une couche de droit supplémentaire après l'expiration des droits patrimoniaux[149].

Enfin le problème du droit d'auteur des agents publics sur les données qu'ils produisent soulève encore quelques incertitudes. Cela concernent notamment certaines métadonnées ou certaines notices d'œuvres. Le titre II de la loi DADVSI[150] du 1er août 2006 apporte un premier élément de réponse mais laisse régner un flou quant à la réutilisation des données par un tiers[151]. En l'absence de textes clairs et précis sur les conditions de réutilisation des données liées à ces cas particuliers, des ambiguïtés demeurent et peuvent donc être invoquées par les institutions productrices comme frein à l'ouverture des données culturelles. C'est la raison pour laquelle, dans son avis n°12 du 5 juin 2012, le CNN recommande dans sa proposition n°3 « d'intégrer les données culturelles relatives à la réutilisation commune[152] ».

Revenir sur l'exception culturelle française

L'ouverture des données culturelles est une question récurrente au sein de la communauté *Open data*. D'après le rapport Ory-Lavollé, il s'agit bien d'une des missions du Ministère de la culture :

> Un principe général doit guider l'action du Ministère et de ses établissements en matière de réutilisation des données publiques : maximiser la présence, la fréquentation et l'appropriation des données publiques culturelles sur les réseaux. Pour la mettre en œuvre, il faut conclure le plus grand nombre possible de licences permettant la réutilisation des données publiques culturelles sur des sites tiers de toute nature et leur appropriation par les particuliers dans le cadre d'usages personnels et collaboratifs[153].

Pourtant l'article 11 de la loi CADA du 17 juillet 1978 repris par la directive européenne de 2003 exclut toujours du champ légal de la réutilisation des données

---

[149] Comme le souligne le groupe Open Glam : ces pratiques ont été déplorées par la Commission européenne et vont à l'encontre de manifeste de document comme le Manifeste du Domaine public du réseau Communia ou la Charte du domaine public d'*Europeana*. http://www.donneeslibres.info/openglamFR.pdf (consulté le 05/11/2012)

[150] Titre II « Droits d'auteurs des agents de l'Etat, des collectivités territoriales et des établissements publics à caractère administratif » de la loi n°2006-291 du 1er août 2006 relative au droit d'auteur et aux droits voisins dans la société de l'information modifiant l'article L-131-3-1 du code de la propriété intellectuelle, JORF n° 178, 3 août 2006, p. 11529 : « Dans la mesure strictement nécessaire à l'accomplissement d'une mission de service public, le droit d'exploitation d'une œuvre créée par un agent de l'Etat dans l'exercice de ses fonctions ou d'après les instructions reçues est, dès la création, cédé de plein droit à l'Etat ».
http://www.legifrance.gouv.fr/affichTexte.do;jsessionid=55585BBAFAB2162DCB8CD7B0E9C5AE5A.tpdjo03v_1?cidTexte=JORFTEXT000000266350&categorieLien=id#JORFSCTA000000900120 (consulté le 02/10/2012)

[151] Avis de la CADA n°20092706 du 5 novembre 2009 : « Leur réutilisation par les tiers (autres que la collectivité employeur), qu'aucune disposition de portée générale n'interdit par ailleurs, n'est pas régie par le chapitre II du titre Ier de cette loi, mais par les règles de droit commun fixées par le code de la propriété intellectuelle, que la commission n'a pas reçu compétence pour interpréter ». Concernant les données culturelles, une des recommandations du groupe Open Glam est justement de revenir sur la loi DADVSI afin que la cession des droits des agents publics soit complète et s'étende aux usages commerciaux comme aux usages extérieurs à l'accomplissement d'une mission de service public.
http://www.cada.fr/conseil-20092706,20092706.html (consulté le 02/10/2012)

[152] Conseil national du numérique (CNN), *Avis n° 12 du relatif à l'ouverture des données publiques* (« Open data »), 5 juin 2012, p. 10.

[153] « Partager notre patrimoine culturel », *Propositions pour une charte de la diffusion et de la réutilisation des données publiques culturelles numériques*, Conclusions du groupe de travail sur le patrimoine culturel numérisé remises en mai 2010 à Madame la Ministre de la culture et de la communication, 2009, p. 5
http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/104000652/0000.pdf
(consulté le 28/10/2012)

(cc) BY-NC-ND

publiques[154] : les établissements et institutions d'enseignement et de recherche (les écoles, les universités, les bibliothèques et les instituts de recherche) et les établissements, organismes ou services culturels (les musées, les bibliothèques, les orchestres, les opéras, ballets et théâtres, les services d'archives). Ces établissements ou services dont font partie les bibliothèques restent libres d'élaborer leurs propres règles de réutilisation « le cas échéant », précise l'article 11, « soit par règlement, soit par voie conventionnelle, dans le respect des législations transversales et de la jurisprudence ». Quelques précédents ont déjà permis de créer un levier en faveur d'une révision de ce régime dérogatoire. L'exemple le plus marquant est celui du conflit opposant le site *Notrefamille.com* aux archives départementales du Cantal[155]. Si le Tribunal administratif de Clermont-Ferrand a estimé que le régime dérogatoire ne permettait pas pour autant aux archives de s'opposer au principe même de la réutilisation, selon le CNN, celui-ci constitue toujours un frein à la mise à disposition des données culturelles.

Plusieurs groupes réfléchissent ou militent aujourd'hui pour un retour sur l'exception culturelle française. Dans la continuité du rapport Boudy-Ory-Lavollé de 2010, la Commission européenne a nommé, dans le cadre du réexamen en 2008 de la directive PSI de 2003[156] un « Comité des sages » chargé d'approfondir la question de la mise en ligne du patrimoine culturel européen[157]. Ce dernier a remis début 2011 un rapport allant fortement dans le sens d'une plus large ouverture à la réutilisation des données culturelles, y compris en France. A l'échelle nationale, le groupe Wikimedia France et l'*Open Knowledge Fondation* ont lancé conjointement un groupe de travail baptisé « *Open Glam*[158] ». Composé sur la base du volontariat de juristes, de professionnels de la culture et de membres d'associations, il se présente lui-même comme un « réseau informel de personnes et d'organisations cherchant à favoriser l'ouverture des contenus conservés ou produits par les institutions culturelles ». A la suite de réunions tenues à Paris entre le 27 avril et le 29 mai 2012, le groupe a rédigé, sous forme de rapport, une série de « Recommandations pour l'ouverture des données et des contenus culturels[159] » tirées des conclusions de leurs réflexions préparatoires. Selon eux, l'ouverture des données culturelles implique une triple coopération entre le Ministère de la Culture, les institutions culturelles et enfin le législateur.

---

[154] http://www.cada.fr/les-etablissements-culturels-ou-d-enseignement,6127.html (consulté le 28/10/2012)

[155] Le Jugement du Tribunal de Clermont-Ferrand rendu le 13 juillet 2011 sur le cas opposant le site de généalogie notrefamille.com au département du Cantal sur la fourniture de registres d'Etat civil numérisés à des fins de réutilisation commerciale crée une jurisprudence en faveur l'ouverture des données. Il lève une ambiguïté autour de l'interprétation du régime dérogatoire en précisant que les services d'archives publics, qui constituent des établissements culturels, « ne disposent pas d'un pouvoir discrétionnaire leur permettant d'apprécier l'opportunité de faire droit ou non à une demande de réutilisation », mais peuvent « encadrer cette réutilisation (...) afin de sécuriser toutes les formes de réutilisation et, d'autre part, de s'opposer aux demandes présentant un caractère abusif ». Cette décision a été pleinement confirmée par un arrêt de la Cour Administrative d'Appel de Lyon en date du 4 juillet 2012.
http://dl.dropbox.com/u/17624307/Trib%20admin%20Clermont%20Ferrand%20Notre%20Famille%20Cantal.pdf
(consulté le 28/10/2010)

[156] Directive européenne 2003/98/CE dite PSI (*Public sector informations*) concernant la réutilisation des informations du secteur public.

[157] « Reflection group on bringing Europe's cultural heritage online, 2011, « The New Renaissance : Report on the « Comité des sages » ».
http://ec.europa.eu/information_society/activities/digital_libraries/doc/refgroup/final_report_cds.pdf (consulté le 24/09/2012)

[158] GLAM : Galleries, Librairies, Archives, Museums.

[159] B. Jean, L. Maurel, P. de Filippi, A. Charmet-Alix, *Recommandations pour l'ouverture des données et des contenus culturels*, «Groupe Open Glam », Wikimedia France, 2012.

(cc) BY-NC-ND

Quelles conséquences économiques pour le secteur culturel ?

Si l'*Open data* se veut un mouvement créateur de valeur et de croissance, la question des retombées économiques pour le secteur culturel vaut le coup d'être posée de manière distincte. Compte tenu du régime dérogatoire dont bénéficient encore les établissements culturels en matière d'ouverture des données, il ne semble pas inutile de s'interroger sur la réalité des revenus qu'ils peuvent espérer, au regard des autres bénéfices qu'ils pourraient retirer d'une politique d'*Open data*. En l'absence d'une dynamique globale de valorisation des données culturelles, de nombreux freins économiques peuvent être ressentis par les institutions au moment d'envisager une démarche d'*Open data* : tri des données, gestion des exceptions, coût des serveurs, formatage des données, respect des standards d'interopérabilité, plateforme de mise en valeur, clarification des droit, contrôle des licences etc.

C'est pourquoi, outre le souci de couvrir ces frais de mise en service, il est tentant pour les établissements du secteur culturel de chercher à commercialiser la plus-value tirée du travail d'enrichissement des contenus et des opérations de numérisation afin d'en dégager d'avantage de ressources financières. La loi CADA précise que les administrations sont libres de fixer des redevances, qui tiennent compte des coûts de mise à disposition des données. Une marge de manœuvre est donc laissée à l'administration pour la tarification. Mais comme le fait remarquer le CNN ces redevances peuvent parfois entraver leur diffusion et l'innovation en créant des barrières d'accès, en favorisant les gros acteurs et encourageant la constitution d'oligopoles. La question à se poser est donc celle formulée par la Fing dans 2011 dans son *Guide pratique de l'ouverture des données publiques territoriales* : « Peut-on gagner de l'argent avec les données culturelles du territoire ? » et surtout de quelle manière ?

Pour Danielle Bourlange directrice-adjointe de l'APIE[160] l'ouverture des données publiques ne doit pas être nécessairement synonyme de gratuité. Une logique commerciale modérée constituerait même les bases saines d'un modèle économique fondé sur un cercle vertueux :

> Si une donnée a une valeur économique, les ré-utilisateurs commerciaux sont prêts à payer le juste prix. Nous préconisons une approche pragmatique et réaliste en matière de tarification. Ce n'est ni une logique du tout gratuit, ni une approche du tout payant. […] A partir du moment où il y a une réutilisation commerciale de ces informations publiques, il est envisageable qu'une partie de la valeur commercialisée revienne à l'administration à travers une redevance équitable. Ceci est de nature à enclencher un cercle vertueux, car dans un contexte budgétaire contraint, ces ressources supplémentaires pour les administrations permettent d'améliorer la qualité des informations produites, et la qualité des conditions de mise à disposition. Avec un bénéfice pour tout le monde : pour l'usager, qui va avoir de meilleures informations, et pour le monde économique qui va incorporer des contenus de meilleure qualité ou mis à disposition dans de meilleures conditions[161].

[160] Agence pour le patrimoine immatériel de l'État (APIE).
[161] Interview de Danielle Bourlange directrice adjointe de l'APIE, dans la *Gazette des Communes*, édition en ligne. http://www.lagazettedescommunes.com/48347/interview-de-danielle-bourlanges-directrice-generale-adjointe-de-l'apie/ (consulté le 02/11/2012)

(cc) BY-NC-ND

De plus, d'un point de vue microéconomique, un modèle fondé sur la gratuité suscite quelques inquiétudes auprès des secteurs, comme les bibliothèques, dont certains revenus reposent sur une logique de *business to business*. Le commerce de données ou de services bibliographiques repose majoritairement sur une économie inter-bibliothèques, or pour Silvère Mercier le risque serait de voir ce marché se privatiser au détriment des institutions publiques, conduisant ainsi les bibliothèques à se retrouver en position de « clients publics d'entreprises qui revendent des données publiques[162] ». Ce sont effectivement des dérives que l'on a pu constater dans le domaine de la documentation et de l'*Open access* comme en témoigne le récent scandale Refdoc/INIST[163]. A l'inverse, la vente de données publiques pose toujours le problème de la double tarification de l'information pour le secteur public. Parallèlement à l'*Open data*, les projets développés autour des licences nationales ou du *Gold open access* témoigne de réflexion sérieuse autour de cette question.

D'un point de vue macroéconomique, le coût initial de la numérisation et de la mise à disposition des données et contenus culturels peut être compensé par une offre de service ou des bénéfices obtenus sur d'autres plans. Comme cherche à le démontrer le CNN, rejoint en cela par les associations citoyennes[164], du strict point de vue économique et financier, la gratuité peut s'imposer comme la solution la plus rationnelle. Non seulement parce que certains établissements considèrent que cela rentre dans le cadre d'une mission de service public, pour laquelle les contribuables ont consenti un effort, mais surtout parce que les coûts de transaction fixes (personnel, facturation, suivi du paiement etc.) peuvent s'avérer supérieurs au montant de la redevance généralement destinée à couvrir les frais de mise en ligne[165]. Tel est en tout cas le constat et le pari fait par Frédéric Blin pour la BNU avant de franchir le cap de l'*Open data* :

> Les sommes récoltées par la BNU chaque année au titre de la redevance d'usage étaient minimes, de l'ordre de 3000€. Elles ne couvraient naturellement pas le temps de travail de la secrétaire chargée de gérer les factures et la correspondance avec les lecteurs, ni le temps des autres personnes – y compris de l'Administrateur – impliquées en cas de demande d'exonération ponctuelle ou systématique. En outre, nous espérons que l'abandon de la redevance d'usage entrainera une augmentation des demandes de numérisation de documents, service qui lui restera payant. Dans notre cas

---

[162]Silvère Mercier, « Bibliothèques publiques et Open data : quels enjeux ? », *Bibliobsession*, 7 juin 2011 : dans le cas présent Silvère Mercier fait un parallèle entre les perspectives pour les bibliothèques et ce qui s'est passé selon lui autour des réseaux autoroutiers dans les années 2000 : « au public l'investissement et au privé l'exploitation commerciale et la rentabilité sans contrepartie ». Une réflexion que s'efforce de nuancer le rapport Ory-Lavollé « Partageons notre patrimoine » avec la théorie néo-classique des externalités.
http://www.bibliobsession.net/2011/06/07/bibliotheques-publiques-et-donnees-ouvertes/(consulté le 11/09/2012)

[163] Depuis octobre 2012, une polémique a éclaté dans le monde de la recherche autour des pratiques documentaires de l'Institut national de l'information scientifique et technique (INIST). Il s'est avéré que l'INIST commercialise via son service de documentation RefDoc, des copies d'articles sans l'autorisation de leurs auteurs et sans même les avertir alors qu'aucun droit d'auteur ne leur est reversé et que beaucoup d'articles vendus sont déjà disponibles gratuitement sur Internet, sur des plateformes en *open access*. Une pratique qui pose certaines questions éthiques autour de la libre réutilisation.

[164] Association LibertTic, « Pourquoi la réutilisation des données publiques à des fins commerciales doit-elle être gratuite ? », 25 février 2011.
http://libertic.wordpress.com/2011/02/25/pourquoi-la-reutilisation-des-donnees-publiques-a-des-fins-commerciales-doit-etre-gratuite/

[165] Ce type de raisonnement coût/avantage est de plus en plus fréquent dans les bibliothèques à propos de service comme le PEB, de moins en moins utilisé et dont la facturation devient parfois plus contraignante que le bénéfice que l'établissement peut espérer en tirer.

(cc) BY-NC-ND

> particulier, nous pensons qu'en autorisant la libre réutilisation, l'établissement sera au final bénéficiaire au strict plan financier[166].

Un premier pas en faveur d'une dynamique axée sur la gratuité a été fait depuis le 30 septembre 2005[167] avec un décret soumettant toute nouvelle redevance postérieure au 1er juillet 2012[168] à la publication d'un décret, après avis consultatif du COEPIA[169]. Si le principe de redevance est accordé, le calcul doit être justifié et viser à une réutilisation maximale des données. De plus, l'*Open data* n'interdit pas aux établissements de chercher à développer eux-mêmes des applications ou de nouveaux services payants à partir des données mises en ligne. Enfin si les retombées économiques ne s'imposent pas immédiatement pour les établissements, dans une logique keynésienne, ils peuvent attendre de l'*Open data* des externalités positives susceptibles de générer des rentrées d'impôts et de nouvelles cotisations. Il reste néanmoins certain qu'embrasser une démarche *Open data*, c'est accepter, pour le secteur culturel, de laisser la possibilité à une concurrence privée de se développer sur un marché jusqu'ici réservé à la sphère publique.

AMELIORER L'OFFRE DE SERVICES

L'*Open data* offre la possibilité aux bibliothèques de toucher des usagers en dehors de son cercle traditionnel de lecteurs et de valoriser, l'échelle du web, le contenu de sa plus-value scientifique. Utiliser à bon escient et dans la limite des moyens à disposition des établissements, l'*Open data* peut devenir un facteur d'amélioration des services offerts par les établissements culturels, comme les bibliothèques, au public.

