

OBSTACLES **FOR CULTURAL OPEN DATA**

elaborated as a result of the OpenGlam Paris workshop on April 27, 2012

Misconceptions / Misunderstandings:

- **The concept of "Open Data"**

Definition of key concepts:

The term "openness" is a bit vague, used to communicate very different things: e.g. opening data and content are two different things.

"Cultural content" must be defined: tangible or intangible cultural content, services, etc.. must be set them in the proper context: open data, but why? when? and for what purpose?

Goals:

We have to explain what is the goal of Open Data; explain the importance and usefulness of open data. Releasing data is not sufficient, it is useless if the data remains unused. We must think about how to give life to data, create added value and encourage re-use.

A. Provide examples of usage:

e.g. access to consultation; reuse; valorisation by the private sector, etc..

Some people think that this may be too prescriptive - although mentioning the uses is different from defining them in advance, it does not circumscribe them.

B. Give way to public initiatives:

e.g. put an infrastructure in place to allow the public to reuse data. The idea is that GLAM institutions do not know how to value their data as well as the public. The aim is to encourage the public to make something that the administration would not have thought of, by providing the "raw material" (i.e. raw data in a format that is the most interoperable as possible).

This is the approach that was taken by the city of Rennes.

Definition of benefits:

Defining the gains for the public, but also for GLAM institutions

Framing the debate:

People use specific terminology when referring to open data, such as 'data liberation' (literal translation from french: *libération des données*).

People at the workshop were divided on this point:

- Some thought it is too strong or violent. This term gives the wrong idea that we are fighting something that has always been "closed" - whereas the sui-generis is actually a recent right.
- Others thought it is important to use strong words in order to be heard

and better understood. It helps to stress out the importance of Open Data.

- **The eco-system where GLAMs operate**

Many GLAM institutions are not fully aware of the ecosystem in which they operate and lack to recognize the opportunity afforded by Open Data.

We must identify who the potential producers and users of cultural data are. GLAMs sometimes have a poor perception of their audience, or the use made of their data.

Open data enables GLAM institutions to be part of an ecosystem that ultimately adds value to their data. Opening up their data help them connect with the other operators of the ecosystem by creating relations and promoting interactions, e.g. by connecting with databases of many other research centers.

It is also important to define the different roles of institutions on the web and beyond.

- **Legal issues**

Lack of training, lack of knowledge concerning legal issues. We must eliminate the legal uncertainty that scares institutions away; clarify the gray areas that keep them passive. There is often no legal obstacle, but rather a mental block, "we cannot do it because we do not have the right to do it."

Legal Uncertainty:

- **Legal status of digital reproductions**

Can we create new rights when scanning content?

Risk of adding rights to the public domain e.g. by protecting the public domain with the laws of cultural data.

Question of Originality?

- **Clearing the rights**

Many rights are often not negotiated upstream, mainly the rights of:

- * state employees and officials (copyright)
- * creators of databases (sui-generis)
- * contributors to databases (copyright - eg for the description of inventory)

- **Content "protected"**

Legal risk (taken by the cultural institution) to make available protected content:

- Copyright
- sui-generis
- privacy (eg private correspondence)
- personal data
- the right of photographers
- image rights, etc..

cf. Gallica - photos that are not in the public domain, public domain books with forewords written by recent authors who are not included in the metadata.

Economic issues:

- **Funding**

Political pressure to reduce the subvention given to GLAM to encourage them to find resources on their own.

Costs of scanning works and digitizing data are high, but only for the first time.
Justification for Non Commercial restrictions: known work pays for the digitization of lesser known works.

- **Compensation**

Open data costs money, how to compensate? How to offset costs by services and the benefits obtained?

- **Business Plan**

Problem with the principle of commercial reuse. By providing content and data under an open licence that does not prohibit commercial uses, GLAMs are afraid to lose their only means of compensation (i.e. by licensing the right of using the data commercially).

Risks involved with the marketing by third parties.

Financial obstacles:

- **Costs**

Administrative Costs: sorting data (e.g. deciding which document can be disclosed and which documents cannot, etc.).

Technical costs: servers, data formatting according to standards for interoperability, standards presentation of data.

Legal Costs: clearing rights, licensing control.

- **Lack of pooled resources**

Control problems:

- **data “Dispossession”**

Some photo libraries have not released their pictures or metadata for fear of being deprived of a resource that will escape beyond their control, that will be crowd-sourced, or mashed-up.

cf. Museum of Natural Sciences of Toulouse: desire to control the reuse of data in order to avoid derogatory uses - that fear was gone after they saw why their photos are actually being used.

- **Integrity**

Fear of losing data integrity due to the reuse by third parties

Technical constraints:

- **Formats**

Which format? MARC, RDF, etc.

Which vocabulary? standard, ad-hoc?

Which serialization? json, xml, etc.

- **Process**

Many Glams are interested in Open data but do not know how to go about it.

We must define the process of production and distribution of open data. Propose a production process that encourages Open Data. Explain how to do it and how to avoid legal problems.

Build example from those who have already opened their data: eg BNF, Strasbourg.

Policy Issues:

- **Various interlocutors**

Different institutions adopt different policies according to local customs: Different policies of the different municipalities and regions. Who should I contact?

- **Lack of political willingness to encourage open data**

Basic policy decisions could resolve many problems, such as:

- A. Removing the special treatment assigned to “cultural data” under French Law - which, as opposed to other public data, does not have to be made freely accessible.
- B. Imposing a condition of interoperable and open format;
- C. Changing procurement codes to impose that by default all results are open.
- D. Changing the law so that the rights belong to institutions and not to their staff.
- E. Specifying that the faithful reproduction in 2D (scan) cannot create new rights.

- **Lobby cultural industries**

The problem is that the ecosystem also includes cultural industries whose profits are based on maximizing intellectual property rights. Economical interests can not be easily challenged.