

A contribution from IAML to the EU reflection about a Digital Single Market for Creative Content Online

1 Why digitised musical objects are so difficult to access

Music is a perfect medium to be transmitted globally over the Internet, because its enjoyment and understanding is not dependent on a verbal language, while in order to enjoy a digitised book the knowledge of its language is essential. This is one of the reasons why many libraries already run extensive projects of digitisation of their musical objects. Digitisation of invaluable and unique treasures offers remote access to the widest range of interested people – from members of the general public to learned scholars - all over the world.

However, if music is a language understood internationally, sadly many of the keys to access are language-dependent.

So while the issue of digital rights barriers is very clearly and well summarized in the EU paper http://ec.europa.eu/avpolicy/docs/other_actions/col_2009/reflection_paper.pdf, when it comes to **access** to digitised musical objects two major issues are never sufficiently taken into account: (i) language barriers; (ii) inconsistency/lack of homogeneity/interoperability of data.

1.1 Language barriers

The Eu has 23 official languages, while the great majority of what is available through the web is described in English.

Musical nomenclature differs from language to language – particularly for classical music - and, for example, German retailers usually complain that German music lovers expects to search and find musical works with a German nomenclature.

Almost half of the Europeana content is described in French, because France is the main contributor to Europeana. It is quite unlikely that a big number of – say – Hungarians, Latvians, Estonians or Slovenians are able to run a proper and effective search in French. Therefore, a big part of the effort of Europeana is not appropriately exploited.

But language barriers are not only a barrier for those who don't know the language of description. They are also an inherent barrier for search engines.

In a normal search engine, that basically runs a textual search, the retrieved result is dependent on the language of the search: a search run in German will retrieve documents described in German etc.

But when it comes to music, its content has a powerful *wordless meaning*, and only its description has a verbal element. If anybody *listens* to the first notes of the EU anthem (i.e.: the final part of Beethoven *Ninth Symphony*) he immediately recognizes *which* music is this. But if he tries to describe it with words, their formulation may be *Ode to the Joy, Ode an die Freunde, Ode à la joie, Oda a la Alegría, Ode aan de vreugde*, etc. etc. etc

What we need, then, is **a concordance which is able to identify music content despite the fact that it may be described in different languages, or even different ways in the same language**

Moreover, there is for sure a very interesting market for digitised content in small countries of Europe like Hungary, Latvia, Estonia or Slovenia mentioned above. The scarcity of traditional retailers is bigger there than – say – in Great Britain or Germany or France, where more widespread languages are spoken. Therefore, in those countries a digital market has more potential than where traditional alternatives are available.

But this would be true only if they could be taken into account as a *single market*: commercial producers like publishers and the recording industry or music retailers like iTunes or Amazon cannot afford the big costs linked to localising music descriptions for such tiny countries with different languages. Here again, a potentially big market – when aggregated – is lost just because of access barriers.

And this is only to stay focused on the EU market, but of course European producers, that are still the great majority - US apart - miss biggest markets like China, India and the rest of the Far-East due to language problems.

1.2 Inconsistency of data

Lack of homogeneity and inconsistency of data not only seriously harms any effort to clearly detect the appropriate rights holders (see http://ec.europa.eu/competition/consultations/2009_online_commerce/afi.pdf: “*Lack of data and complete information on right holders represents another obstacle to accurate remuneration and correct management of rights*”), but most of all it is one of the main obstacles for exhaustive, thoroughful search.

Each content producer (i.e. record companies, music publishers, radio stations, concert societies...) follows its own rules, because it has no authority on other producers. For each of them, the creation / correction of metadata is a major part of the effort alongside the digitisation of audio (and video) content, and therefore it represents also a barrier due to its high costs in terms of manpower.

Both libraries and retailer sites like Amazon or i.Tunes are *dependent* on content producers, and therefore obliged to cope with this huge quantity of different descriptions of the same work.

None of them can afford the immense task of really uniform all of these titles, only to find them inconsistent again the next day.

Hence, users are terribly frustrated when they cannot effectively retrieve all the different versions of a work to choose their favourite one to buy it/listen/download/read: therefore, a very promising market, or the true fruition of cultural objects go at a slower pace in respect with its true potentialities.

A contribution from IAML to the EU reflection about a Digital Single Market for Creative Content Online

2 Why music librarians could help easing the problem

2.1 An EU music portal

Producers present at the EU round-table on music of October 20th (<http://europa.eu/rapid/pressReleasesAction.do?reference=IP/09/1548&format=HTML&aged=0&language=EN&guiLanguage=en>) already "agreed to set up a common, non-mandatory and non-exclusive portal to include the largest possible repertoire," said Bernard Miyet, head of French collecting society Sacem.

But a portal that includes *the largest possible repertoire* as it is described now, far from easing any of these problems, rather compounds them. In the future, after so many years of contributing digitised data to ever larger datasets, no such contributor will be able or be prepared to change/correct/adapt their data just for the sake of being consistent with other players. Neither, really, can they be expected to take on such a universal obligation. It would be too big and too expensive an effort just to conceive it. Aggregation of data from many sources is a wonderful goal and achievement, but at the same time simple aggregation of its descriptions (metadata) without control is a major problem that prevents optimal exploitation of the content and any **correct attribution of 'intellectual property rights'** to rights owners: easing access to resources and unequivocally identifying them is the best way to improve it.

This 'EU music infrastructure' would also represent an interesting scientific innovation towards a true *semantic web*, because music, as it is managed within the network of music libraries, is likely to be one of the few cases where an ontology already exists, it is accepted and shared by a large body of users and it is fairly easy to formalize

2.2 The role of librarians for a EU music infrastructure

Music librarians are accustomed to cope with both problems, and therefore they are in a very good position to help making the portal – or any other solution – into an effective tool.

Libraries usually receive/host material described in every language on earth, and when cataloguing they cite the title in **the language of the item** (e.g.: Italian for a Ricordi score, French for a Harmonia Mundi CD, English for a Decca digital file), but connect it to the appropriate uniform title which is usually in **the local language** (i.e.: the language of the country where the library is located), and has mainly the scope to aggregate together different descriptions of the same work, or different namings of the same person (author/composer/performer).

This approach illustrates one of the many **simple, cost-effective, ways in which libraries have had to develop concordances of data across languages to allow accurate search and retrieval of highly complex data. This is ever more important as increasingly** libraries form more interconnected systems of cataloguing, retrieval and even portals themselves.

This approach and this habit is a powerful 'drive belt' to really aggregate disparate contents and increase the **interoperability among different players' metadata**.

There is a wonderful opportunity here for an innovative, close, **cooperation between public bodies like libraries** - traditionally hosts, keepers and disseminators of cultural heritage - and **private, independent bodies like content producers and right owners**. Libraries are also well placed to provide very good feedback from end-users, with whom they have a strong, daily contact. Libraries are service- and user-driven.

The **desired result** of this EU music infrastructure is an increased use of all kinds of musical objects, truly **integrating** what is publicly available in **e.g. libraries with music content from the commercial sector, such as** online music stores or record labels, music publishers, broadcasters, concert societies, opera houses, websites.