

Why Metadata Matter for the Future of Copyright

Kluwer Copyright Blog
November 27, 2020

Martin Schaefer (Boehmert & Boehmert)

Please refer to this post as: Martin Schaefer, 'Why Metadata Matter for the Future of Copyright', Kluwer Copyright Blog, November 27, 2020, <http://copyrightblog.kluweriplaw.com/2020/11/27/why-metadata-matter-for-the-future-of-copyright/>

The legislative agenda of the past two decades – both in Europe and further afield – has been about adapting copyright to the requirements of the information society. The administrative means to make use of those new opportunities by licensing at the right source and allocating revenues to the right recipients, in a world of interactive and intertwined content, have not been harmonised at the same pace. In the copyright industries of the 21st century, metadata are the grease required to make the engine of copyright run smoothly and powerfully for the benefit of creators, copyright industries and users alike.

What is the problem?

The challenge is greater than it appears at first sight. Imagine a user who wants to do something as simple as using a photo and a piece of music for his or her website: How would this user find out from whom to ask permission? Is it a Collective Management Organisation (CMO) or an individual rights holder? If so, which CMO and which rights holder? And if you find the right person, will there be a way to get licences for exactly the use required? For a single pop song, you will need data about the composer, the corresponding music publisher and often a CMO. The same applies for the lyricist. In addition, there is the performer, the label and maybe yet another CMO representing his or her rights or claims for remuneration. Often rights ownership will differ from country to country. To make things yet more complicated, even for a short pop song there will frequently be not one but numerous co-composers and co-lyricists, each of whom might be represented by different music publishers.

Why are centralised database projects likely to fail?

CMOs and other groups of rights holders have tried in vain to create large-scale “definitive” databases designed to establish standard points of reference. Hundreds of millions of dollars and euros have been sunk into such projects. None of them so far have worked sufficiently well. There is a simple reason for that: Metadata are not stable, as rights in content are a tradeable asset. Accordingly, ownership in content changes frequently (e.g. if publisher A sells part of its catalogue to publisher B), just as changing user demand is likely to require new sets of data. For example, while the new owner of a right is usually keen to inform CMOs and other operators of central databases as quickly as possible after rights acquisition, the former owner does not necessarily feel the same pressure to act – with the result that two different owners will show up for the same content.

The analysis

If centralised databases are not the answer, what else could solve the problem? Even though this issue is common throughout all copyright industries, this post will use examples from the music industry. This field has shown to be the most complex of all, so it might be justified to say that “if we can solve it there, we’ll solve it anywhere.”

There are plenty of proprietary databases everywhere in the music business. Each CMO owns one and each label and music publisher does as well. There are also public institutions running non-profit databases of metadata that are often of superior quality. However, almost none of the database owners are likely to allow searches of their databases, let alone extraction of their data, even in part.

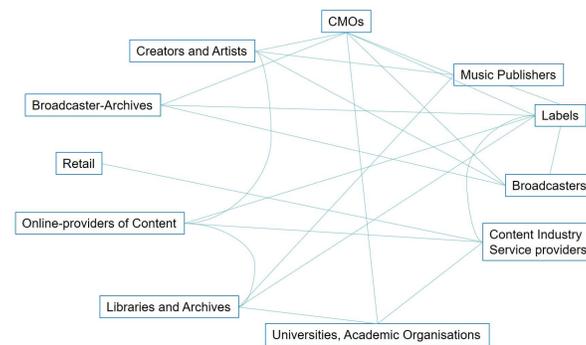


Fig. 1: Numerous players own databases and exchange metadata for their business relations but do not know how reliable those metadata are.

A possible solution

If the central “super database” cannot solve the problem, and direct exchange of larger sets of metadata among the holders of proprietary databases does not seem feasible, how can we emerge from the deadlock? A possible solution could lie in a metadata search and enhancement tool that could constitute a buffer that safeguards the interests of the various proprietary database owners. They want to keep out intruders who are simply interested in freeloading from their stock of information. That is what the new approach can achieve. Instead of allowing everyone to look into everyone else’s metadata stocks (and to blatantly copy from them), it would be a central trustworthy system operating a new search tool that conducts searches in all databases connected to it, ranking the data based on a highly sophisticated yet transparent algorithm.