Améliorer la recherche des utilisateurs

L'utilisation couplée du *Linked data* et l'Open data ouvre des horizons nouveaux pour le monde de la recherche et de la navigation sur le Web. A l'occasion du lancement en France du nouveau *Graph of Knowledge* du moteur de recherche Google, Amit Singhal, vice-président de Google rappelle justement les perspectives attendues de ces nouvelles technologies pour les utilisateurs :

> Nous espérons que cette intelligence de notre moteur aidera à découvrir encore plus les choses qui vous intéressent et piquera votre curiosité pour des sujets nouveaux. Nous sommes fiers, au travers du *Knowledge Graphe*, de faire notre premier (tout petit) pas vers une recherche plus intelligente, et plus proche de "l'ordinateur de Star Trek" que j'ai toujours rêvé de construire. Nous espérons qu'ainsi, le moteur de recherche de Google pourra vous simplifier les recherches et les découvertes qui sont les vôtres tout au long de

[166] Interview de Frédéric Blin, directeur de la conservation et du patrimoine de la BNU à Strasbourg par Rémi Mathis, 21 mars 2012. http://alatoisondor.wordpress.com/2012/03/21/il-est-de-notre-mission-de-service-public/ (consulté le 26/11/2012).

[167] Décret n°2005-1755 du 30 septembre 2005 relatif à la liberté d'accès aux documents administratifs et à la réutilisation des informations publiques, pris pour application de la loi 78-753 du 17 juillet 1978 ; modifié par décret n°2011-577 du 26 mai 2011.

[168] Circulaire du 26 mai 2011 : elle fixe le principe pour les administrations centrales (ministères et établissements publics sous tutelle) d'autoriser la réutilisation gratuite de leur données à partir du 1er juillet 2012. En conséquence toute fixation de redevance ne peut se faire que par décret après avis du Conseil d'orientation de l'édition publique et de l'information administrative (COEPIA).

[169] Conseil d'orientation de l'édition publique et de l'information administrative (COEPIA).

(cc) BY-NC-ND

votre vie, et vous permettre de passer moins de temps à chercher pour en passer plus à faire ce que vous aimez[170].

Rendre les contenus plus visibles et accessibles pour les utilisateurs

Comme l'explique Emmanuelle Bermès dans un article de la revue *Arabesque*[171], produire et échanger en ligne des données structurées et normalisées, est une des vocations premières du web sémantique. Exprimé comme tel, celui-ci se présente comme une réponse toute indiquée aux besoins des bibliothèques. Or encore aujourd'hui, les catalogues de bibliothèques sont, en général, entreposés dans des bases de données relationnelles tout simplement inexploitables par des moteurs de recherche. Les utilisateurs ne peuvent accéder à ces données que par le biais des portails de bibliothèques, dont ils ne connaissent pas forcément l'existence. Dans la pratique, les internautes ont donc peu de chance de trouver les ressources à partir d'une interface de moteur de recherche, à moins de déjà connaître l'institution et ses outils.

En permettant de dépasser un stockage des données en « silos »[172], le *Linked Open data* autorise la création d'un espace d'information et de recherche homogène et unifié. Il permet de décloisonner les outils de recherche, aussi spécialisés soient-ils. Il offre des qualités de recherche améliorées pour l'internaute. Des études ont en effet montré que les démarches de recherche sur internet ne sont généralement pas centrées sur les institutions mais bien sur les contenus. En permettant aux moteurs de recherche de navigateur d'interroger directement les contenus des catalogues de bibliothèques, le *Linked Open data,* permet de les faire remonter immédiatement dans les résultats auxquels accèdent prioritairement les internautes. Il offre ainsi aux collections des bibliothèques ou de tout autre institution culturelle un meilleur signalement et la possibilité de toucher un public autre que celui fréquentant traditionnellement la bibliothèque. Comme l'explique la Bnf au sujet de son programme *data.bnf.fr* avec le web de données « Nous voulons donc que les ressources de la BnF soient aussi visibles sur le web que la bibliothèque dans la ville[173] ».

De plus, par l'*Open data*, les données culturelles peuvent être relayées par des acteurs commerciaux ou communautaires proposant des services initialement à vocation non culturels comme GoogleMap, ViaMichelin ou d'autres sites fondés sur la géolocalisation pour ne citer qu'un exemple. En effet, une quantité considérable de données bibliographiques partagent des informations communes ou convergentes avec d'autres types de ressources sur le web tels que des dates, des noms de personnes, d'organisations ou encore des informations géographiques. A l'avenir le *Linked Open data* devraient permettre à tous ces points communs de s'interconnecter. Dans cette logique, le géo-référencement des données publiques culturelles est déjà une stratégie de valorisation adoptée par de grands établissements comme la BnF. Associer à chaque objet des coordonnées

---

[170] http://googlefrance.blogspot.fr/2012/12/lancement-en-france-du-knowledge-graph.html (consulté le 28/12/12)

[171] Emmanuelle Bermès, « Vers le Web sémantique ? Les bibliothèques face aux normes du web », *Arabesque*, n°62, juin 2011.

[172] Fabien Gandon, Catherine Faron-Zucker et Olivier Corby, *Le Web sémantique : comment lier les données et les schémas sur le web ?*, Paris, Dunod, coll.« InfoPro. Management des systèmes d'information », 2012, vol. 1/ et Hind Bouchareb, Aurore Cartier, Julie Garel-Grislin, Alicia Leon y Barella, Lucie Leprevost, Fanny Mion-Mouton, Cécile Quach, Web sémantique, web de données : quels intérêts pour les bibliothèques ?, travail d'étude DCB21, Enssib, 2012. http://websmantique.unblog.fr/

[173] Bibliothèque nationale de France, *Présentation générale du projet data.bnf.fr*, 15 novembre 2011, p.2. http://data.bnf.fr/docs/databnf-presentation.pdf (consulté le 20/10/2012)

(cc) BY-NC-ND

géographiques comme lieu de naissance de l'artiste, le lieu de la scène représentée, le lieu de création, d'exposition, de diffusion ou de conservation etc., permet de multiplier les occasions de réutilisation en augmentant les chances d'être repérées sur l'internet. Proposer de telle information sous forme de données ouvertes ouvre des possibilités de développement d'applications mobiles très intéressantes, notamment pour le secteur du tourisme et de la culture, mais pas uniquement.

Améliorer son image et sa notoriété

Faire le choix de rendre accessibles en ligne et réutilisables des fonds documentaires ou des données scientifiques et techniques parfois exposées mais souvent méconnus et sous-utilisés peut considérablement contribuer à l'accroissement da la notoriété et de la visibilité d'un territoire ou d'une l'institution. De plus, comme le constate Bernadette Kessler, « indépendamment des retombées économiques, les institutions et les collectivités ont beaucoup à gagner de l'Open data en terme d'images[174] ». Comme le confirme le rapport Boudy-Ory-Lavollé, mettre en ligne des données bien exposées et bien référencées est donc un biais pour stimuler les réutilisations et améliorer les remontées dans les moteurs de recherche qui accroissent la notoriété des données comme celle de l'institution culturelle productrice. Encourager la fréquentation des sites en ligne peut aussi être un moyen d'encourager la fréquentation physique des lieux, des œuvres et des documents publics. En retour, cette double fréquentation permet à d'accroître le désir et les opportunités de réutilisation des œuvres numérisées.

ENRICHIR LES CATALOGUES

Compte tenu de leur vocation à partager et communiquer des données, les bibliothèques s'intéressent depuis longtemps aux bénéfices des technologies issus du *linked Open data* et peuvent prétendre figurer parmi les établissements publics les plus investis dans cette démarche. Si le *linked Open data* présente un intérêt évident pour les utilisateurs en terme de recherche et de traitement de l'information, il possède également des avantages directs pour les professionnels des bibliothèques.

Mutualiser des acquis, agréger des ressources

Dans le strict domaine de la bibliothéconomie et du catalogage, les données liées ouvertes permettent aux bibliothèques de mutualiser d'avantage leurs outils et leur production, dans la mesure où les ressources peuvent y être décrites en collaboration avec d'autres bibliothèques et liées à des données produites par d'autres établissements ou communautés. Dans le cadre de la production de notices bibliographiques ou d'autorité, il devient possible d'envisager la création d'une réserve mondiale de données partagées qui pourrait être utilisée et réutilisée pour la description des ressources, réduisant ainsi le travail redondant, inhérent aux processus actuels de catalogage. Plusieurs décisions abondent déjà en ce sens, comme l'annonce par l'OCLC en avril 2012 par OCLC de la publication des données du *WorldCat* sous licence *Open data Commons* (ODC). Comme l'explique Richard Wallis, « après de nombreuses recherches et discussions il est apparu

[174] Entretien avec Bernadette Kessler.

(cc) BY-NC-ND

clairement que l'ODC-BY était le meilleur choix de licence pour beaucoup de services de données de l'OCLC. La recommandation faites aux établissements membres d'adopter également cette stratégie claire et cohérente de licence ouverte, dérivée du *WordCat*, découle naturellement de cette expérience[175] ».

Ces décisions sont à mettre en regard du projet VIAF[176] piloté par OCLC. Ce projet regroupe des autorités produites par 27 bibliothèques nationales, rediffusées gratuitement en RDF et sous licence ouverte. C'est un exemple de valeur ajoutée de *Linked Open data* car chaque Bibliothèque nationale étant la meilleure dans sa langue nationale il est de l'intérêt de tous de pouvoir récupérer les données produites par autres. Personne n'aurait à gagner à une tarifier ces données dans la mesure où la vraie plus-value scientifiques de ce système vient des partenaires les plus marginaux et souvent les moins riches : petits pays, langues rares etc.). Grâce au partage des identifiants, les établissements pourront assembler des descriptions à partir de ressources qui se trouvent hors de leur propre domaine, dans tous les jeux de données du patrimoine culturel, et même dans l'ensemble du web. Les catalogueurs pourront alors concentrer leurs efforts sur leur domaine d'expertise locale, plutôt que d'avoir à recréer des descriptions déjà faites par d'autres. Différents types de données sur le même fonds peuvent être produits de manière décentralisée par différents acteurs, puis regroupés dans un seul graphe. C'est donc un moyen pour les institutions d'optimiser les processus internes de traitements des données.

Pour l'*OKFN Working Group on Open Bibliographic Data,* le *Linked Open data* est donc un outil privilégié pour stimuler la coopération entre groupes et établissements au sein et au-delà du monde des bibliothèques. De cette sorte, « les bibliothèques seraient parmi les premières institutions à tirer profit de la libéralisation des données bibliographiques[177] ». Outre la poursuite logique des idées de coopération et de partage qui ont existé de manière constante entre bibliothèques et réseaux de bibliothèques, elle offre par ailleurs des opportunités pour élargir le champ des usages possibles par d'autres institutions comme *Wikipedia* ou encore l'*Internet Archive*[178] pour ne citer que quelques exemples fréquents. Une mission intitulée « Bibliothèque et Web de données » a été confiée à un groupe d'incubation au sein du W3C entre mai 2010 et août 2011 pour réfléchir aux perspectives offertes par cette technologies pour les bibliothèques et pour « contribuer à améliorer de façon globale l'interopérabilité des données de bibliothèques sur le web en rassemblant de professionnels travaillant sur le web sémantique – notamment sur le web de données- qu'ils appartiennent à des

---

[175] David Wallis : « After much research and discussion, it was clear that ODC-BY was the best choice of license for many OCLC data services. The recommendation for members to also adopt this clear and consistent approach to the open licensing of shared data, derived from WorldCat, naturally flowed from this experience ».
http://www.oclc.org/news/releases/2012/201248.htm (consulté le 15/12/12)

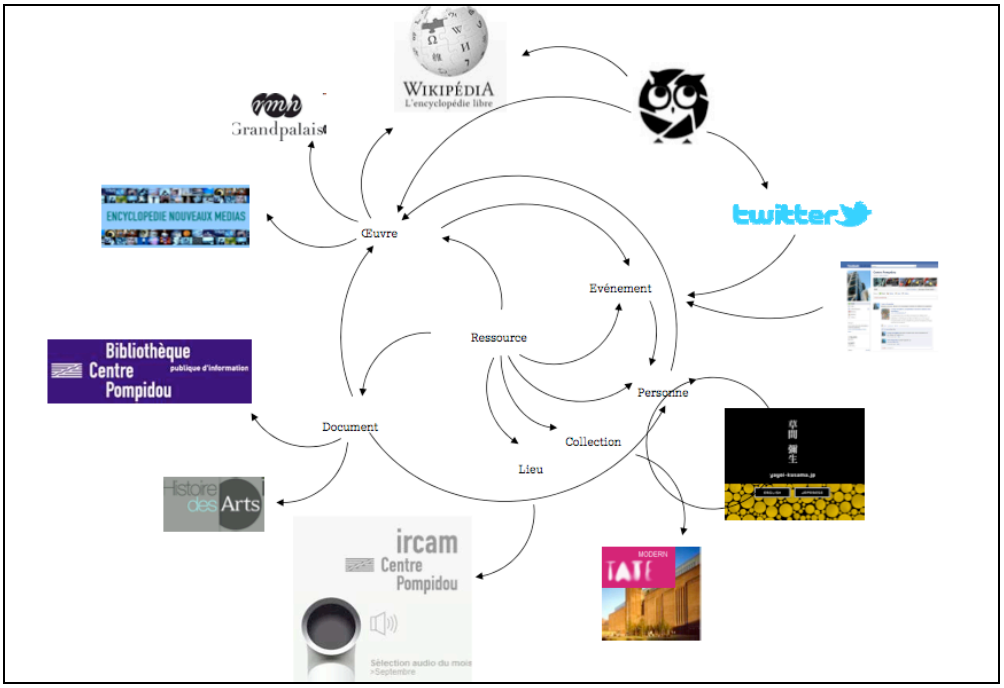
[176] http://viaf.org

[177] *OKFN Working Group on Open Bibliographic Data* : « Library institutions are among the first to benefit from a liberalization of bibliographic data » in *Open Bibloigraphic Data Flyer*, 10 septembre 2010.
http://openbiblio.net/2010/09/10/open-bibliographic-data-flyer/ (consulté le 15/12/12)

[178] Au sein de la BnF des passerelles existent déjà entre Open data et archivage du Web. Le décret d'application n° 2011-1904 du 19 décembre 2011 consécutif à la loi DADVSI n° 2006-961 du 1er août 2006 confie à la BnF la responsabilité de l'archivage du Web français au titre du dépôt légal. Lors de la campagne présidentielle de 2012, une campagne de collecte d'URL a été mise en place et à donner lieu à la publication, sous licence ouverte, d'un fichier de données issus des collectes du web électoral réalisé ou acquis par la Bnf depuis 2002 jusqu'aux législatives post présidentielles de 2012. Ce fichier est accessible via la plateforme nationale data.gouv.fr à l'adresse suivante : http://www.data.gouv.fr/donnees/view/Collectes-du-Web-électoral-par-la-BnF-551866?xtmc=élection%20nationale&xtcr=3

(cc) BY-NC-ND

bibliothèques ou non[179] ». Grâce à l'enrichissement obtenu par des liens avec des données complémentaires provenant de sources fiables, les bibliothèques peuvent accroître la valeur de leurs propres données au-delà du périmètre constitué par l'addition de toutes leurs sources.



Source: Emmanuelle Bermès, *Web de données, web sémantique*, 25 octobre 2011.

Pérenniser l'information

Comme le rappelle le groupe d'incubation du W3C, l'histoire a montré que toutes les technologies sont éphémères et que les formats de données spécifiques sont en générale de courte durée. Or dans la mesure où le web de données s'attache à décrire la signification des données indépendamment des structures spécifiques, celles-ci conservent leur signification lors des modifications de format. Les données liées sont donc plus pérennes et plus robustes que les formats de métadonnées qui dépendent d'une structure particulière de données. Par ailleurs, le fait de travailler dans le web est un garant de stabilité, car le système de noms de domaine permet, par exemple, de placer les identifiants dans un cadre de maintenance réglementé et clair en ce qui concerne les propriétaires et les responsabilités. Un cadre qui tend par ailleurs à se généraliser sous l'influence des travaux de standardisation du W3C dont la légitimité dans ce domaine fait désormais consensus. Ainsi, le fait de disposer de données ouvertes et structurées dans un format reposant sur un langage de balisage XML présente des garanties renforcées de conservation et de pérennisation de l'information.

Les institutions patrimoniales, parmi lesquelles figurent de nombreuses bibliothèques sont parmi les premières à pouvoir tirer parti de ces bénéfices, avec

[179] Rapport final du groupe d'incubation « Bibliothèques et Web de données », Groupe d'incubation, 25 octobre 2011. http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/ (consulté le 14 décembre 2012)

(cc) BY-NC-ND

la production sur le long terme de métadonnées fiables pour le domaine culturel[180]. Une part importante des projets de *Linked Open data* en bibliothèques concernent d'ailleurs des collections patrimoniales[181]. Nous avons déjà évoqué le fonds Trudat de la bibliothèque de Toulouse, mais il convient également de mentionner le projet de bibliothèque numérique patrimoniale de la bibliothèque des Champs Libres qui devrait bientôt voire le jour. Les données y seront structurées en RDF et placées sous licence domaine public comme le souhaite la bibliothèque et selon la dynamique insufflée par la collectivité. Le projet prévoit un enrichissement des données par des liens systématiques vers d'autres catalogues d'incunables et d'autres sites comme Gallica, Google Books ou encore Internet Archives[182].

