

Open data in the cultural field

Descrizione

French cultural public policies are characterized by a strong public intervention, at the State or at the local level. Their aim is to make culture and art accessible to the public. With the introduction of new technologies of information and communication, administrations in the cultural field can promote culture through an innovative process: the open data. This movement of releasing public data has been adopted by the French administration since 2010 in local collectivities. Under the Fillon's government, a law was passed on May 26th, 2011 to create an agency called Etalab to open the new French open data website data.gouv.fr, whose role is to collect and publish data of the Ministers. François Hollande and his Prime Minister Jean-Marc Ayrault with a Charte of Deontology pursue this movement.

1) Partnerships with the private sector to digitize the data to be released: issues about free accessibility of the patrimoine to the citizens

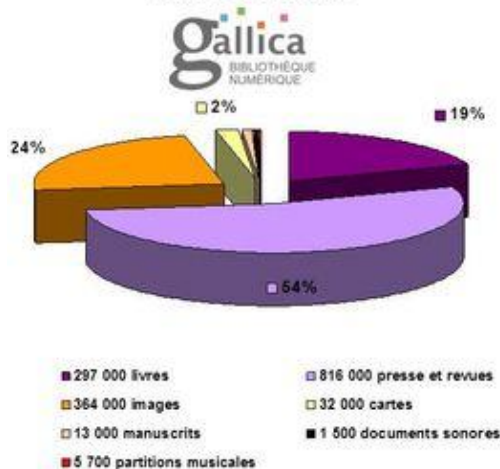
The Bibliothèque Nationale de France followed this movement of releasing public data, through an association to the private sector to digitize its data, to make them accessible to citizens, through its own portal data.bnf.fr.(1)

The BnF has the missions to collect, to conserve and to spread the document heritage of France. The BnF has to describe all the edited documents it is receiving and it has also make its heritage accessible to the public. With the Internet raising, the BnF decided to release not only the documents it receives inside its walls, but also in the Web, through data.bnf.fr, after having also created another portal, called Gallica (the numeric library, which is used to get directly to digitized documents).

So, the BnF decided to release two sets of data: digitized documents and metadata.(2)

Les données publiques culturelles disponibles à la BnF

Les documents



Les métadonnées

- Les catalogues de la BnF contiennent aujourd'hui environ 11 millions de notices bibliographiques et 5,1 millions de notices d'autorité.
- Parmi les notices d'autorités, on compte :
 - Noms de personnes : 3,7 millions
 - Noms de collectivités : 642 000
 - Noms de marques : 23 800
 - Noms géographiques : 259 000
 - Titres (tous types) : 215 000
- Pour les sujets, deux types d'autorités recensés :
 - Dewey : 138 500
 - Rameau : 168 200

(BnF

In July 2011, the BnF launched a partnership with the private sector to digitize those documents, for a total amount of 150 million euros.(3) Five centuries of literature are concerned and ancient manuscripts must be extra taken care of. Those documents will be added to the portal Gallica, which already gathers 1,5 million documents.

However, some organizations highlight the paradox of such a digitization. Whereas the purpose of such a process is to make those data accessible to French citizens, the BnF is first obliged to concede for several years the exploitation of those data to the private companies that digitized the data. The BnF has created a subsidiary "BnF-Partenariats", which will deal with the private companies chosen to digitize and which will have the exclusivity and the commercial re-use of the data they digitize.(4)

A deputy from the Parti socialiste declared that those projects « constituent une régression manifeste par rapport à la politique de diffusion du patrimoine conduite jusqu'à présent. Les modalités de ces partenariats public-privé sont par ailleurs incompatibles avec les recommandations émises par le comité des sages européens

qui, sans exclure la possibilité de concéder des exclusivités commerciales pour un temps limité aux partenaires privés, avait insisté sur la nécessité de préserver l'accès gratuit en ligne aux documents du domaine public(5)». (6)

This exploitation of the public domain, to make those data profitable to private firms, underline the difficulties to conceal first the interests between the public sector and the private sector, and the issue to make profit on public goods. However, it is also a way to make those data better known.

2) Linked open data: the challenge to open cultural data in an ecosystem to make them visible to the citizens

As seen above, the Gallica portal gathers 1,5 millions documents and the metadata represents 11 million bibliographical notices (that's to say documents owned by the BnF). All those documents are put on data.bnf.fr. The idea is to open to the outside library world the data of the BnF as the catalogs are most of the times excluded from search engines. It was rare than BnF collections were on the top of the Google web ranking when Internet users are looking for information on authors for example. The core-idea of data.bnf.fr is to create a similar website as Wikipedia about authors and works, to gather all useful information on them.(7) Internet users can download the data and re-use them.(8)

Using the semantic Web and the linked open data, the data.bnf.fr portal is covering more and more authors and works. The BnF is now inside an ecosystem as its own web pages are redirecting to other external websites, to create a cloud of cultural data.

Immerged in this environment, the BnF isn't anymore alone, but is using the network of other organizations, either public or private, to make its work known and especially to the citizens. The BnF signed a partnership with Wikipedia to correct collaboratively 1400 digitized documents that are difficult to identify.(9) This partnership ecosystem is used to widespread the BnF work and legitimacy, in order for citizens to get to know this huge library of data.

Endnotes

- (1) Bibliothèque Nationale de France, « Des fiches de référence sur les auteurs, les œuvres et les thèmes », <http://data.bnf.fr/>
- (2) LatitudeFrance, « L'open data culturel en débats », 07/11/11, <http://www.latitudefrance.org/L-open-data-culturel-en-debats.html>
- (3) Le Monde.fr, « La BNF fait appel à des partenaires privés pour numériser ses collections », 06/07/2011, http://www.lemonde.fr/technologies/article/2011/07/06/la-bnf-fait-appel-a-des-partenaires-privés-pour-numeriser-ses-collections_1545674_651865.html
- (4) Liberation.fr, « La BnF appelle le privé pour numériser », 09/07/2011, <http://www.liberation.fr/culture/01012348028-la-bnf-appelle-le-privé-pour-numeriser>
- (5) Actualitte, « Partenariat public-privé à la BnF : exploitation du domaine public », 24/10/12, <http://www.actualitte.com/bibliotheques/partenariat-public-privé-a-la-bnf-exploitation-du-domaine-public-37698.html>
- (6) TRANSLATION: “these projects are a clear regression in relation to the “dissemination of heritage” policy conducted so far. The terms of these public-private partnerships are also inconsistent with the recommendations of the comité européen des sages who, without excluding the possibility to grant exclusive commercial for a limited time private partners, insisted on the need to preserve the free online access to documents in the public domain.”
- (7) See for example, the notice about Alexandre Dumas on the data.bnf.fr website : http://data.bnf.fr/11901063/alexandre_dumas/
- (8) Data.gouv.fr, « Bibliothèque Nationale de France », <http://www.data.gouv.fr/Producteurs/Bibliotheque-nationale-de-France>
- (9) LatitudeFrance, « L'open data culturel en débats », 07/11/11, <http://www.latitudefrance.org/L-open-data-culturel-en-debats.html>

Campi meta