The most known challenges of PSI Access and Re-USE: Intellectual Property (and Data Protection)

Facilitatior: Freyja van den Boom

Scribe: Lorenzo Canova

IPR is traditionally perceived as a tool for development; however they reveal to be an obstacle for an easy access and reuse of PSI. The same goes for Privacy and Personal Data protection rules. How can we overcome these challenges?

The discussion is aimed at identifying best practices on legal rules, contractual transfers of rights and other issues in the area of public sector works in different European Union countries. A particular attention is devoted to cultural institutions legal rules and practices as well. More precisely, this session focusses on three different levels of the public sector works production and circulation:

    i.   legal rules on protectability of works;

    ii.   legal rules and contractual practices on rights ownership;

    iii.   legal rules and other practices implemented within the cultural institutions.

The LAPSI 2.0 network can share its experience in identifying some good practices, based on a questionnaire that was prepared and answered by the partners of the network. In particular, the LAPSI 2.0 network identified some best legal rules on protectability of public sector information eligible for copyright protection, some good and bad legal rules on rights ownership of PSI covered by copyright protection and some best practices on the transfer of such rights. In the area of cultural institutions the LAPSI 2.0 network identified some best practices as to the metadata exploitation, access and re-use to cultural content.

SCRIBE

Intro

This is the last Lapsi 2.0 Workshop.
Suggested topics: Barriers to reuse PSI, such as Intellectual Property and Data Protection (the focus is on IP, since there is no much time). Best practices in different member states, protectability, with a specific look at cultural institutions.

Issues

- Protectability and international rules: there is no harmonization in member states (it's addressed differently in the different states)
- DB (database) protection
- Data protection in reuse of PSI: Any info about an identifiable person is considered personal data, this could be a barrier to reuse of public data. example: if you wanna use maps, that could fall in certain circumstances fall under personal data protection Directive.
  There are certain requirement to be fulfilled, there could be some problem with that.
- Copyright, 3 scenarios: PSI (1) is exempted from copyright (2) fall sunder copyright (thus you need to ask for reuse of data) (3) is a mixture of 1 and 2
  To have copyright is not bad as long as there is open access and it's made clear to re-users. Some countries have no IP protection on PSI, but specific agreements limit the access to it or its re-use
- Issue with metadata: does it fall under copyright?
- IP and digitalization: what access should the public have?

Grants in research and educational license should produce works in CC0 or CCBY
There are no licensing issues if copyright does not apply. If copyright does apply, there are no licensing issues only if there are open licenses

Round table

Ireland-Data analyst-Interface for government companies.

In Ireland, public sector does **not allow commercial usage** of PSI. It's not possible to sell services based on PSI data.
However it's easy to obtain some sort of data, e.g it's possible to go to a Public Administration to ask physically for architectural data, they will give them to you and you could use them on Openstreetmap.

The strategy often used says that commercial reuse is not allowed. There is still hope they will change to CC0 / BY. If you use PS data you always must say when it is public domain

Questions: what about db right even if is public domain data? You have IP right but who is holding the right? in Ireland the rule is basically: "Do what you want but not too much"

Finland-financial department

They are running a new OD (Open Data) program.They just released the first version of national *** (expenditures?), they used the PSI recommendation for the sector. License: they use a modified license based on CC 4.0, like CC0. The OD program will run until June 2015.
This OD program is a national implementation, they recommend to reuse of data

3 major issues:
There is no license for describing the national repository.
Hype about OD has been dying down .
Huge municipal sector, they need support for implementation

Questions: why Finland has its **own license** instead of using CC? It's a national license, actually but is just a translation of CC. There shouldn't be any difference with CC (except that it's not CC).

Latvia - National library

 They are trying to make easier for people to work with OD. There are lots of sources available (it's quite easy for people to use this data). They also added info on the okfn global ** sensor (Global Open Data Index?). Most of the key dataset in that sensor are available in some form, but there are some issues in providing usable data, i.e. lot of datasets are not machine readable (quality problems).
None of the datasets have OD licenses. However in some cases there are regulations saying that you can use the data. Basically there are confusing laws, nobody knows what is allowed and what not.
The library is publishing semantic data, from the library perspective license is a nightmare. Only if something is of public domain has open license, but the border of "public domain" is not clear. Talking about museums' data about museum there is the (fashionable) issue of thumbnail copyright. If a certain museum wants to put a catalog with **thumbnails** describing some artwork, they have to pay for each thumbnail they put, moreover even if is public domain, it must be proven.

Questions: is there an exception for low resolution images? No, in Finland and Latvia there is no exception, it is still copyright protected material.

In **Italy** there was the same problem, but an exception has been made for low resolution images and sounds. This achievement however was the result of a long discussion: the decree was available quite quickly, but it took a long time to apply it.

Quotation issue: Quotation of images is not clear, for some countries if you quote an image you are just reproducing it. In the **UK** an exception for quotation was introduced.