Le projet OpenCat de la BnF : un catalogue ouvert [183]

Fin 2011, le Ministère de la culture et de la communication lance, dans le cadre de son programme de numérisation, un appel à projet pour l'année 2012, pour des « Services numériques culturels innovants ». Le 28 novembre 2011, la Bibliothèque nationale de France répond à cet appel en proposant un projet baptisé « OpenCat : les données ouvertes de la BNF au service des bibliothèques publiques ». Comme nous l'avons vu, depuis juillet 2011, la BnF expose librement via le site http://data.bnf.fr/ des extraits des données de ses catalogues selon les technologies et l'ergonomie du web de données, en permettant à tous de les réutiliser. Par le projet OpenCat, elle espère pouvoir tester et adapter ses avancées faites dans ce domaine dans l'environnement d'une collectivité pilote. L'enjeu est double : promouvoir auprès des autres bibliothèques les atouts du web de données ouvert pour la recherche documentaire et s'inscrire en tête de pont dans le transfert et le partage de ses technologies et de ses données à l'échelle du territoire auprès d'autres bibliothèques. En mars 2012, le projet OpenCat est sélectionné par le Ministère de la Culture. Il doit permettre à terme, de mettre les données ouvertes présentes dans *data.bnf.fr* au service des bibliothèques publiques en utilisant des logiciels libres, dont CubicWeb[184]. Le projet en lui même consiste en un partenariat de recherche entre deux bibliothèques – la BnF et la bibliothèque municipale de Fresnes-, avec le travail d'une entreprise spécialisée dans les technologies du Web sémantique, Logilab, afin d'expérimenter les nouvelles techniques de structuration et d'exposition des données de la BnF.

La bibliothèque de Fresnes a été une des institutions pionnières dans la récupération des notices bibliographiques et des notices d'autorités de la BnF. Elle les présentait jusqu'ici à ses lecteurs comme des résultats « à plat » comme le permettent classiquement les SIGB en format MARC mais désireuse de s'aligner sur les nouveaux modes de recherche et de consultation plus intuitives, promues par le modèle FRBR. Elle souhaiterait ainsi bénéficier des données et du modèle

---

[180] Cf. notamment le projet d'archivage réparti de la Bnf, SPAR (Système de pérennisation et d'archivage réparti) et le recourt aux identifiants ARK : « La BnF garantit la pérennité des identifiants ARK attribués […] le niveau déclaré des garanties pourra évoluer au fur et à mesure que les normes et méthodes de pérennisation seront mieux maîtrisées, et que la BnF mettra en application des procédures pour maintenir ces garanties et les associer aux objets qu'elle gère.]

[181] Emmanuelle Bermès, « Bibliothèques, archives et musées : l'enjeu de la convergence des données du patrimoine culturel », Documentaliste – sciences de l'information, 2011, vol. 48, n°11.

[182] Entretien avec Sarah Toulouse, responsable du projet de bibliothèque numérique patrimoniale à la bibliothèque des Champs Libres, 6 décembre 2012.

[183] Bibliothèque nationale de France, *Les données ouvertes de la BnF au service des bibliothèques publiques, Culture Labs,* juin 2012.
http://culturelabs.culture.fr/fiche.html?id=176&tri=cat (consulté le 17/12/12)

[184] CubicWeb est une plate-forme de développement d'applications web sémantique, développé en Python par la société française Logilab.

(cc) BY-NC-ND

déjà disponible à la BnF et des possibilités d'enrichissement de la recherche que permet le web de données grâce à l'alignement avec d'autres réservoirs de données disponibles à la bibliothèque ou bien sur le web. Par la mise au point d'un programme de recherche adapté, il deviendrait possible d'extraire les notices de Fresnes pour les comparer avec celle du corpus de *data.bnf.fr*, d'agréger dans une base externe les données locales du SIGB aux notices FRBRisées de la BnF, d'identifier les notices de Fresnes qui ne sont pas présentes dans *data.bnf.fr* et d'utiliser les technologies du Web sémantique pour lier les notices de Fresnes à celles de la BnF et d'autres données culturelles. Par ce partenariat, le rôle de la bibliothèque de Fresnes serait également de faire remonter les besoins des utilisateurs, les fonctionnalités souhaitées, les informations sur les systèmes et les collections pour travailler ensemble à l'évaluation des usages.

VERS UNE REVISION DES PRATIQUES ET DES OUTILS PROFESSIONNELS

Faire évoluer les normes bibliographiques

Si l'on a vu l'intérêt des bibliothèques à s'investir dans une démarche de *Linked Open data*, il est important de prendre conscience également de l'impact que celle-ci aura à terme tant sur les pratiques professionnelles que sur les partenaires traditionnels des bibliothèques que sont les organismes responsables des normes bibliographiques et les concepteurs de systèmes de gestion de données, notamment les éditeurs de SIGB[185]. Comme le fait judicieusement remarquer le groupe d'expert du W3C, les organismes responsables des normes bibliographiques doivent être les premiers à envisager le développement de normes de description compatibles avec le web de données et développer de bonnes pratiques en matière de conception de schémas pour les données liées de la bibliothèque. Les réflexions menées actuellement par l'IFLA depuis 2009 autour des nouveaux principes internationaux de catalogage que sont les modèles FRBR[186] et le format RDA[187] participent pleinement de cette logique. Ils ont pour vocation non seulement d'améliorer la présentation des catalogues en regroupant les informations relatives à une même œuvre (par exemple, traductions, révisions, éditions critiques, différentes présentations matérielles) mais également d'ouvrir la voie vers une interopérabilité accrue des données des catalogues de bibliothèques, par la médiation des modèles et ontologies de référence, et vers une ouverture des catalogues sur le web de données.

Vers une rénovation du marché logiciel

Il est important que les concepteurs de systèmes et de données prennent également en compte les implications techniques et juridiques de ces nouveaux usages, en premier lieu en intégrant les enjeux d'ouverture des données et métadonnées bibliographiques transitant par leurs systèmes (format non propriétaire, base ouverte attaquable par des API, faciliter d'export et de mise à jour régulière etc.) en améliorant la performance des services aux utilisateurs et

[185] Système intégré de gestion de bibliothèques.
[186] *Functional Requirements for Bibliographic Records* – Fonctionnalités requises des notices bibliographiques
[187] *Ressources : Description and Access* – Ressources : Description et accès.

(cc) BY-NC-ND

surtout en travaillant sur la performance du volet « services aux utilisateurs » de leur produit, très souvent prioritairement tournés vers les interfaces professionnelles. Il est important que ces services intègrent également les possibilités offertes par le web de données : créer des URI pour les *items* appartenant aux jeux de données des bibliothèques, mettre en œuvre des politiques pour la gestion des vocabulaires RDF et leur URI, intégrer les ontologies dans les systèmes d'indexation pour permettre la mise en correspondance de l'expression des données de bibliothèques avec celles existants dans le web de données[188]. Comme l'explique Frédéric Blin[189], c'est une démarche d'anticipation qu'adoptent de plus en plus de bibliothèques. Il serait souhaitable que les acteurs du marché soient en mesure d'emboiter le pas. Par ailleurs, l'intérêt pour les bibliothèques de passer à des technologies de *Linked Open data* se trouve également dans le plus grand choix de fournisseurs qui s'offrent alors elles. Si jusqu'à présent les bibliothèques utilisent des technologies et des formats de données spécifiques, en conséquence produits par des fabricants spécialisés de systèmes intégrés (SIGB), l'adoption des technologies du web sémantique leur ouvre la possibilité de recourir à des fournisseurs plus génériques dont bon nombre sont en libre accès, et ce tant pour l'équipement que pour la maintenance des systèmes. Le partenariat de la BnF avec la société LogiLab sur le projet OpenCat, est un exemple parmi d'autres de ces nouveaux types de partenariats.

A l'instar du succès rencontré par des solutions *Open source* comme Koha, il est même envisageable de pousser la réflexion jusqu'à concevoir des solutions progicielles entièrement conçues par des bibliothèques, pour des bibliothèques dans une logique de mutualisation. Comme le souligne le groupe d'incubation du W3C, le *Linked Open data* « sera peut-être la première étape d'une approche « dans les nuages » de la gestion de l'information culturelle, qui pourrait s'avérer plus rentable que les systèmes autonomes des institutions. Les petites institutions ou les projets individuels gagneraient en visibilité et en capacité d'interaction, tout en réduisant leurs coûts d'infrastructure[190] ». Cette option est d'ailleurs actuellement très sérieusement étudiée par l'ABES dans le cadre de son projet de « SGB mutualisé », posant notamment question quant à l'articulation entre le commun et le local. Dans le cadre de ce projet, l'ABES envisagerait de substituer une base globale, partagée par tous les établissements à travers l'Europe et au-delà, aux bases nationales traditionnelles.

Comme l'explique Raymond Bérard, directeur de l'ABES, « un scénario maximaliste consisterait, pour l'ABES, au moins à moyen ou long terme, à utiliser un de ces systèmes pour ses propres besoins en remplacement partiel ou total du SUDOC[191] ». Afin de faciliter la transition vers la constitution d'une base commune, sur le modèle du *WorldCat*, une diffusion des données du SUDOC sous

---

[188] Cela implique notamment une prise en compte des langages SKOS (*Simple Knowledge Organization System* - Système simple d'organisation des connaissances) conçu selon une structure OWL (Ontology Web langage). SKOS est une famille de langages formels permettant une représentation standard des thésaurus, classifications ou tout autre type de vocabulaire contrôlé et structuré. D'autres ontologies doivent également être prises en compte comme le *Dublin Core* pour le catalogage de données bibliographiques des bibliothèques et des ressources électroniques ou encore FOAF pour la description des personnes et les relations qu'elles entretiennent entre elles.

[189] Frédéric Blin, directeur de la BNU : « Le développement du web sémantique et la question de l'adoption du format RDA en France nous semblent devoir aboutir à terme à une nécessaire libération des données bibliographiques : nous anticipons d'une certaine manière sur ces évolutions prévisibles. »
http://alatoisondor.wordpress.com/category/le-travail-du-conservateur/ (consulté le 14/12/12)

[190] Collectif, *Bibliothèques et Web de données*, Rapport final du groupe d'incubation W3C, 25 octobre 2011

[191] Raymond Bérard, « SIGB vs Cloud computing », Éditorial in *Les licences nationales*, Arabesque, n° 66, avril-mai-juin 2012, p.3

(cc) BY-NC-ND

licence ouverte Etalab, devrait permettre une diffusion plus libre et plus large des métadonnées bibliographiques produites par le secteur de l'enseignement et de la recherche, sous réserve d'une mention d'attribution. Première étape de cette démarche, le 26 novembre 2012 l'ABES a déposé un premier *dump* RDF du SUDOC sur Sindice[192] permettant d'afficher les 10 millions de notices du SUDOC en RDF/XML. Accessible pour les programmes informatiques, ce fichier ne permet pas aujourd'hui d'explorer systématiquement le SUDOC ni d'y faire une recherche mais il constitue néanmoins une première étape vers la constitution d'une base de référence commune libre et structurée.

[192] Sindice : site d'indexation du Web sémantique. http://sindice.com/

(cc) BY-NC-ND

CONCLUSION

Aujourd'hui, en l'absence d'un cadre juridique défini, engager sa bibliothèque et ses données dans une démarche d'*Open data*, c'est avant tout réfléchir au sens que l'on souhaite donner au terme de « service public ». Si l'on se réfère aux principes établis par le manifeste de l'UNESCO[193] et l'IFLA[194] sur la bibliothèque publique, les bibliothèques ont, entre autres, pour mission « d'assurer l'accès des citoyens aux informations de toutes catégories issues des collectivités locales », de « fournir aux entreprises locales, aux associations et aux groupes d'intérêt les services d'information adéquats » et de contribuer au « progrès d'une société démocratique en donnant aux individus accès à une large gamme de savoirs, d'idées et d'opinions ». Au regard, de ces éléments, des enjeux, expériences et promesses de l'*Open data* établis précédemment, il semble exister une convergence naturelle entre l'ouverture des données publiques et les missions de « service public », qui dépassent le simple cadre des bibliothèques.

Sans réfuter totalement la dimension économique de l'*Open data*, il serait fallacieux de considérer cet argument comme un élément premier de motivation pour le lancement d'une démarche *Open data*. En revanche, on ne peut que saluer le dynamisme, l'esprit de collaboration et la démarche solidaire et citoyenne qui animent, en règle général, les différents acteurs du mouvement. Leur démarche reste en grande partie volontaire et non lucrative. Beaucoup de développeurs, sont issus du milieu du logiciel libre et de l'*Open source*. Par l'*Open data*, ils ont dans l'idée de proposer de nouveaux services aux habitants et non celle d'obtenir une rémunération. Si cette attitude est ce qui fait aujourd'hui la force de l'*Open Data*, c'est aussi ce qui constitue à terme sa faiblesse, car comme toute activité bénévole, l'investissement dans les projets doit faire face à des abandons réguliers, un enlisement des projets ou un important *turn-over*. Autant de tendances qui menacent de s'accentuer si le mouvement ne parvient à créer de synergie à plus grande échelle.

Pourtant derrière ces initiatives, qui d'apparence confinent parfois aux loisirs, à l'expérimentation ou à l'effet de mode, nous nous sommes efforcés, au cours de ce travail, de mettre en avant la plus value que peut réellement représenter l'*Open data* pour les bibliothèques, tant dans les services qu'elles peuvent offrir au public que dans leurs démarches de coopération professionnelle. Il convient néanmoins de garder un certain regard critique et l'esprit pragmatique quant à l'investissement que nécessite de telles opérations. L'ouverture des données est une démarche contraignante, coûteuse en temps, en effectif et en matériel. Mettre en ligne des données demande une réflexion importante sur le choix des informations, un travail de mise en forme, de simplification et d'actualisation régulier des données, parfois aussi exigeant qu'un travail d'édition, et enfin un effort de médiation et d'animation autour des données. Une démarche dont les objectifs doivent se concevoir, dès le départ, à long terme.

Par manque de temps, autant que de convictions, beaucoup trop de démarches d'ouverture des données, en bibliothèque comme ailleurs, ne sont conçues qu'à

[193] Manifeste de l'UNESCO sur la bibliothèque publique. http://www.unesco.org/webworld/libraries/manifestos/libraman_fr.html (consulté le 23/12/12)

[194] *Les services de la bibliothèque publique : principes directeurs* de l'IFLA/UNESCO, avril 2001, 95 p.

(cc) BY-NC-ND

très court terme, comme des opérations uniques. Si l'idée de l'*Open data* est bien de faire appel au dynamisme et à la créativité du public, en lui laissant accéder clé en main aux données des institutions, l'expérience a montré, que c'était une erreur de considérer que le travail des institutions productrices (bibliothèques, État, collectivités etc.) s'arrêtait avec la mise en ligne. Il est donc important de ne pas négliger le coût de la démarche, ni de sous-estimer le besoin d'animation qu'elle nécessite pour ne pas la voir péricliter avant même d'avoir pu porter ses fruits. En d'autres termes, pour les établissements culturels comme les bibliothèques, encore plus que pour les autres, la participation à l'*Open data* doit être pensée dans un ratio coût/avantage : A combien estimer le coût de l'ouverture des données, en temps, en effectif, en maintenance ? Quel est le retour sur investissement attendu ? En d'autres termes, l'amélioration de l'offre de service pour le public ou les bénéfices professionnels escomptés seront-ils à la hauteur des sommes et des moyens investis dans une démarche d'*Open data* ?

Difficile d'apporter des réponses certaines à toutes ces interrogations. Néanmoins, la création de groupe de réflexion et la publication des conclusions tirées des premiers retours d'expérience devaient permettre non seulement d'éviter certaines dérives constatées dans le domaine de la numérisation[195] mais surtout de proposer des critères relativement objectifs quant à l'opportunité ou au coût estimé pour tel établissement de se lancer dans une politique d'*Open data*. Au vu des exemples et des observations dressées dans ce mémoire, plusieurs critères peuvent déjà permettre d'évaluer le degré de pertinence et les facteurs de réussite d'une démarche d'*Open data* en bibliothèques.

En premier lieu, si leurs missions, leur vocation et leur logique professionnelle en font des candidates idéales à l'ouverture des données publiques, les bibliothèques ne peuvent se passer d'un travail d'animation qui ne peut se concevoir efficacement qu'en lien avec le tissu administratif et associatif local ou national. A l'image de leurs catalogues, les bibliothèques ont besoin d'ouverture. Par ailleurs, l'*Open data* présente l'avantage de s'inscrire dans la logique des réflexions menées actuellement autour du renouveau des normes et standards bibliographiques, de description et d'échange de données en bibliothèques et au-delà. En conséquence, il est de l'intérêt des établissements déjà investis dans ces démarches prospectives, de prendre en compte, dès maintenant, les enjeux de l'*Open data*, de l'*Open access* et et de l'*Open source*.