To use the above example of a catalogue of works that is sold from one music publisher to another, the algorithm would suggest that the newer owner is more likely to be the actual owner than the earlier one. Access to this network would be granted to anyone who is willing to contribute with their respective databases. Only those who are willing to give access to their data should be allowed to benefit from the “machine wisdom” of the search tool in combination with the data available in that closed network, based on a trusted operator.

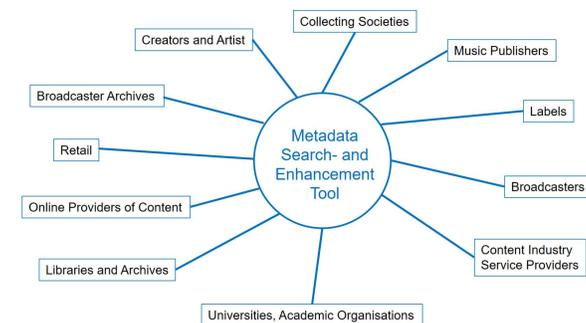


Fig. 2: If all players connect to a trusted service, they can benefit from the data of all databases connected to establish reliable metadata for their respective business relations.

In late summer 2020, the German Ministry of Justice and Consumer Protection commissioned a study for a model based on this approach (“Conception of a European, decentralised Copyright Platform for Music”) from a working group presided over by Prof. Dr. Gronau of Potsdam University. This was presented at a workshop called “Copyright Infrastructure”, at a virtual conference hosted by the Ministry as an [online event in September 2020](#).

The Study

Based on a thorough analysis, the study arrived at a number of prerequisites needed to make the concept a success:

• Neutrality

As an independent entity, the platform should stay neutral. In order to attract the broadest possible circle of participants, it must avoid even the impression that it could be aimed at levelling or re-defining the market status of any stakeholders that take part in it. Most stakeholders required to make the system work have a commercial agenda, and the platform should neither aim at nor effectively act in a way that levels the (relative) market power between different stakeholders.

• Respecting the decentralised structure and the independence of all databases connected

The platform must always be a tool for the participants’ business and should enhance the scope and metadata quality of their databases. If the platform’s operator were to ever use its central position in the metadata network to store data centrally, it would become a threat to them and thus be abandoned.

• No monetary transactions – payment by mutual provision of data

The currency in this system is metadata. If metadata were bought with money, users would no longer have the incentive to provide their metadata in order to improve the overall quality of the data. Moreover, other users with less money as a resource would be disadvantaged. The platform should be a neutral space for all users. Users should have a mutual incentive to exchange metadata with one another.

• Precautions against creating a data pool outside the system

Technological and contractual measures need to be taken against the risk that powerful players could try to drain the connected databases in a disproportionate manner to create a large-scale database outside the network.

• Transparent algorithm and non-commercial setup of the platform

Only a non-commercial setup of the platform and full transparency of the algorithm will guarantee the degree of trust needed for owners of large stocks of metadata in proprietary databases to join the network.

Effectively, a system established along these lines would allow machine learning and eventually improve the quality of metadata throughout all databases connected to it.

The role of the EU

Since the Romanian EU Council Presidency of the first half of 2019, issues like those discussed above have

been dealt with on an EU level in more detail than before. However, in the [Conclusions of the Romanian presidency](#), reference is still be made to a "database at EU level" that could provide for the reliability and sustainability of the data collected.

Over the course of the Finnish EU Council presidency in the second half of 2019 and the Croatian Presidency of the first half of 2020, the issue of data economy has gained a high rank on the [priority list of the EU](#). This led to the study, which was commissioned in preparation for the German Council Presidency that is due to end this year.

Conclusion

Only recently, the EU Commission's Directorate-General for Education, Youth, Sport and Culture issued a ["Feasibility study for the establishment of a European Music Observatory"](#) that might be open to accommodating a test run of a "Metadata Search and Enhancement Tool".

However, the scope of the idea goes far beyond the music industry alone and can be applied throughout the copyright sector (and even beyond). In light of all that has been said, it seems obvious that large international public organisations or NGOs would be ideally suited to set up a project like that outlined above.