Possible improvements for this issue? Italian example, or exception for quotation like in the UK. In Latvia this will not be so easy since people are getting money from that (from thumbnails copyright).
Is there a clear definition of cultural institution in Latvia? Pretty much, but still, it would be hard to have an exception for cultural institutions.
However different museums have different practices: some of them publish images and they take them off only if somebody complains about it...by now lots of stuff has been published and there have not been (legal) issues.

Pictures copyright: In Italy, Spain, Austria, Germany when you take a picture you have a copyright for 70 years; when it is not creative you have a neighbouring right lasting 20-25 years.

Provenance of originality: who guarantees the originality of data, with gov data you can do something. there are certain responsibility for org that create that, you are accountable for rightness of data etc.
In Finland there is no organization that to proves data originality. A curious case in Finland is the one about data on rare protected species. Data is protected, therefore government data cannot publish it. However communities (interested in rare species) have more accurate and recent data and publish them. So it is quite nonsense what is happening.

Asun - Spain - **licensing and rdf**: rdf allows provenance information and also license description. Data should not only be for humans but also from machines. Machine can find incompatibilities between licenses (if data is exposed in rdf). In the open data cloud 50% of resources are claimed to be open but don't specify license, there needs to be a standard of representation. One of the problem is communication between technical and law people.
On datahub there is already a dataset about all kind of licenses.
We could also include economist in the discussion...its hard to make each other understand.
From a technical point of view it's hard to express when you don't know sure stuff, like what kind of license is used or it's supposed to be used.

Economist - part of the Tourism Agency of Croatia

Government does his job publishing lots of dataset in the transparency area, but there should be some boundaries for users' request to open new data. Some people ask too many times for (granular) data, government people bothered by **too many requests** (from people that want to analyze too much in depth).

It must be considered that the Government doesn't have all the requested dataset. Many (high value) dataset are given to **other institutions** that don't receive any money from the State. They sell those kind of data, e.g. cadastre, data on highways. These datasets are really interesting for public, but there is no way they are going to be opened since is a big part of the business of some public company.

It happens the same in **UK**, some **public companies** make money out of valuable datasets, those are high value, they permit to have a good source of revenue and nobody will make them open.

It must be said that public companies often are not covered by PSI legislation.

In **Latvia** public companies are covered by PSI legislation, they started to publish in order to be compliant with PSI standards. The problem here is the outsourcing of the data, those private companies that run public DBs have no incentive in free the info.

In LAPSI 1 in the final paper there was the **request to please include public company** in the PSI legislation, however that was not done even if they frequently have the most valuable datasets.

In the end Public Sector Bodies cannot refuse to publish data, instead public companies can (they are not subject to article 8 and 10).

Example from **Croatia** - tourism public company has budget from taxation to tourists and from data on tourism. They were asked to open it, because company from other states could use tourism data to create new activities, the answer was simply "No".

Licences and standards

From the technical point of view, what is the minimal set of info for licenses to define constraints? metadata, registry?

It's national, every nation has its licenses and usually they are not machine readable.

In Finland dcat has been adopted. Moreover they are using the ckan platform, Actually it is an adapted ckan basic. There is a need of coordination for organizing OD platforms, in this way municipalities will not have all their own portal, creating the usual interoperability issues. It is also important to provide mandatory fields (?) (in Finland it has already been done).

What about mandatory fields for licenses? It would be good if it was the same as in dcat. Libraries use marc21.

Technical opinion about licenses. There are too many standard for labels and metadata for expressing the licenses.

It is possible to infer that two licenses are the same if they have the same uri, however this not enough. Specifically dc and dcterm are not enough for licenses, it's a matter of granularity of info about licenses. In Spain, in fact, they are using as standard ODRL.

It would be beneficial trying to define standard for representation of licensing like "machine readable rights statement vocabulary" (ODRS).

Malta - PSI

It's important to clean up the legislative tools (e.g: foia), copyrights can be great barriers.
It could be an idea to fuse new psi provision and national law in one piece of legislation.
Regarding the "request for access" there are lots of different mechanism, however psi open by default seems to be the best solution.

IP and databases

If you are a IP right owner you can impose the price you want. In Italy, IP right is a fundamental right, but there is no mention how to protect it and public institution could use this argument for limiting access to information.

IP on DB: In Malta if you create DB and you are a public body you have no right on it, it's public domain, but it doesn't work like this in every country. At the European level is not easy to find a solution. Malta approach is best practice. In Italy art.29 states that for 20 year what a public employee produces is property of the country for 20yrs and then the rights get back to the employee.
Interpretation of contracts is another thing to clarify…

WHAT SHALL BE DONE TO ENCOURAGE REUSE OF PSI?
Allow re-use of relevant information even if it is in the hands of private companies (for example through digitization or outsourcing or privatization of governmental tasks in the past; royal mail example)
WHY?
It makes no sense that this valuable public data is in the hands of private companies and cannot be re-used freely.
HOW
Perhaps by not focusing on who owns the data but what data should be made available in the public interest regardless of the ownership.

A second point for discussion was the linking of all the opendata platform initiatives