Enfin, il semble assez évident qu'en bibliothèque, sans doute encore plus que dans les autres institutions culturelles, l'*Open data* n'a d'intérêt réel qu'en complément des technologies du *Linked data*. Sous l'influence des enjeux foncièrement politiques de démocratie, de transparence et de croissance économique, l'accent a d'avantage été mis sur la notion « d'ouverture » que sur celle de « valorisation » des données publiques. Comme le soulignait Romain Wenz, « il y a quelques années on ne parlait que de *Linked open data*, puis on s'est mis à concevoir les deux séparément l'un de l'autre[196] ». Or cette solution de continuité entre « ouverture » et « valorisation » est aujourd'hui ce qui pénalise le

[195] Par effet de mode et souci de valorisation de l'établissement beaucoup d'opération de numérisation ont été conçue en en dehors de toute logique de concertation. En conséquence, beaucoup d'œuvres ont été numériser plusieurs fois, par des institutions différentes, selon des formats et des qualités variés, en l'absence de cadre juridique homogène, en contradiction de toute logique d'interopérabilité et de coopération. Compte-tenu du coût de ces opérations, ce manque de concertation est parfois regrettable.

[196] Entretien avec Romain Wenz, mars 2012.

(cc) BY-NC-ND

développement de l'*Open data*. Si cette distinction a pu néanmoins se concevoir, un instant, pour publier des données de fonctionnement ou des données chiffrées, la véritable valeur ajoutée des bibliothèques, pour le public comme pour les professionnels, porte moins sur les données de fonctionnement que sur leurs données scientifiques, à savoir leurs métadonnées bibliographiques : notices d'œuvres, notices d'autorité auteur, matière, données d'indexation etc. En conséquence, optimiser, à l'avenir, l'ouverture, la valorisation et la réutilisation de ce types des données en bibliothèques, indépendamment des technologies du *Linked Open data* serait une gageure.

Cette évolution du web, des modes de recherche, de navigation semble d'autant plus irrémédiable qu'elle est suivie par les principaux *leaders* dans le domaine de l'Internet, à savoir Google, Wikipédia ou encore le W3C pour ne citer que quelques noms. Depuis mai 2011, une initiative des moteurs de recherche Google, Bing et Yahoo a même conduit à la définition du « schema.org »[197], un langage permettant d'intégrer des balises « sémantiques » dans une page Web. Déjà environ 10% des sites Web intègrent ce type de balises à l'heure actuelle. C'est une manière directe d'amener ceux qui publient des données à fournir des données correctement structurées, afin d'être mieux noté et référencé dans les moteurs et gratuites, car ce sont par essences, des pages Web, non commerciales. Dans le domaine de la culture et de la connaissance, le lancement du *Google Graph of Knowledge*, du Centre Pompidou virtuel, de Data.bnf.fr ou encore de conférence comme l'*International Linked Open Data in Libraries Archives and Museums Summit,* montre à quel point, en France comme à l'étranger, le *linked open data* s'impose comme une solution vers laquelle tendent progressivement tous les grands établissements tournés vers l'avenir. En conséquence, ce mémoire ne saurait que rappeler l'importance pour les bibliothèques de faire évoluer leurs catalogues en cohérence avec les technologies du web et les nouvelles possibilités offertes pour la recherche, car qu'est ce qu'un catalogue ou un inventaire si ce n'est un instrument de recherche et de signalement ?

En ce sens, les bibliothèques ont semble-t-il tout à gagner à s'approprier, pour leurs propres catalogues et services, les technologies développées et adoptées par les grands moteurs de recherche sur le net. Ce n'est que de cette manière que les grandes institutions culturelles pourront être en mesure d'offrir une qualité de services publics équivalente à celle proposer par les grands opérateurs privés dans ce secteur. Néanmoins, compte-tenu de la taille et des moyens dont disposent aujourd'hui les bibliothèques publiques comme universitaires, cette transition ne peut s'envisager que dans une logique de coopération. Dans cette optique, le projet OpenCat de la BnF se conçoit comme un premier pas incitateur vers une politique de coopération et de mutualisation des données et des compétences, en vue d'une généralisation des technologies du *Linked Open data* dans les bibliothèques.

[197] Se reporter aux travaux du groupe « schemabibex » du W3C : http://www.w3.org/community/schemabibex/

(cc) BY-NC-ND

(cc) BY-NC-ND

Sources

Entretien avec Simon CHIGNARD, La cantine numérique rennaise, 19 juillet 2012

Entretien avec Frédérique SCHLOSSER, responsable du département des réseaux en charge de la politique *Open data* de la bibliothèque des Champs Libres, Rennes Métropole, 19 juillet 2012

Entretien avec Sarah TOULOUSE, responsable du projet de bibliothèque numérique patrimoniale à la bibliothèque des Champs Libres, Rennes Métropole, 6 décembre 2012.

Entretien avec Léa LACROIX, ancienne stagiaire de la DG Com sur le projet *Open data*, membre du Collectif citoyen Open data Rennes Métropole, 20 juillet 2012

Entretien avec Bernadette KESSLER, responsable *Open data* à la direction générale de la Communication de Rennes Métropole.

Entretien avec Romain WENZ, responsable du projet *Data.bnf.fr* au sein du département de l'information bibliographique et numérique de la Bibliothèque nationale de France, mars 2012.

Échange avec Jean-Marie BOURGOGNE, responsable *Open data* à Montpellier.

Entretien avec Julien SEMPERE, membre du groupe de réflexion de l'ABES sur le SGB mutualisé.

(cc) BY-NC-ND

Bibliographie

Généralités

BROUDOUX Évelyne, CHARTRON Ghislaine, *Enjeux politiques du document numérique : actes de la troisième conférence Document numérique et société,* « Documents, contenus numériques : politique en question », Aix-en-Provence, 15-16 novembre 2010, Paris, ADBS éd., coll. « Sciences et techniques de l'information, 2010.

CHIGNARD Simon, *L'Open data : Comprendre l'ouverture des données publiques*, FYP éditions, 2012.

COLLECTIF DU NUMERIQUE (LE)

http://www.collectifdunumerique.fr/

EAVES David, *Un État des lieux de l'Open data, OWNI*, 2 novembre 2011.

http://owni.fr/2011/11/02/etat-des-lieux-open-data-eaves/

GUILLAUD Hubert, « Quelles données libère-t-on ? », Critiques du Web^2 (1/4), 16 septembre 2009

http://owni.fr/2009/09/16/critiques-du-web^2-14-quelles-donnees-libere-t-on/

GUILLAUD Hubert, « Mesurer le résultat de la libération des données », *Critique du Web^2 (2/4)*, 16 septembre 2009.

http://www.internetactu.net/2009/09/16/critiques-du-web^2-24-les-effets-de-la-liberation-des-donnees/

GUILLAUD Hubert, « Toutes les données sont devenues personnelles ? », *Critique du Web^2 (3/4),* 21 septembre 2009.

http://www.internetactu.net/2009/09/21/critique-du-web^2-34-toutes-les-donnees-sont-devenues-personnelles/

GUILLAUD Hubert, « Que faire face à la puissance des données ? », *Critique du Web^2 (4/4)*, 26 janvier 2009.

http://www.internetactu.net/2009/10/26/critiques-du-web^2-44-que-faire-face-a-la-puissance-des-donnees/

GUILLAUD Hubert *, Open data : Où en est-on ? (1/4)*, 15 juin 2012.

http://internetactu.blog.lemonde.fr/2012/06/15/open-data-ou-en-est-on/.

GUILLAUD Hubert, « Open data: l'enjeu de la coproduction », *Open data : Où en est-on ? (1/4)* 6 juin 2012

http://www.internetactu.net/2012/06/06/open-data-34-lenjeu-de-la-coproduction/

GUILLAUD Hubert, « Open data: animer, animer, encore animer », *Open data : Où en est-on ? (2/4),* 31 mai 2012

http://www.internetactu.net/2012/05/31/open-data-24-animer-animer-et-encore-animer/

GUILLAUD Hubert, *Open data, un premier bilan français*, 31 mai 2012.

(cc) BY-NC-ND

http://owni.fr/2012/05/31/open-data-france-premier-bilan-francais-etalab/

JOSPIN Lionel, « Préparer l'entrée de la France dans la société de l'information », *Discours de Hourtin du 25 août 1997,* Université de la Communication.

KAPLAN Daniel, *L'ouverture des données publiques et après ?*, OWNI, 18 janvier 2011.

http://owni.fr/2011/01/18/louverture-des-donnees-publiques-et-apres/

LACOMBE Romain, BERTIN Pierre-Henri, VAUGLIN François, VIEILLEFOSSE Alice, *Pour une politique ambitieuse des données publiques. Les données publiques au service de l'innovation et de la transparence,* Rapport à la délégation aux usages de l'Internet dans le cadre du portail Proxima Mobile, remis au Ministre de l'Industrie, de l'énergie et de l'économie numérique, le 13 juillet 2011.

LIBERTIC, *Open data : 20 questions qui se posent au gouvernement français*, 17 juin 2012.

http://libertic.wordpress.com/2012/06/17/open-data-20-questions-qui-se-posent-au-gouvernement-francais/

MAGAZINE DE L'OPEN DATA (LE)

http://123opendata.com/blog/

MANDELKERN (Dieudonné), MARAIS Bertrand (DU), *Diffusion des données publiques et révolution numérique : rapport de l'atelier « Des moyens nouveaux au service de la diffusion des données publiques »*, Paris, la Documentation française, 1999, 123 p.

ORGANISATION D'OBSERVATION DES PARLEMENTS (OOP), *Déclaration pour l'ouverture et la transparence parlementaire*, 29 août 2012, 16 p.

PASQUESOONE Valentine, « En France, l'Open data en marche », *Le Monde.fr,* 21 mai 2012

http://www.lemonde.fr/politique/article/2012/05/31/en-france-l-open-data-en-marche_1709874_823448.html

Bibliothèques

BIBLIOTHEQUE NATIONALE DE FRANCE (BNF), *Les données ouvertes de la BnF au service des bibliothèques publiques*, Culture Labs, juin 2012.

http://culturelabs.culture.fr/fiche.html?id=176&tri=cat

BIBLIOTHEQUE NATIONALE DE FRANCE, *Présentation générale du projet data.bnf.fr*, 15 novembre 2011.

http://data.bnf.fr/docs/databnf-presentation.pdf

CHAIMBAUD Thomas, « Open data et bibliothèques », *Vagabondages,* 17 juillet 2012.

http://www.vagabondages.org/post/2012/07/17/Bibliothèques-et-Open-data

MERCIER Silvère, *Bibliothèques publiques et Open data : quels enjeux ?*, Bibliobsession, 7 juin 2011.

(cc) BY-NC-ND

http://www.bibliobsession.net/2011/06/07/bibliotheques-publiques-et-donnees-ouvertes/

WENZ Romain, *L'avenir des catalogues : formats, données, outils, usages*, mémoire d'étude DCB, janvier 2009.

http://www.enssib.fr/bibliotheque-numerique/document-21205

Collectivité

Généralités

ANONYME, *Open data: quel impact pour les administrations publiques ?*, 29 novembre 2011

http://rh.sia-conseil.com/20111129/l%E2%80%99open-data-une-opportunite-pour-la-collectivite-et-les-administrations-publiques-l%E2%80%99open-data-quels-impacts-pour-les-administrations-publiques/

BERDOT Vincent, *Pourquoi les collectivités rechignent à mutualiser leurs plates-formes Open data*, 01Net, 4 juin 2012.

http://pro.01net.com/editorial/567381/pourquoi-les-collectivites-rechignent-a-mutualiser-leurs-plates-formes-open-data/

COLLECTIF OPEN DATA FRANCE, « Data sets des Data sets »

http://opendatafrance.net/les-ateliers/atelier-3-dataset-des-datasets/

FING, *Guide pratique de l'ouverture des données publiques territoriales*, Guide à l'usage des territoires et de leurs partenaires, , Version de travail n°1, Janvier 2011, 66 p.

PARMENTIER Philippe, *Consultation pour un premier projet de normalisation*, Données publiques, Localtis, 24 avril 2012.

OBSERVATOIRE DES TERRITOIRES NUMERIQUES, *Données publiques : clés et repères*, janvier 2012, 178 p.

Rennes

BLANC Sabine, *Open data péRennes*, 21 février 2012.

http://owni.fr/2012/02/21/open-data-perennes/

COLLECTIF OPEN DATA RENNES METROPOLE, « Biblioviz: visualiser les données de la bibliothèques »

http://opendatarennes.wordpress.com/operation-biblioviz /

JOLY Julien, *Open data : 4 projets qui vont hacker la vie rennais*, Rennes 1720, 20 juin 2012.

http://rennes1720.fr/2012/06/20/open-data-4-projets-qui-vont-hacker-la-vie-rennaise-63126/

JOLY Julien, *Les bibliothécaires de Rennes sont tristes le vendredi*, Rennes 1720, 15 juillet 2012

http://rennes1720.fr/2012/07/15/les-bibliothecaires-de-rennes-sont-tristes-le-vendredi-52036/

(cc) BY-NC-ND

LACROIX Léa, *Open data : Comment impliquer les habitants d'un territoire dans le processus d'ouverture des données publiques ?*, Mémoire Universitaire, Licence pro USETIC, juin 2012, 29 p.

http://www.scribd.com/doc/96964660/Memoire-Universitaire-Open-Data-Comment-impliquer-les-habitants-d-un-territoire-dans-le-processus-d-ouverture-des-donnees-publiques

RENNES-METROPOLE, *Rapport d'activités*, 2009, 106 p.

RENNES-METROPOLE, *Rapport d'activités*, « Créatrice, apprenante, accueillante, solidaire, entreprenante et ouverte », 2010, 92 p.

Autres collectivités

CLEMENT Jean-Philippe, *Atelier EtaLab : Retour d'expérience Open data Paris*, 2011

LiberTIC, *Présentation Opendata de la ville de Nantes*, 2011

MONTPELLIER TERRITOIRE NUMERIQUE, *Montpellier Territoire Numérique et l'Open data*, 2012

Mouvements citoyens

REGARDS CITOYENS

http://www.regardscitoyens.org/

COLLECTIF POUR LES DONNEES OUVERTES DANS LE RHONE

http://opendata69.org/

LE COLLECTIF OPEN DATA DE TOURAINE

http://opendata-touraine.fr/

LIBERTIC - NANTES

http://libertic.wordpress.com/libertic/

LE BLOG DU COLLECTIF DE L'OPEN DATA A RENNES

http://opendatarennes.wordpress.com/

Economie - Entreprise

BERDOT Vincent, *Open data : un modèle économique encore incertain*, 11 juillet 2011.

http://pro.01net.com/editorial/535620/open-data-un-modele-economique-encore-incertain/

BOYER Anne-Sophie, *Open data: citoyenneté et économie sont-ils antinomiques ?*, 30 janvier 2012.

http://googlemystate.wordpress.com/2012/01/30/open-data-citoyennete-et-economie-sont-ils-antinomiques/

CHIGNARD Simon, « Où est le marché de l'Open data ? », *Données ouvertes*, 12 mars 2012.

(cc) BY-NC-ND

http://donneesouvertes.info/2012/03/12/ou-est-le-marche-de-lopen-data/

FING, *L'Open data. Pérenniser les actions. Au cœur de l'animation de l'activité économique locale*, OPEN DATA WEEK - Plénière 1.

http://fr.slideshare.net/slidesharefing/open-data-week-plnire-1-lopen-data-au-coeur-de-lanimation-de-lactivit-conomique-locale

GAZETTE DES COMMUNES (LA), « Données publiques : une nouvelle économie des services publics en débat », *Réutilisation des données publiques : des promesses vertigineuses*, 23 novembre 2010.

http://www.lagazettedescommunes.com/48192/donnees-publiques-une-nouvelle-economie-des-services-publics-en-debat/

GUILLAUD Hubert, « Open data: le monde de l'entreprise face au défi de l'ouverture des données publiques, , *Open data : Où en est-on ? (4/4)*, 7 juin 2012

http://www.internetactu.net/2012/06/07/open-data-44-le-monde-de-lentreprise-face-au-defi-de-louverture/

GUIMARD Emmanuel, « Open data » : les collectivités sont en quête d'un modèle économique », *Les Echos*, n° 21196, 1 Juin 2012, p. 5

LIBERTTIC, « Pourquoi la réutilisation des données publiques à des fins commerciales doit-elle être gratuite ? », 25 février 2011.

http://libertic.wordpress.com/2011/02/25/pourquoi-la-reutilisation-des-donnees-publiques-a-des-fins-commerciales-doit-etre-gratuite/

TRANSPORTS DU FUTUR (LES), *Livre blanc : Open data: quels enjeux, quelles opportunités pour les entreprises ?*, Bluenove.

http://fr.slideshare.net/transportsdufutur/livre-blancbluenove?ref=http://lecercle.lesechos.fr/entreprises-marches/high-tech-medias/internet/221140403/livre-blanc-open-data-quels-enjeux-et-opport

Aspects juridiques

Généralités

FING, *Présentation généraliste sur les réutilisations des données publiques.*

GUILLAUD Hubert, «Toutes les données sont devenues personnelles », *Critique du Web² (3/4)*, 21 septembre 2009.

http://www.internetactu.net/2009/09/21/critique-du-web²-34-toutes-les-donnees-sont-devenues-personnelles/

OPEN KNOWLEDGE FOUNDATION, *Legal tools for Open data*.

http://opendatacommons.org/

MAUREL Lionel, *L'Open data à la croisée des chemins juridiques*, 7 juin 2012.

http://owni.fr/2012/06/07/l%E2%80%99open-data-a-la-croisee-des-chemins-juridiques/

(cc) BY-NC-ND

Textes réglementaires

COMMISSION EUROPEENNE, Réutilisation des informations du secteur public : réexamen de la directive 2003/98/CE, 7 mai 2009.

http://ec.europa.eu/information_society/policy/psi/docs/pdfs/opendata2012/revision_of_PSI_Directive/proposal_directive_FR.pdf

CONSEIL NATIONAL DU NUMERIQUE, Avis n° 12 relatif à l'ouverture des données publiques (« Open data »), 5 juin 2012, 23 p.

Loi CADA n°78-753 du 17 juillet 1978 portant diverses mesures d'amélioration des relations entre l'administration et le public et diverses disposition d'ordre administratif, social et fiscal.

http://www.legifrance.gouv.fr/affichTexte.do;jsessionid=4D4218C9137FAC6B6F657E4E5DF39060.tpdjo17v_1?cidTexte=JORFTEXT000000339241&dateTexte=20991231

Loi CNIL, n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.

http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006068624&dateTexte=vig

Directive PSI 2003/98/CE du 17 novembre 2003 concernant la réutilisation des informations du secteur public.

http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32003L0098:fr:HTML

Ordonnance 2005-650 du 6 juin 2005 relative à la liberté d'accès aux documents administratifs et à la réutilisation des informations publiques

http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000629684

Décret n°2005-1755 du 30 septembre 2005 relatif à la liberté d'accès aux documents administratifs et à la réutilisation des informations publiques

http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000265304&dateTexte=&categorieLien=id

Aspects techniques – Web de données

BAUER Florian, KALTENBÖCK Martin, *Linked Open data : The Essentials. A Quick Start guide for Decision Makers*, janvier 2012, 24 p.

BERMES Emmanuelle, « Vers le Web sémantique ? Les bibliothèques face aux normes du web », *Arabesque*, n°62, juin 2011.

BIBLIOTHEQUE NATIONALE DE FRANCE, *Le Web de données à la BnF : data.bnf.fr.*

http://www.bnf.fr/fr/professionnels/modelisation_ontologies/a.web_semantique_bnf.html

BIBLIOTHEQUE NATIONALE DE FRANCE, *Web sémantique, Web de données : présentation*.

http://www.bnf.fr/fr/professionnels/web_semantique_donnees/s.web_semantique_intro.html

(cc) BY-NC-ND

BOUCHAREB Hind, CARTIER Aurore, GAREL-GRISLIN Julie, LEON Y BARELLA Alicia, LEPREVOST Lucie, MION-MOUTON Fanny, QUACH Cécile, Web sémantique, web de données : quels intérêts pour les bibliothèques ?, travail d'étude DCB21, Enssib, 2012.

http://websmantique.unblog.fr/

COLLECTIF, *Bibliothèques et Web de données*, Rapport final du groupe d'incubation W3C, 25 octobre 2011.

http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/

DEKKERS Maks, *Linked (open) data : opportunities and challenges*, journée ABES 20 juin 2011.

http://fr.slideshare.net/abesweb/1-dekkers-jabes11

GANDON Fabien, FARON-ZUCKER Catherine et CORBY Olivier, *Le Web sémantique : comment lier les données et les schémas sur le web ?*, Paris, Dunod, coll.« InfoPro. Management des systèmes d'information », 2012, vol. 1/.

KEPEKLIAN Gabriel et LEQUEUX Jean-Louis, *Déployer un projet Web 2.0 : anticiper le Web sémantique (Web 3.0)*, Paris, Éd. D'organisation, 2008.

LODLAM, *International Linked Open Data in Libraries Archives and Museums Summit*, 2-3 juin 2011 San Francisco.

http://lod-lam.net/summit/

International

ETALAB, *Data.gov.uk l'ouverture des données publiques au royaume uni*, 2011

http://www.slideshare.net/Etalab/datagovuk-louverture-des-donnes-publiques-au-royaume-uni

SHADBOLT Nigel, *L'Open data n'est plus une chimère* », Entretien, Regards sur le numérique, 24 février 2011.

http://www.rslnmag.fr/post/2011/02/24/171;-Lopen-data-nest-plus-une-chimere-187;-entretien-avec-Nigel-Shadbolt.aspx

KUNDRA Vivek, *La libération de données a donné naissance à une économie des applications*, Regards sur le numérique, 22 février 2011.

http://www.rslnmag.fr/post/2011/3/16/vivek-kundra_la-liberation-de-donnees-a-donne-naissance-a-une-economie-des-applications_.aspx

Culture

GOETA Samuel, *Compte-rendu de l'atelier sur les données culturelles*, Semaine de l'Open data 21-26 mai 2012, juin 2012.

http://www.opendataweek.org/wp-content/uploads/2012/06/CR-AtelierB-2505.pdf

JEAN Benjamin, MAUREL Lionel, FILIPPI Primavera (de), CHARMET-ALIX Adrienne, *Recommandations pour l'ouverture des données et des contenus culturels*, Rapport du Groupe *Open Glam*, Wikimedia France, 2012, 14 p.

MAUREL Lionel, *Et si la culture s'ouvrait à l'Open data ?*, 25 mai 2011.

(cc) BY-NC-ND

http://owni.fr/2011/05/25/et-si-la-culture-souvrait-a-lopendata/

MINISTERE DE LA CULTURE ET DE LA COMMUNICATION, *Culture et Médias 2020 :Un ministère nouvelles génér@tion Secrétariat général*, Mission Stratégie & prospective, mars 2012, 302 p.

http://www.culturecommunication.gouv.fr/.../Minist%C3%A8re+nouvelle+g%C3%A9n%C3%A9r@a...

NEPOTE Charles, *La culture est-elle le parent pauvre de l'Open data ?*, 14 avril 2011.

http://www.reseaufing.org/pg/blog/openid_82/read/62085/la-culture-estelle-le-parent-pauvre-de-lopen-data-

NIGGEMAN Elisabeth, DECKER (DE) Jacques, LEVY Maurice, *The new renaissance,* Report of the « Comité des sages », reflection group on bringing Europe's cultural heritage online, Bruxelles, 10 janvier 2011, 45 p.

http://ec.europa.eu/information_society/activities/digital_libraries/comite_des_sages/index_fr.htm

ORY-LAVOLLE Bruno , « Partager notre patrimoine culturel » - *Propositions pour une charte de la diffusion et de la réutilisation des données publiques culturelles numériques*, Rapport Boudy-Ory-Lavollé, remis en mai 2010 à la Ministre de la culture et de la communication, La documentation française, 2009, 65 p.

ORY-LAVOLLE Bruno, *La diffusion numérique du patrimoine dimension de la politique culturelle*, Rapport à Mme la Ministre de la Culture et de la Communication, janvier 2002, 243 p.

PENE Camille, F*ree et Open data dans le secteur culturel : quelques réalisations*, Conférence Atelier français, Open data Design Group – ENSCI, 21 octobre 2011

http://fr.slideshare.net/Atelier-Francais/free-open-data-dans-le-secteur-culturel-quelques-ralisation#btnNext

ZELNICK Patrick, TOUBON Jacques, CERUTTI Guillaume, *Création et Internet*, Rapport au Ministre de la culture et de la communication, Janvier 2010, p.147.

Webographie

BOUTANY Joumana, *L'accès et la réutilisation des données publiques : état des lieux et des pratiques en France*, Biennale du numérique, 2011.

http://www.enssib.fr/bibliotheque-numerique/document-56432

EUROPEANA, *Linked Open data : what is it ? Why is it good for you ?*, Vimeo

http://vimeo.com/36752317

PEARLTREE : *Open data*

http://www.pearltrees.com/#/N-p=21099068&N-pw=-1&N-f=1_2875646&N-s=1_2875646&N-fa=2875646&N-u=1_283462

PEARLTREE : *Web sémantique, web de données*

http://www.pearltrees.com/#/N-p=20447754&N-f=1_1596265&N-s=1_1596265&N-fa=927907&N-u=1_83863

LIBERTIC : *Le film de l'Open data*

(cc) BY-NC-ND

http://libertic.wordpress.com/2011/12/12/le-film-de-lopen-data/

GOOGLE, *Introducing the Knowledge Graph*, 16 mai 2012.

http://googlefrance.blogspot.fr/2012/12/lancement-en-france-du-knowledge-graph.html

(cc) BY-NC-ND

Table des annexes

Trois scénarios pour l'Open Data 86

De l'accès aux documents administratifs à la libre réutilisation des données publiques : un long cheminement juridique 88

Opération #Biblioviz, les réalisations 94

Licence de réutilisation des informations publiques de Rennes Métropole..... 95

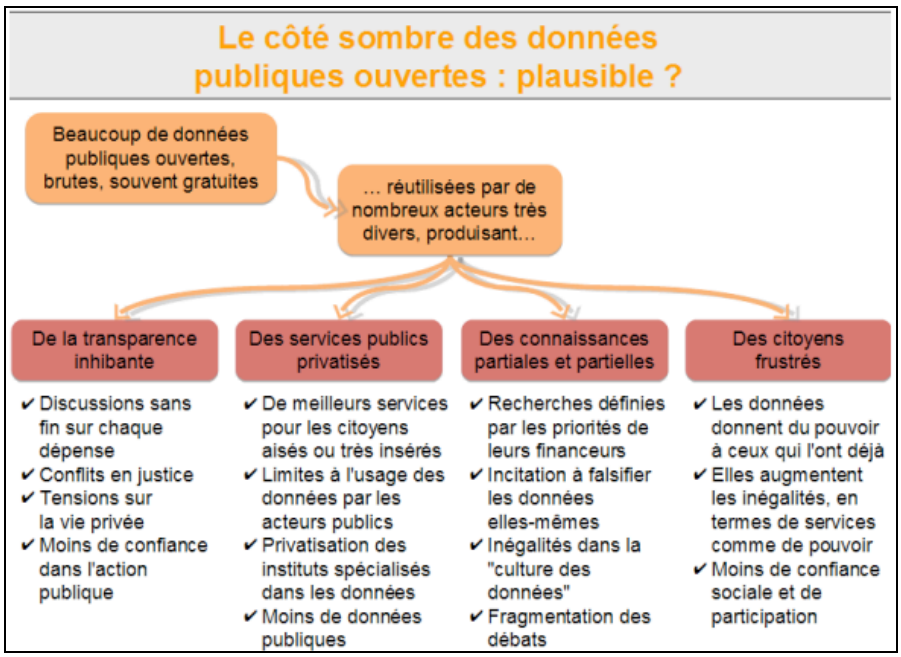
Carte de France de l'Open data 97

Licence ouverte ETALAB..... 98

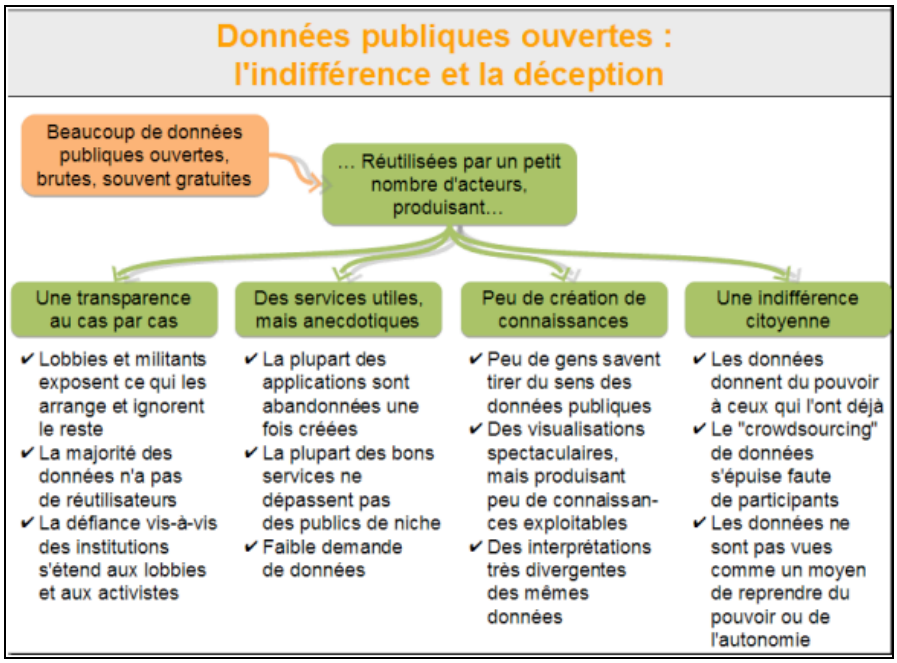
Linked Open Data : enjeux et fonctionnement 101

(cc) BY-NC-ND

TROIS SCENARIOS POUR L'OPEN DATA



(cc) BY-NC-ND



Source : Daniel Kaplan, *L'ouverture des données publiques et après ?*, OWNI, 18 janvier 2011 http://owni.fr/2011/01/18/louverture-des-donnees-publiques-et-apres/ (consulté le 09/09/2012)

(cc) BY-NC-ND

DE L'ACCES AUX DOCUMENTS ADMINISTRATIFS A LA LIBRE REUTILISATION DES DONNEES PUBLIQUES : UN LONG CHEMINEMENT JURIDIQUE

DE L'ACCES AUX DONNEES A LA REUTILISATION

Le droit d'accès à l'information publique en France, n'est pas un phénomène récent, quoi que souvent méconnu. Il s'inscrit dans la lignée des fondements révolutionnaires posés par l'article XV de la Déclaration des droits de l'homme et du citoyen de 1789 : « La Société a le droit de demander compte à tout agent public de son administration ». Mais passer d'un principe d'accès à celui d'une réutilisation à des fins potentiellement commerciales n'est pas une chose aisée et l'on retrouve ce cheminement autant intellectuel que politique au travers de chacun des textes juridiques qui ont jalonné cette période.

Le point de départ de ces réflexions a été posé par **la loi n°78-753 du 17 juillet 1978 sur l'accès aux documents administratifs dite loi CADA[198].** Cette loi portant sur « diverses mesures d'amélioration des relations entre l'administration et le public et diverses dispositions d'ordre administratif, social et fiscal »[199] reste encore aujourd'hui un texte fondateur des relations entre administrations et usagers dans le domaine de l'information. S'il est bien question de droit d'accès, le texte ne traite de la diffusion qu'à la marge et seulement par la négative[200] puisque celle-ci est laissée à l'initiative de l'usager. De fait, la première version de la loi CADA est encore loin d'une démarche Open data. Non seulement, il n'est pas encore question de « données » mais seulement de « documents publics » et de plus la réutilisation reste encore très fermée puisqu'en sont exclues les entreprises chargées d'une mission économique et commerciale ce qui ne permet pas de commercialisation des données publiques pourtant un des enjeux cruciaux de la démarche Open data.

> Art. 10 - Les documents administratifs sont communiqués sous réserve des droits de propriété littéraire et artistique. L'exercice du droit à la communication institué par le présent titre exclut, pour ses bénéficiaires ou pour les tiers, la possibilité de reproduire, de diffuser ou d'utiliser à des fons commerciales les documents communiqués[201].

Par ailleurs, comme l'analyse Maxime Michali[202], la réutilisation, même à des fins non commerciale peut donner lieu au paiement d'une redevance, dont le

---

[198] Commission d'accès aux documents administratifs (CADA)

[199] http://www.legifrance.gouv.fr/affichTexte.do;jsessionid=5D05BF827EBE061D3ECDC7BFA2E64337.tpdjo06v_1?cidTexte=JORFTEXT000000339241&dateTexte=20991231 (consulté le

[200] Commissariat général au Plan, *Rapport sur la diffusion des données publiques et révolution numérique*, présidé par Dieudonné Mandelkern, La documentation française, Novembre 1999, p. 69.

[201] Version initiale de la loi n° 78-753 du 17 juillet 1978, dite loi CADA, JORF du 18 juillet 1978, p. 2851. http://www.legifrance.gouv.fr/jopdf/common/jo_pdf.jsp?numJO=0&dateJO=19780718&numTexte=&pageDebut=02851&pageFin= (consulté le 21/09/2012)

[202] Maxime Michali, « Edile Max, Joies et peines d'un doctorant en droit », *Blog d'un thésard sur l'Open data.* http://edilemax.wordpress.com/ (consulté le 21/09/2012). Actuellement en thèse de droit à l'Université d'Aix-Marseille 3 sous la direction de Laure Merland, Maxime Michali travaille depuis le 1er septembre 2011 sur « L'open-data :

(cc) BY-NC-ND

montant est assez librement fixé par l'administration, puisque la loi n'impose aucune gratuité de principe, contrairement à l'accès. S'il y a une redevance, l'administration doit fournir une licence de réutilisation et le non respect de cette licence expose le fraudeur à une amende appliquée par la CADA. Néanmoins, le champ d'application de la loi dans le texte et la jurisprudence est très large puisqu'il concerne : les administrations de l'Etat, les collectivités locales, les établissements publics et les organismes de droit privé chargés d'une mission de service public. Un périmètre que l'on tente aujourd'hui d'atteindre dans le domaine de l'*Open data*.

Dans la continuité de la loi CADA, **la loi n° 79-18 du 3 janvier 1979 sur les archives** parfois omis des textes de référence de l'*Open data*, vient pourtant préciser les conditions de diffusion des données publiques puisqu'elle régit l'accès à une partie des ressources publiques. Elle établit que tout document public devient une archives dès après sa publication[203] et précise également que les conditions d'accès et de diffusion de certains documents publics peuvent faire l'objet de délais particuliers et dérogatoires à ceux prévus par le droit commun.

> Art. 1 – Les archives sont l'ensemble des documents, quels que soient leur date, leur forme et leur support matériel produits ou reçus par toute personne physique ou morale, et par tout service ou organisme public ou privé, dans l'exercice de leu activité. […]
>
> Art. 6 – Les documents dont la communication était libre avant leur dépôt aux archives publiques continueront d'être communiqués sans restriction d'aucune sorte à toute personne qui en fera la demande. Les documents visés à l'article 1er de la 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 demeurent communicables dans les conditions fixées par cette loi. Tous les autres documents d'archives publiques pourront être librement consultés à l'expiration d'un délai de trente ans ou des délais spéciaux prévus à l'article 7 ci-dessous[204].

Si ce cadre juridique de la fin des années 1970 est resté pendant plus de 10 ans inchangé, les questions d'accès et de diffusion des données publiques n'a pour autant cesser d'animer les débats au sein des instances ministériels. Parmi les nombreuses circulaires émises pour tenté de définir les droits et obligations des administrations d'Etat en matière de diffusion de l'information publique, la circulaire du 14 février 1994 relative à la diffusion des données publiques[205], dite « Circulaire Balladur » pose un jalon sur le plan de la diffusion puisqu'elle fait non

---

l'encadrement et le enjeux de la libération des données publiques ». C'est à l'occasion d'une rencontre à Lyon, le 12 septembre 2012, que celui-ci m'a parlé de son blog et de sa démarche de « médiation juridique » vis-à-vis des enjeux de l'Open data.

[203] Commissariat général au Plan, *Rapport sur la diffusion des données publiques et révolution numérique*, présidé par Dieudonné Mandelkern, La documentation française, Novembre 1999, p. 71.

[204] Version initiale de la loi n° 79-18 du 3 janvier 1979 sur les archives, JORF du 8 janvier 1979, p. 43. http://www.legifrance.gouv.fr/jopdf/common/jo_pdf.jsp?numJO=0&dateJO=19790105&numTexte=&pageDebut=00043&pageFin= (consulté le 21/09/2012)

[205] Cette circulaire a notamment été élaborée à la suite du *Rapport sur la commercialisation des données publiques*, remis en 1992 par le professeur Gaudrat à l'Observatoire juridique des technologies de l'information. Il y fait le point sur trois grands axes : la nature des informations à transmettre ; dans quelles mesures il est possible (juridiquement) de les valoriser économiquement ; et enfin, quels sont les problèmes posés par la tarification de ce service en regard de la vocation de gratuité des administrations.

(cc) BY-NC-ND

seulement apparaître la notion de « donnée » mais tend aussi vers un principe de gratuité[206] en rappelant néanmoins que cette diffusion doit se faire dans le respect des principes de liberté du commerce et de l'industrie, du droit de la concurrence et avec des moyens proportionnés à l'intérêt pour le public.

> Pour exercer ses missions, l'administration collecte, enregistre, conserve et traite d'importantes quantités de données. […] Dans cette perspective, les administrations ou services publics qui disposent d'informations non confidentielles présentant un intérêt pour le public et notamment pour les entreprises doivent pouvoir en envisager la diffusion. […] Par elle-même, la détention d'information par l'administration ne confère pas à cette dernière le droit de l'exploiter directement de façon commerciale. La commercialisation des données publiques ne s'inscrit pas en effet dans la vocation première d'une collectivité publique. […] Sauf circonstances particulières, je ne vois que des avantages à ce que l'administration détentrice confie la diffusion de données publiques aux organismes publics ou privés spécialisés. En particulier, les entreprises privées spécialisées: disposent d'un équipement informatique adapté encourent une responsabilité commerciale de droit commun à l'égard des utilisateurs, peuvent offrir un service plus étendu et plus personnalisé qu'une administration classique[207].

Ce besoin d'approfondissement et de retour sur la question des données publiques et les textes législatifs de 1978 et 1979 se retrouvent également dans le Rapport Braibant sur les Archives de France remis au premier ministre Alain Juppé le 28 mai 1996. En dressant le bilan de la loi de 1979, le rapport relie étroitement le la question des données publiques aux enjeux de démocratie, de transparence et de mémoire tant à l'échelle nationale qu'en terme de comparaison internationale.

> L'accès aux documents et archives publics apparaît désormais comme une garantie fondamentale des droits des citoyens dans les sociétés démocratiques. L'évolution politique des pays d'Europe centrale et orientale s'est accompagnée depuis 1989 d'une ouverture spectaculaire en la matière. Le mouvement général vers la transparence en France et dans les pays comparables, l'ouverture des archives à l'étranger qui, combinée avec le développement d'Internet, conduit à un système d'«.archives sans frontières.», l'accélération de l'Histoire et le vieillissement rapide des secrets, tout porte à remettre en cause un dispositif qui a constitué un progrès en son temps mais qui est en train de devenir anachronique[208].

---

[206] *Rapport sur la diffusion des données publiques et révolution numérique*, présidé par Dieudonné Mandelkern, La documentation française, Novembre 1999, p. 74 : La circulaire n'affirme pas de principe de gratuité mais prévoit en revanche que « la tarification ne doit pas procuré de bénéfice à l'administration, que les coûts de production ne soient pas répercutés lorsqu'ils ont déjà été financés sur fonds publics et que la contribution aux frais généraux s'élève aux coûts directement imputables à la fourniture du service. Commissariat général au Plan.

[207] Circulaire Balladur du 14 février 1994 relative à la diffusion des données publiques, JORF n°42 du 19 février 1994, p. 2864.
http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000181007&dateTexte=&categorieLien=id (consulté le 21/09/2012)

[208] Guy Braibant, *Rapport sur les archives de France* remis à Alain Juppé, La documentation française, 28 mars 1996, p. 47.
http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/964093000/0000.pdf (Consulté le 22/09/2012)

(cc) BY-NC-ND

Si cette urgence d'ouverture, exprimée dans le rapport Braibant, est, dans ce cas précis, surtout reliée à la question des archives, ces dernières n'en constituent pas moins un type de données publiques à partir duquel le débat peut être élargi.

En 1998, un projet de loi en date du 13 mai [209] relatif aux droits des citoyens dans leurs relations avec les administrations, susceptible de modifier une partie des éléments du cadre juridique applicable à la diffusion des données publiques, est adopté en Conseil des Ministres. L'exposé des motifs insiste sur :

> […] la volonté de combattre le secret administratif et de développer corrélativement la transparence des administrations face aux citoyens […] En effet, les citoyens et les entreprises, malgré les efforts engagés, et ceux déployés par les agents publics, perçoivent encore les administrations comme lointaines et impersonnelles[210].

Le projet est finalement voté comme la loi éponyme n° 2000-321 du 12 avril 2000. L'article 2 du texte tente de répondre au constat précédent, en prévoyant en matière de diffusion que :

> Art. 2 - Le droit de toute personne à l'information est précisé et garanti par le présent chapitre en ce qui concerne la liberté d'accès aux règles de droit applicables aux citoyens. Les autorités administratives sont tenues d'organiser un accès simple aux règles de droit qu'elles édictent. La mise à disposition et la diffusion des textes juridiques constituent une mission de service public au bon accomplissement de laquelle il appartient aux autorités administratives de veiller[211].

Au moins pour les documents d'ordre législatif, l'initiative de la diffusion est donc replacée du côté des institutions et non plus des citoyens. Cette priorité législative est confirmée par la circulaire du 28 janvier 1999 relative à diffusion gratuite des rapports officiels sur internet. Cette circulaire organise la création d'une « bibliothèque numérisée des rapports officiels » géré par la Documentation française.

> La diffusion gratuite des données publiques essentielles sur l'Internet est l'un des objectifs prioritaires retenus par le programme d'action du Gouvernement pour la société de l'information. Des innovations importantes ont déjà été réalisées dans ce sens, notamment avec la création des sites Légifrance et Admifrance. Il convient maintenant d'appliquer les mêmes principes à la diffusion des rapports officiels. C'est le moyen d'accroître la participation des citoyens au débat politique et à la préparation des décisions administratives, de s'affranchir des limites inhérentes à la diffusion des documents imprimés et

---

[209] *Rapport sur la diffusion des données publiques et révolution numérique*, présidé par Dieudonné Mandelkern, La documentation française, Novembre 1999, p. 72 : Ce projet de loi a été adopté par le Conseil des Ministres le 13 mai 1998 après discussion en première lecture par le Sénat le 10 mars 1998 et par l'Assemblée nationale le 27 mai 1999, puis retransmis au Sénat pour un seconde lecture.

[210] Projet de loi du 13 mai 1998 relatif aux droits des citoyens dans leurs relations avec les administrations, délibéré en Conseil des ministres après avis du Conseil d'Etat. http://www.assemblee-nationale.fr/11/projets/pl0900.asp (consulté le 23/09/2012)

[211] Loi n° 2000-321 du 12 avril 2000 relative aux droits des citoyens dans leurs relations avec les administrations, JORF n°88 du 13 avril 2000, p. 5646 http://www.legifrance.gouv.fr/affichTexte.do;jsessionid=DFEBFA9EEBB3A85BDC9324031E9AA270.tpdjo05v_3?cidTexte=JORFTEXT000000215117&categorieLien=id (consulté le 23/09/2012)

(cc) BY-NC-ND

de donner une meilleure vue sur l'action publique, notamment depuis les pays étrangers[212].

Pour le reste des documents, les tarifications de référence sont précisées par l'arrêté du 1 octobre 2001 relatif aux conditions de fixation et de détermination du montant des frais de copie d'un document administratif[213].

L'INFLUENCE EUROPEENNE

La gestion des données publiques est une préoccupation qui anime également les débats communautaire, or en tant que membre de l'Union Européenne et par respect du principe de subsidiarité, la France a pour obligation légale de transposer les décisions qui en sont issues pour les rendre applicables par ses autorités administratives et judiciaires. C'est en grande partie par le biais de la réglementation communautaire que le processus d'ouverture des données publiques a pris un nouveau tournant en France depuis le milieu des années 2000. Plusieurs sources législatives européenne relatives à la diffusion des données publiques ont ainsi fait l'objet de transposition en droit français. En premier lieu, **la directive européenne 90/313/CEE** concernant l'accès aux données environnementales de 1990 qui a abouti en 2002 sur la Convention d'Aarhus, elle-même transcrite en 2003 dans la directive européenne 2003/4/CE et en 2005 en droit français. Elle définit « le droit de toute personne d'accéder aux informations relatives à l'environnement détenues, reçues ou établies par les autorités publiques» (Art. L. 124-1)[214]. Elle est suivie, en 2007, par la directive européenne Inspire n°2007/2/CE du 14 mars 2007 établissant une infrastructure d'information géographique dans la communauté européenne, transposée en droit français par l'ordonnance n°2010-1232 du 21 octobre 2010 dans les articles L.127-1 à L.127-10 du code de l'environnement des produits.

Mais la directive majeure dans le domaine reste néanmoins **la directive européenne PSI n°2003/98/CE du 17 novembre 2003**[215] concernant, la réutilisation des informations du secteur public. Transposée en droit français par l'**ordonnance n°2005-650 du 6 juin 2005** relative à la liberté d'accès aux documents administratifs et à la réutilisation des informations publiques, elle vient modifier la loi CADA en instaurant non seulement une obligation d'ouverture mais

---

[212] Circulaire du 28 janvier 1999 relative à la diffusion gratuite des rapports officiels sur l'Internet, JORF n°27 du 2 février 1999 p. 1678.
http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000743424&dateTexte
(Consulté le 23/09/2012)

[213] Arrêté du 1er octobre 2001 relatif aux conditions de fixation et de détermination du montant des frais de copie d'un document administratif, JORF n°228 du 2 octobre 2001, p. 15496.
http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000772503&dateTexte=vig (consulté le 23/09/2012)

[214] Loi n° 2005-1319 du 26 octobre 2005 portant diverses dispositions d'adaptation au droit communautaire dans le domaine de l'environnement, JORF n°251 du 27 octobre 2005, p.16929
http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000635725&dateTexte=&categorieLien=id
(consulté le 23/09/2012)

[215] Plusieurs réexamens de cette directive ont donné lieu à des demandes de précisions sur des termes trop ambigus comme « la définition des documents, des missions de service public, de la tarification fondée sur les coûts marginaux ou du retour sur investissement raisonnable. » Un réexamen est de nouveau en cours et pourrait mener à une extension du champs d'application de la directive en terme de tarification, réutilisation ou encore de listes de ressources. Cf. Proposition de directive du parlement européen et du conseil modifiant la directive 2003/98/CE concernant la réutilisation des informations du secteur public, Bruxelles, 12 décembre 2011.
http://ec.europa.eu/information_society/policy/psi/docs/pdfs/directive_proposal/2012/fr.pdf
(Consulté le 23/09/2012)

(cc) BY-NC-ND

aussi un droit de réutilisation des données publiques. Sans faire mention explicite de l'exclusion de réutilisation à des fins commerciales, elle ouvre ainsi la voie vers les enjeux économique de l'ouverture des données publiques, un argument majeur dans l'adoption d'une démarche Open data.

> Art.10 - Les informations figurant dans des documents élaborés ou détenus par les administrations mentionnées à l'article 1er, quel que soit le support, peuvent être utilisées par toute personne qui le souhaite à d'autres fins que celles de la mission de service public pour les besoins de laquelle les documents ont été élaborés ou sont détenus[216].

Dans la continuité le décret n°2005-1755 du 30 décembre 2005[217] puis la circulaire du Premier Ministre n°5156/SG du 29 mai 2006[218] sont venus apporter quelques précisions notamment pour clarifier la réutilisation à des fins commerciales autant que non commerciales des données publiques. Les principes posés par le droit européen ont ainsi fourni le cadre nécessaire à la mise en place d'une démarche *Open data* en France, en inscrivant juridiquement que les données publiques doivent être accessibles et réutilisables, à des fins commerciales ou non, d'une manière, non discriminatoire et non exclusive à des tarifs qui n'excèdent pas leur coût de mise à disposition et éventuellement de production.

---

[216] Ordonnance n° 2005-650 du 6 juin 2005 relative à la liberté d'accès aux documents administratifs et à la réutilisation des informations publiques, JORF n°131 du 7 juin 2005 p. 10022.
http://www.legifrance.gouv.fr/affichTexte.do;jsessionid=5CEAE63657B7220DA2A438FF7F1EC87A.tpdjo05v_3?cidTexte=JORFTEXT000000629684&categorieLien=id (consulté le 23/09/19)

[217] Décret n° 2005-1755 du 30 décembre 2005 relatif à la liberté d'accès aux documents administratifs et à la réutilisation des informations publiques, pris pour l'application de la loi n° 78-753 du 17 juillet 1978, JORF n°304 du 31 décembre 2005 p. 20827.
http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000265304&dateTexte=&categorieLien=id (consulté le 23/09/2012)

[218] La circulaire vient notamment préciser les conditions de tarifications et le cadre légal de leur réutilisation tant relativement aux licences choisies par les administrations que par le respect de la concurrence en matière de diffusion et de réutilisation des données publiques.
http://www.uved.eseq.fr/assets/pdf/circulaire_du_premier_ministre_n_5156_sg_du_29_mai_2006_relative_a_la_reforme_des_dispositions_regissant_l_acces_aux_documents_administratifs_et_a_l_institution_d_un_droit_de_reutilisation_des_informations_publiques.pdf (consulté le 23/09/2012)

(cc) BY-NC-ND

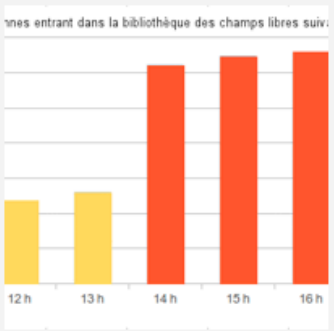
OPERATION #BIBLIOVIZ, LES REALISATIONS

[Infographie] Fréquentation de la bibliothèque par mois en 2011. Auteur : Auregann. Licence CC-BY-SA.

[Infographie] Fréquentation de la bibliothèque par semaine en 2011. Auteur : Auregann. Licence CC-BY-SA.

[Infographie] Fréquentation de la bibliothèque par tranche horaire dans la semaine en 2012. Auteur : Auregann. Licence CC-BY-SA.

[Application] Fréquentation moyenne par jour de la semaine. Auteur : Auregann. Licence CC-BY-SA.



[Infographie] Entrées par tranche horaire. Auteur : Nieluje.

Source : http://opendatarennes.wordpress.com/operation-biblioviz-les-realisations/

(cc) BY-NC-ND

LICENCE DE REUTILISATION DES INFORMATIONS PUBLIQUES DE RENNES METROPOLE

Direction Générale de l'Information et de la Communication
Licence« *Rennes Métropole en libre accès* »

LICENCE DE REUTILISATION DES INFORMATIONS PUBLIQUES
RENNES METROPOLE EN ACCES LIBRE

Pour favoriser la création de nouveaux produits et services et contribuer au développement économique, la loi n°78-753 du 17 juillet 1978 crée un droit de réutilisation des informations publiques. Ainsi, les informations qui entrent dans le champ de cette loi peuvent être exploitées à d'autres fins que celles de la mission de service public pour les besoins de laquelle elles ont été produites, y compris commerciales.

Les présentes dispositions rappellent les droits et obligations des réutilisateurs des informations publiques accessibles sur le site http://www.data.rennes-metropole.fr

Article 1 - Définitions
Dans le cadre de la présente licence de réutilisation d'informations publiques délivrée en application de la loi n°78-753 du 17 juillet 1978 les termes ci-dessous ont la signification suivante :

Administration désigne la personne publique mettant à disposition les informations publiques en vue de leur réutilisation.
Informations désignent les informations publiques y compris leurs mises à jour successives, objet de la Licence.
Licence désigne les conditions générales de réutilisation des Informations et de leurs mises à disposition.
Licencié ou Réutilisateur désigne la personne physique ou morale souhaitant réutiliser les Informations dans le cadre de la Licence.

Article 2 - Objet
La présente licence a pour objet d'autoriser les utilisateurs à partager, modifier et utiliser librement les données publiques disponibles sur le site www.data.rennes-metropole.fr, sous réserve que la modification ne constitue pas une altération des données ou une dénaturation de leur sens.
Les présentes conditions générales précisent les droits et obligations applicables à la réutilisation des informations publiques, au sens de la loi n° 78-753 du 17 juillet 1978.

Article 3 - Droits conférés
La Licence confère au Licencié un droit personnel et non exclusif de réutilisation des Informations mises à sa disposition sur le site www.data.rennes-metropole.fr.

La réutilisation des Informations publiques est gratuite. Ces droits permettent expressément les utilisations commerciales ou non commerciales.

Les droits octroyés concernent :

- L'extraction et la réutilisation de la totalité ou d'une partie substantielle du contenu ;
- La création de bases de données dérivées;
- La création de bases de données collaboratives;
- La création de reproductions temporaires ou permanentes, par tous moyens et sous quelque forme que ce soit, en tout ou en partie, y compris de toute base de données dérivée ou en tant que partie d'une base de donnée collaborative;
- La distribution, la communication, l'affichage, la location, la mise à disposition ou la diffusion au public, par tous moyens et sous quelque forme que ce soit, en tout ou en partie, y compris de toute base de données dérivée ou en tant que partie d'une base de données collaborative.

La Licence ne confère en aucun cas les droits de propriété intellectuelle de ces données.

(cc) BY-NC-ND

Le Licencié est autorisé dans les mêmes conditions à exploiter les Informations publiques lorsque l'Administration est titulaire de droits de propriété intellectuelle, sous réserve du droit moral. La reproduction des marques et logos qui peuvent y figurer reste interdite.

Article 4 - Obligations du Licencié
Conformément à l'article 12 de la Loi, le Licencié s'engage à ce que les Informations ne soient pas altérées ni leur sens dénaturé.

Le Licencié s'engage à indiquer leur source et la date de leur dernière mise à jour.

Le Licencié fait son affaire, le cas échéant, des moyens à mettre en œuvre pour que ses systèmes informatiques soient en adéquation avec le format des Informations et/ou les modalités de mises à disposition des Informations.

Dans l'hypothèse où des évolutions, liées notamment au changement de format, à la modification de structure des bases de données, ou de fichier et/ou d'évolution des modalités de mise à disposition des données, induiraient pour le Licencié des adaptations techniques de ses équipements, le surcoût qui en découlerait ne pourrait en aucun cas être mis à la charge de l'Administration. Le Licencié ne peut prétendre à aucune indemnité ni compensation à ce titre.

Les obligations ci-dessus mentionnées demeurent applicables pendant toute la durée de réutilisation des Informations, y compris en cas de cessation de mise à disposition des Informations pour quelque cause que ce soit.

Le Licencié s'engage, sans restriction ni réserve, à respecter la Licence et la réglementation en vigueur. Le Licencié s'abstient de tout usage contraire aux lois et règlements ou portant atteinte à l'ordre public.

Toute base de données dérivée doit impérativement respecter les conditions de la présente Licence ou d'une licence compatible.

Article 5 - - Garanties et responsabilités
Le Licencié reconnaît et accepte que les Informations sont fournies par l'Administration en l'état, telles que détenues par l'Administration dans le cadre de sa mission, sans autre garantie, expresse ou tacite.
Le Licencié exploite les Informations publiques sous sa seule responsabilité.
L'Administration et le Licencié s'informeront réciproquement de toute difficulté qui surviendrait dans la mise en œuvre de la réutilisation des Données dans les conditions prévues par la présente Licence.

Article 6 - Modifications
Le Licencié est averti que l'administration se réserve le droit à tout moment, et pour quel que motif que ce soit, de modifier les conditions de réutilisation fixées dans la présente Licence, de mettre un terme à leur production ou modifier leur contenu ou la fréquence de leur mise à jour, sans que le réutilisateur ne puisse prétendre à aucune indemnité, .
Cette modification fera l'objet d'une mise en ligne sous la forme d'une alerte qui sera diffusée sur le site www.data.rennes-metropole.fr.

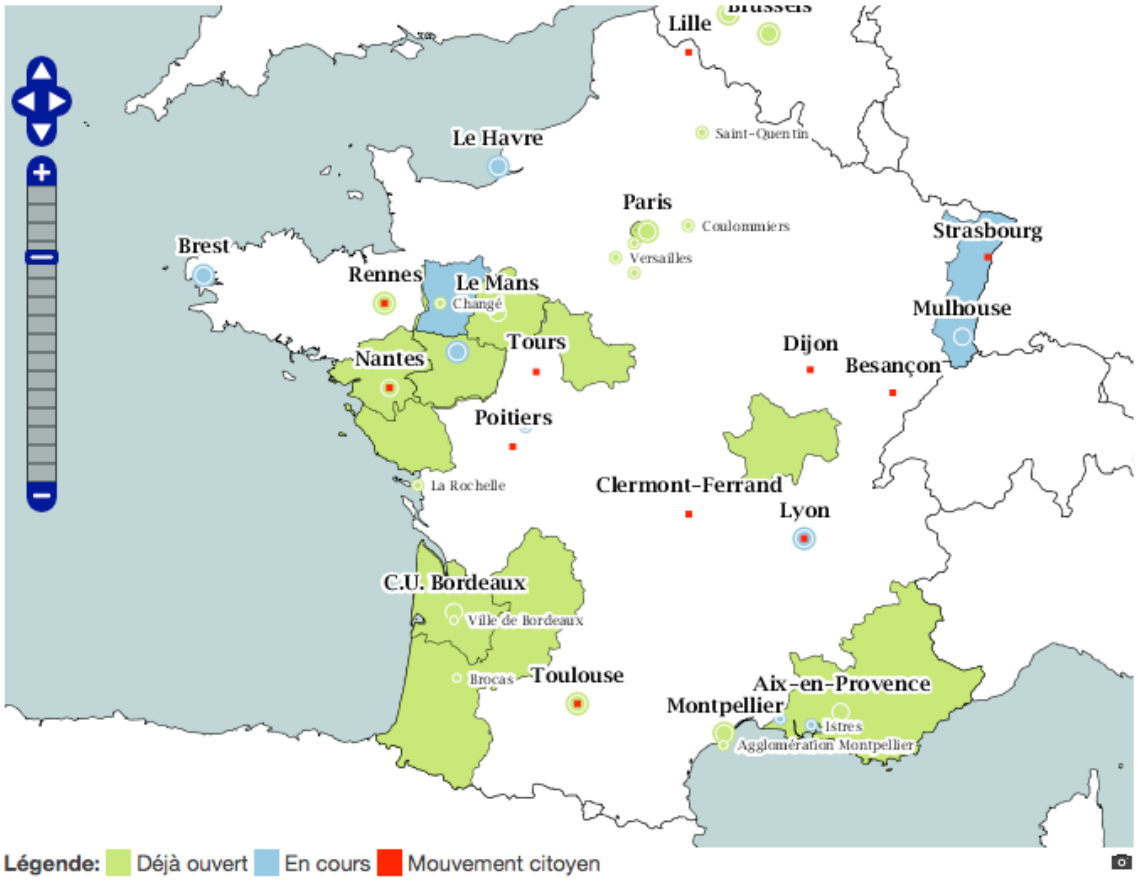
Article 7 - Durée
Le Licencié est autorisé à exploiter les Informations publiques sans limitation de durée.

Article 8 - Loi applicable et recours
Les Conditions générales sont soumises à la seule loi française.

Les différends ou litiges en relation avec les Conditions générales peuvent être soumis à la Commission d'accès aux documents administratifs en application des dispositions de l'article 20 de la loi n°78-753 du 17 juillet 1978.

(cc) BY-NC-ND

CARTE DE FRANCE DE L'OPEN DATA



Source : http://libertic.wordpress.com/2012/07/16/la-carte-de-france-de-lopen-data-au-quotidien/

(cc) BY-NC-ND

LICENCE OUVERTE ETALAB

LICENCE OUVERTE
OPEN LICENCE

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

LA RÉUTILISATION DE L'INFORMATION DIFFUSÉE SOUS CETTE LICENCE

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

VOUS ÊTES LIBRE DE RÉUTILISER « L'INFORMATION » :

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

SOUS RÉSERVE DE :

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

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

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

(cc) BY-NC-ND

RESPONSABILITÉ

« L'Information » est mise à disposition telle que produite ou reçue par le « Producteur », sans autre garantie expresse ou tacite qui n'est pas prévue par la présente licence.

Le « Producteur » garantit qu'il met à disposition gratuitement « l'Information » dans les libertés et les conditions définies par la présente licence. Il ne peut garantir l'absence de défauts ou d'irrégularités éventuellement contenues dans « l'Information ». Il ne garantit pas la fourniture continue de « l'Information ». Il ne peut être tenu pour responsable de toute perte, préjudice ou dommage de quelque sorte causé à des tiers du fait de la réutilisation.

Le « Réutilisateur » est le seul responsable de la réutilisation de « l'Information ». La réutilisation ne doit pas induire en erreur des tiers quant au contenu de « l'Information », sa source et sa date de mise à jour.

DROITS DE PROPRIÉTÉ INTELLECTUELLE

Le « Producteur » garantit que « l'Information » ne contient pas de droits de propriété intellectuelle appartenant à des tiers.

Les éventuels « Droits de propriété intellectuelle » détenus par le « Producteur » sur des documents contenant « l'Information » ne font pas obstacle à la libre réutilisation de « l'Information ». Lorsque le « Producteur » détient des « Droits de propriété intellectuelle » sur des documents qui contiennent « l'Information », il les cède de façon non exclusive, à titre gracieux, pour le monde entier et pour toute la durée des « Droits de propriété intellectuelle », au « Réutilisateur » qui peut en faire tout usage conformément aux libertés et aux conditions définies par la présente licence.

COMPATIBILITÉ DE LA PRÉSENTE LICENCE

Pour faciliter la réutilisation des « Informations », cette licence a été conçue pour être compatible avec toute licence libre qui exige *a minima* la mention de paternité. Elle est notamment compatible avec les licences « Open Government Licence » (OGL) du Royaume-Uni, « Creative Commons Attribution 2.0 » (CC-BY 2.0) de Creative Commons et « Open Data Commons Attribution » (ODC-BY) de l'Open Knowledge Foundation.

DROIT APPLICABLE

La présente licence est régie par le droit français.

OCTOBRE 2011

BY-NC-ND

DÉFINITIONS

DROITS DE PROPRIÉTÉ INTELLECTUELLE*
Il s'agit des droits identifiés comme tels par le Code de la propriété intellectuelle (droit d'auteur, droits voisins au droit d'auteur, droit sui generis des bases de données).

INFORMATION*
Il s'agit des données ou des informations proposées à la réutilisation dans les libertés et les conditions de cette licence.

INFORMATIONS DÉRIVÉES*
Il s'agit des nouvelles données ou informations qui ont été créés soit directement à partir « d'Informations », soit à partir d'une combinaison « d'Informations » et d'autres données ou informations qui ne seraient pas soumises à cette licence.

PRODUCTEUR*
Il s'agit de l'entité qui produit « l'Information » et l'ouvre à la réutilisation dans les libertés et les conditions prévues par cette licence.

RÉUTILISATEUR*
Il s'agit de toute personne physique ou morale qui réutilise « l'Information » conformément aux libertés et aux conditions de cette licence.

À PROPOS DE LA LICENCE OUVERTE

Etalab est la mission chargée sous l'autorité du Premier ministre d'ouvrir le plus grand nombre de données publiques des administrations de l'Etat et de ses établissements publics. Elle a réalisé la Licence Ouverte pour faciliter la réutilisation libre et gratuite de ces informations publiques, telles que définies par l'article 10 de la loi n°78-753 du 17 juillet 1978.

Dans le cadre de leurs missions de service public, les administrations produisent ou reçoivent des informations publiques qui peuvent être réutilisées par toute personne physique ou morale à d'autres fins que celles de la mission de service public.

Ne sont pas des informations publiques au sens de la loi du 17 juillet 1978 les informations contenues dans des documents dont la communication ne constitue pas un droit (en application de la loi du 17 juillet 1978 ou d'autres dispositions législatives, sauf si ces informations font l'objet d'une diffusion publique), celles contenues dans des documents produits ou reçus par les administrations dans l'exercice d'une mission de service public à caractère industriel ou commercial, et celles contenues dans des documents sur lesquels des tiers détiennent des droits de propriété intellectuelle.

Ne sont également pas des informations publiques susceptibles d'être réutilisées celles qui contiennent des données à caractère personnel, sauf lorsque les personnes intéressées y ont consenti, ou lorsqu'elles ont fait l'objet d'une anonymisation par l'administration, ou lorsqu'une disposition légale ou réglementaire le permet (dans ces trois cas, la réutilisation est subordonnée au respect de la loi n°78-17 du 6 janvier 1978).

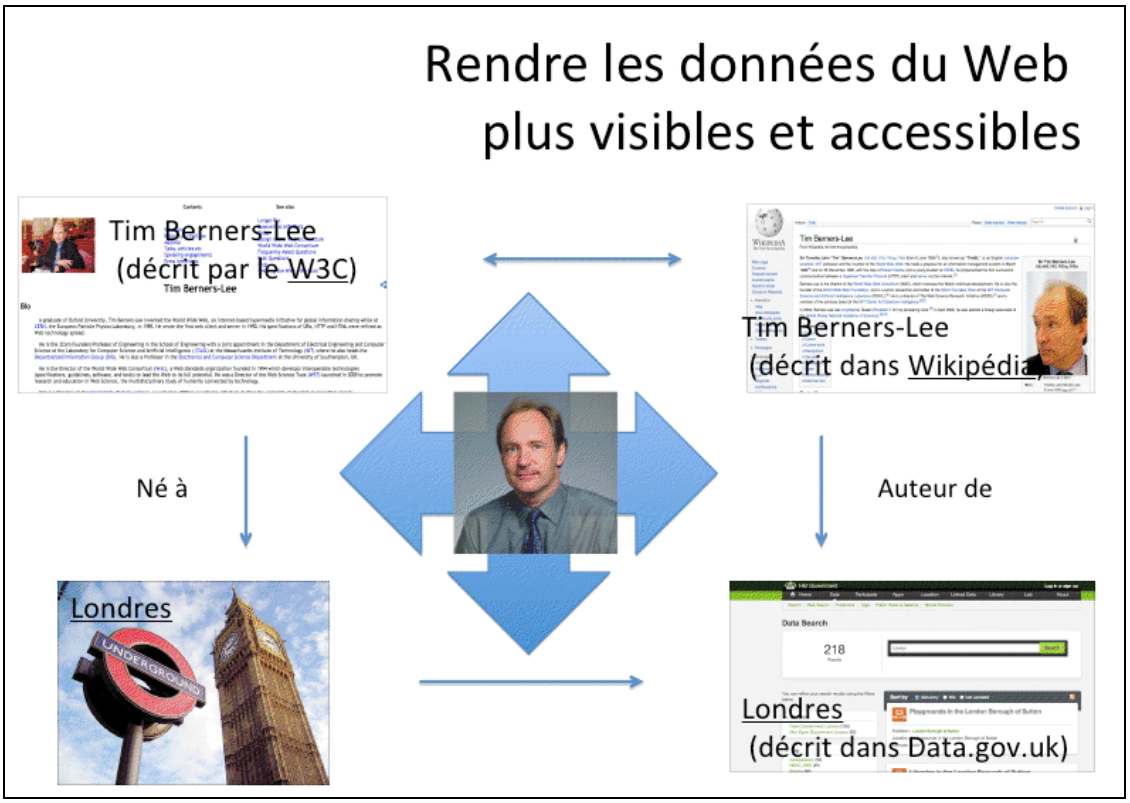
Cette licence est une version 1.0 de la Licence Ouverte. Etalab se réserve la faculté de proposer de nouvelles versions de la Licence Ouverte. Cependant, les réutilisateurs pourront continuer à réutiliser les informations disponibles sous cette licence s'ils le souhaitent.

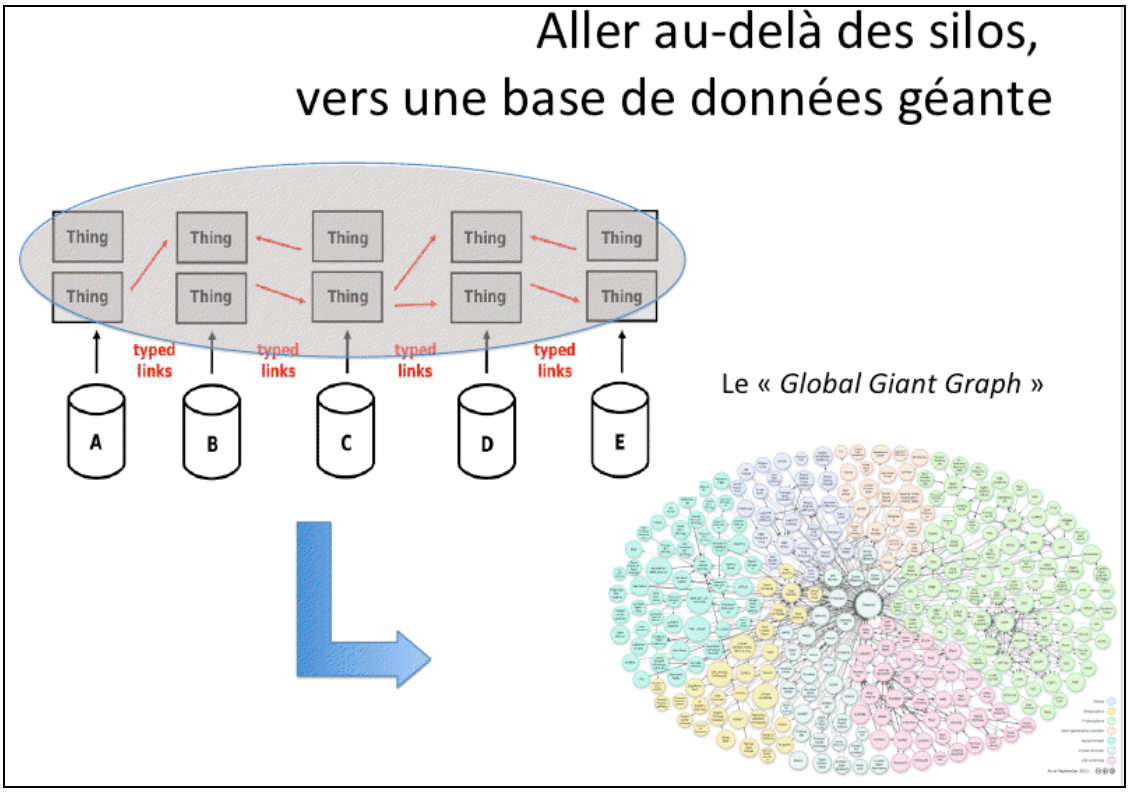
OCTOBRE 2011

etalab gouv.fr — data.gouv.fr BETA innovation transparence ouverture

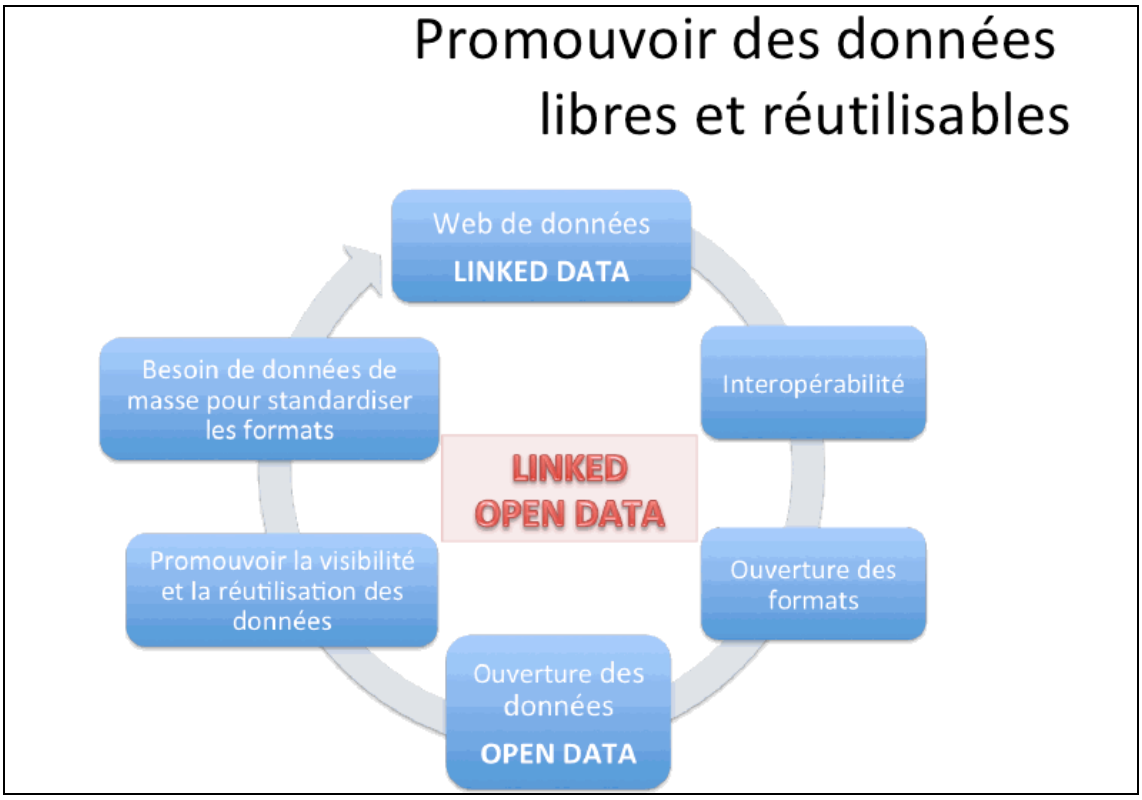
(cc) BY-NC-ND

LINKED OPEN DATA : ENJEUX ET FONCTIONNEMENT





(cc) BY-NC-ND


Promouvoir des données libres et réutilisables
Web de données
LINKED DATA
Interopérabilité
Ouverture des formats
Ouverture des données
OPEN DATA
Promouvoir la visibilité et la réutilisation des données
Besoin de données de masse pour standardiser les formats
LINKED OPEN DATA

(cc) BY-NC-ND

Glossaire

API (Application Programming Interface)

Interface à un ou plusieurs jeux de données permettant à un programme ou une application d'y accéder. Les jeux de données peuvent être rendus accessibles soit par téléchargement (pour les jeux de données raisonnablement stables dans le temps) soit par API (pour les jeux de données très volumineux ou très volatiles).

APIE (Agence pour le patrimoine immatériel de l'État)

Créée en 2007 pour valoriser le patrimoine immatériel de l'Etat (brevets, marques, œuvres d'art, données, etc). Ses missions sont d'optimiser l'impact de la gestion du patrimoine immatériel sur l'économie, de tirer parti d'une meilleure valorisation des actifs pour moderniser les services publics, de soutenir la conduite des politiques publiques au profit des usagers et contribuer au désendettement, et de prémunir l'État et les usagers contre d'éventuels risques de détournement. L'APIE a été initialement chargée de mettre en place le répertoire des données publiques, mission qui a finalement été confiée à Etalab.

CADA (Commission d'Accès au Documents Administratifs)

Elle sert de recours aux citoyens qui ont des problèmes pour accéder aux données publiques. La CADA est une autorité administrative indépendante dont le rôle est consultatif. Elle peut aider un citoyen à obtenir un document administratif individuel qui lui a été refusé. Elle répond aux consultations des administrations qui la saisissent sur le caractère communicable ou réutilisable des documents qu'elles détiennent ou sur les modalités de leur communication ou de leur réutilisation. Elle intervient pour tous les documents détenus par un service de l'Etat, une collectivité territoriale, un établissement public ou un organisme chargé de la gestion d'un service public, que cet organisme soit public ou privé.

CKAN (Comprehensive Knowledge Archive Network)

Logiciel open source pour portail Open Data permettant de publier, de partager et de trouver des données. Il inclut une base de données pour cataloguer et stocker des jeux de données.

COEPIA (Conseil d'orientation de l'édition publique et de l'information administrative)

Créé le 10 janvier 2010, placé auprès du Premier ministre, exerce une fonction d'évaluation, d'expertise et de conseil dans les domaines de l'édition publique et les publications administratives, de l'information et le renseignement administratifs et de la mise à disposition des données publiques. Le Coepia a notamment avalisé la liste des jeux de données publiques payantes.

CROWDSOURCING

Externalisation de masse : recueil de données par une communauté de volontaires.

DATAVIZ

Technologies, méthodes et outils de visualisation des données.

(cc) BY-NC-ND

DUBLIN CORE

Système simple pour décrire les ressources (de tout type) diffusées sur le web. La structure générale comprend 15 éléments qui constituent le "Dublin Core non qualifié" auxquels il est possible d'affecter des qualicatifs et attributs qui permettent de préciser la nature du contenu de l'élément ou de lui attacher un vocabulaire contrôlé : "Dublin Core qualifié". Le projet est animé dans le cadre d'un travail coopératif (DCMI = Dublin Core Metadata Initiative) mené à l'échelle international.

FRBR (Functional Requirements of Bibliographic Records)

Modèle conceptuel de données bibliographiques, validé en 1997 par l'IFLA. Ce modèle définit les entités sur lesquelles porte l'information catalographique, ainsi que les relations qui existent entre ces entités. Il analyse la notice catalographique comme un ensemble de données reflétant essentiellement les attributs et les relations de quatre entités distinctes: Œuvre, Expression, Manifestation et Item.

RDA (Resource Description and Access)

Nouveau code de catalogage, à vocation internationale, ayant pour but d'adapter les catalogues à l'environnement numérique et aux nouveaux modes d'accès à l'information.

RDF (Resource Description Framework)

Modèle pour l'échange de données sur le Web. C'est un des éléments constitutifs du Web sémantique, qui décrit les données, ainsi que les relations qui existent entre elles, selon la syntaxe : sujet – prédicat – objet.

SKOS (Simple Knowledge Organisation System**)**

Langage permettant la représentation formalisée de thésaurus, répertoires de vedette-matière et autres types de vocabulaires contrôlés et structurés, ainsi que l'expression de correspondances (alignements) entre vocabulaires. SKOS est développé par le W3C dans le cadre du Web sémantique.

WEB SEMANTIQUE/ WEB DE DONNEES

Ensemble de technologies visant à rendre le contenu des ressources du World Wide Web accessible et utilisable par les programmes et agents logiciels, grâce à un système de métadonnées formelles, utilisant notamment la famille de langages développés par le W3C. Le Web sémantique repose notamment sur le modèle RDF (Resource Description Framework) destiné à décrire les données et leurs relations entre elles selon la syntaxe suivante : sujet - prédicat - objet. L'intérêt du Web sémantique pour les bibliothèques réside dans la capacité à créer davantage de liens entre les données des catalogues, dans la même conception que les modèles de données tels que les FRBR, et à les rendre plus visibles sur le Web.

(cc) BY-NC-ND

Index

A

ABES 11, 51, 68, 75, 83
Accountability 20
ALBANEL, Christine 45
API 11, 26, 67
ARK 11, 47, 66
ASSANGE, Julian 19
AYRAULT, Jean-Marc.....13

B

BADGER, Emilie 26
Bangalore.....22
BERMES, Emmanuelle.....62, 65, 66, 82
BERNERS-LEE, Lee.....14, 21, 39, 46
Bibliothèque du Congrès 48, 51
Bibliothèque nationale de France 47, 62, 64, 66
Bibliothèque nationale de France, BnF 11, 45, 46, 47, 48, 53, 56, 62, 64, 66, 67, 68, 78, 82
Bibliothèque nationale Universitaire (BNU) 11, 43, 60, 61, 68
BOURGOGNE, Jean-Marie 29, 75
BOURLANGE, Danielle.....59
British Library.....48

C

CADA... 11, 14, 28, 31, 32, 33, 57, 59, 90, 91, 94
CADA (Loi).....32
Cantal (Archives départementales).....58
Cantine Numérique Rennaise 39, 40, 41
CC-BY-SA 11, 27
Catalogue Collectif de France (CCFr).....52
Champs-Libres (Bibliothèque) 15, 16, 37, 39, 40, 41, 42, 66, 75
Chicago 26
CHIGNARD, Simon 25, 26, 34, 39, 75, 77, 80
Comité interministériel pour la société de l'information (CISI)..... 11, 18
Clermont-Ferrand (Tribunal administratif .58
CNIL (Loi) 11, 19, 27, 30, 32, 37, 53, 82
COEPIA..... 11, 61
Collectifs .. 11, 30, 35, 39, 40, 41, 42, 68, 75, 77, 79, 80, 83
Comité des sages.....22, 49, 58, 84
Commission européenne ... 25, 49, 57, 58, 82
Conférence des Bibliothèques nationales européennes (CENL)..... 11, 49, 50
Conseil national du numérique (CNN) 11, 28, 32, 55, 57, 58, 59, 60
Crowdsourcing 23
Culture 15, 18, 27, 40, 45, 48, 49, 53, 55, 56, 57, 58, 63, 66, 78, 83, 84

D

DADVSI 11, 18, 57, 64
Danemark 49

E

Étalab ... 13, 28, 29, 30, 33, 34, 43, 47, 69, 83
États-Unis 22, 25, 26, 48, 52
Europeana 49, 50, 57, 84
Externalités (Théorie) 24, 60, 61

F

FILLON, François..... 29
FING 11, 30, 35
Flickr 43, 44
FRBR.....11, 46, 52, 66, 67

G

Gallica 45, 46, 47, 66
GALLON, Claire 30
Google Maps 26
Grande-Bretagne 21

H

Hourtin (Discours) 18, 31, 78
HTML.....11, 14, 46, 47, 82

I

IFLA 11, 43, 52, 67
Internet Archive 64
ISBD.....11, 52

J

JOSPIN, Lionel 18, 31, 78

K

KAPLAN, Daniel..... 22
KAPLAN, David..... 22, 35, 78, 89
KESSLER, Bernadette.....32, 34, 37, 63, 75
KOSCIUSKO-MORIZET, Nathalie..... 29
KUNDRA, Vivek 20, 83

L

LACROIX, Léa 38, 39, 40, 41, 42, 75, 80
LiberTIC 28, 30, 80
Licence 12, 27, 33, 34, 43, 47, 80, 97, 100
Linked data 21, 48, 49, 52, 62, 63, 64, 66, 67, 68
Logilab..... 66

M

Mandelkern 23, 78, 90, 91, 92, 93
MARC 12, 52, 66
MAUDE, Francis..... 23
Mémorandum.....20, 21

(cc) BY-NC-ND

MEPSIR (Étude) 25
MERCIER, Silvère 60, 78
MOATTI, Sandra 21
Modernisation..... 13, 22, 28
Montpellier 27, 29, 30, 31, 39, 75, 80

N

Nantes 13, 29, 30, 80
NAUDET, Séverin 13, 29
Néo-classique (Théorie).....24
New-York 26
NIGGEMANN, Elisabeth..... 49, 50

O

O'REILLY, Tim 22
OBAMA, Barack 20
OCLC 12, 51, 63, 64
Open database Licence (DDbL)..... 12, 27, 33
Open access 14, 33, 42, 60
Open Knowledge Foundation ..12, 14, 21, 27, 33, 34, 81
Open source 14, 17, 26, 35, 68, 70
OpenCat (Projet) 66, 68
Ory-Lavollé (Rapport) .. 56, 57, 58, 60, 63, 84
OWNI 13, 22, 31, 77, 78, 89

P

PAGSI 12, 18
Paris.... 23, 27, 30, 31, 40, 58, 62, 77, 78, 80, 83
Patrimoine 11, 18, 27, 34, 43, 44, 49, 53, 55, 56, 57, 58, 59, 60, 61, 64, 66, 84
PRISME (Atelier) 12, 54
PSI (Directive) 25, 58, 82, 94

R

RACINE, Bruno 45, 50
RAMEAU..... 47, 52
RDA..... 12, 67, 68
RDF 12, 14, 31, 38, 46, 47, 66, 68, 69
Regards Citoyens (Association) 21, 23
RENAUD, Clément.....35
Rennes 15, 26, 27, 30, 32, 34, 37, 38, 39, 40, 41, 42, 75, 79, 80, 97
Réunion des musées nationaux (RMN) 12, 53, 56

S

San Francisco.....7, 26
SARKOZY, Nicolas 13, 29
SCHLOSSER, Frédérique 38, 39, 75
Service public . 14, 23, 24, 43, 44, 53, 54, 56, 57, 60, 91, 93, 94, 95
SHADBOLT, Nigel 21, 83
SIGB 12, 52, 66, 67, 68
SKOS..... 12, 47, 68
SNCF.....12, 25
Socrata 25
SUDOC.....12, 51, 52, 53, 68
Suède 49

T

TED..... 39
TOULOUSE, Sarah 30, 38, 40, 43, 44, 66, 75
Transparence.. 18, 19, 20, 21, 22, 23, 24, 27, 28, 31, 35, 78, 92, 93
TRUTAT, Eugène 44

U

Universcience 54

V

VEYRON, Thierry 44

W

W3C 12, 31, 46, 64, 65, 67, 68, 83
WALLIS, Richard 63, 64
Washington 20, 26, 48
Web de données, web sémantique……46, 64, 65, 68, 82, 83
Linked data. 31, 46, 55, 62, 64, 65, 66, 67, 68, 83, 84
Wikileaks 19, 23
Wikimedia 58, 83
Wikipedia 64
WorldCat 63, 64, 68

Y

Yale (Université) 48

(cc) BY-NC-ND

(cc) BY-NC-ND