Conclusions and recommendations
Guibault, L.; Margoni, T.; Spindler, G.

Published in:
Safe to be open: study on the protection of research data and recommendations for access and usage

Citation for published version (APA):
5. Conclusions and Recommendations

In the following, we will briefly describe the conclusions of our study on the European legal framework and thereafter give some recommendations on how the rights situation as to research data can be improved. The focus will be on the interests of the scientific community as well as on the aspect of legal security. Some recommendations will be given to the European legislator on how to improve the European regulations. In addition, specific recommendations on how to enable the use and reuse of research data between repositories/data providers and e-infrastructure providers on a contractual basis will be given.

5.1 Conclusions on the legal framework

Taking into account the legal framework referred to above, whether on the level of Acts and directives or on the contractual level of licences, some major pitfalls are obvious which can be structured along the lines of the intended mutual recognition and use of research data and databases. Before analysing these pitfalls and giving some recommendations we highlight briefly the results of the legal analysis:

a) Research data itself is not protected by copyright law and seldom by other legal norms. Only databases and their structures are protected (if sufficient investments have been undertaken for establishing the database).

b) (Massive) extraction of data for the purpose of analysis is not being covered by specific limitations and exceptions of the Database Directive such as the limitations for temporary copies enshrined in Article 5(1) Info Directive.

c) Scientific use of databases is just an optional limitation (Article 9 Database Directive) and is not fully harmonised. Hence, the scope of this limitation is quite opaque across the different Member States.

d) Scientific use does not cover the electronic infrastructure as such. Hence, the entire reproduction of a database cannot be justified by scientific use. The envisaged reproduction of other databases in OpenAIREplus and vice versa is thus not being covered by the limitations.

e) Specifically, only direct scientific use is being covered by the limitations of the Database Directive. Hence, indirect scientific use, such as scientific databases using another database, cannot benefit from the limitations.
Conclusions and recommendations

Moreover, linking to research data and/or publications is not clearly regulated either in the Info Directive or in the Database Directive. Obviously, some Member States have the tendency to qualify linking as another way of making works publicly available, thus burdening the link setter with the obligation of obtaining specific licences/agreements of the author. However, an ECJ decision could clarify these issues in the near future.

On the level of licences some deficits have also appeared:

a) As demonstrated, the Creative Commons License version 3.0 explicitly does not cover database rights as they are reserved in the case of the unported version and waived in the case of the EU ported versions.

b) Other licences such as the ODC Licences suffer from a lack of interoperability with machines and do not cover specific scientific purposes or they do not refer to copyrights as well as to database rights so that different licences have to be used.

c) Even carefully crafted licences such as the German DPPL do not deal with database rights explicitly. Moreover, these national licences are scarcely known outside the relevant Member State, here Germany.

Given these deficits a set of actions may be considered that could be structured roughly according to the different roles/levels of actors: be it legislation or be it contractual.

5.2 Recommendations to the European legislator

Legislative actions should be considered only on the European level due to the fact that the relevant legal acts are harmonised by European directives. Thus, there is scarcely any leeway for national legislators to introduce new limitations; however, new protection rights may be created, such as the recent new right for press publishers in Germany\textsuperscript{396}.

This fact points to the first question to be answered, namely whether there is a need for protecting research data, e.g. by introducing a new sui generis right. Starting with the idea that research data should be freely (re-)usable by the scientific community, such a new right would be highly counter-productive to the envisaged goal as it allows other (scientific) users to be blocked from analysing and using the data. Such a protection right to data would encompass any kind of data that could be created, such as meteorological data, health data, traffic data etc. A definition

of research data is hard to find. Thus, the sensitive balance between freedom of access to ideas and to data etc. on one side and protection for works on the other side would be heavily affected. Hence, any kind of protection of data should be restricted to privacy law and contractual provisions (such as know-how protections), if any.

More relevant to the OpenAIREplus concept would be the introduction of new limitations for scientific use of databases: starting with the limitation for scientific purposes enshrined in Article 9(a) Database Directive, the scope of privileged uses should be enlarged, explicitly encompassing the analysis, the re-utilisation and the whole reproduction of a database. However, only use for scientific purposes as well as intermediary scientific purposes such as non-commercial scientific databases should be privileged. By enlarging the scope of privileged users to infrastructure operators such as non-commercial scientific databases, the barriers to exchange data and also metadata could be overcome. Moreover, the right to make the data (and database) available to the public on a non-commercial-use basis should be integrated as well, as it is not actually subject to the limitations for scientific purposes.

At least the limitations of Article 9(a) of the Database Directive should be made mandatory and be crafted in such a way that the limitation is a full harmonisation.

5.3 Recommendations to data- and e-infrastructure providers

However, these legal options may come too late to enhance research in the EU as it has proved to be very difficult from a political angle to reform EU directives in IP rights. Hence, in order to foster the OpenAIREplus e-infrastructure it is highly recommended that a contractually based framework for open exchange of data and databases be established. As these networks encompass potentially vast number of participants it could be based only upon the type of Open Source/Open Access/Creative Commons licences, which require the user to use the same type of licences in order to benefit from the rights transferred, thus ensuring the spreading of the licence and the boosting of exchange on the same grounds.

From this starting point the choice is easily made: given the pitfalls of the licences described, such as ODC etc., and taking into account the recent reforms of the Creative Commons Licence it is clear that the new CC License version 4.0 is the most suited to the purposes of the OpenAIREplus project. As the new CC Licence now encompasses database rights as well as copyrights it guarantees the

397 Cf. Chapter 1 above.
398 Thus, it is no wonder that some legal scholars even qualified Open Source Systems as a sort of partnership, constituted by a subsequent agreement to the same licence.
free exchange of research data, albeit on condition that users follow the same approach. Thus, it is ensured that a commercial exploitation of non-commercial databases is excluded and a free flow of information is guaranteed.

During the development of the OpenAIREplus project, thanks to the collaboration between the different teams involved, and especially the scientific teams collaborating on Working Package 3 and the legal team, a list of databases has been gathered that will be used for the OpenAIREplus project\textsuperscript{399}. The list is not complete and never will be, given the ambition of scalability of OpenAIREplus.

However, it proves to be a very useful element for an exercise of analysis and understanding of how to implement legal research at the level of the databases used by OpenAIREplus.

Of the 19 listed databases, none currently fulfils the conditions of Open Access\textsuperscript{400}. A structural absence of legal expertise in the specific field of regulation of the copyright and related rights aspects deriving from the use of the databases of the participating partners can generally be observed. The lack of expertise and of specific legal competences is particularly serious, especially in light of the fact that in a few cases a sensibility towards Open Access goals is clearly present and can be identified in the terms of use with expressions such as: “the data are freely available”, “the system is operated in the sense of the Berlin Declaration on Open Access”, or even references to a Creative Commons licence, unfortunately without an indication of the correct licence, for the reference is to the 3.0 unported version, which is, as we have seen above, a version that does not license the \emph{sui generis} right on databases.

The consequences of this lack of clear Open Access conditions are serious. A lack of Open Access in this context means that no reuse, mining or other forms of analysis of contents of databases are in accordance with the terms of use of the databases used. Therefore, the OpenAIREplus infrastructure is not effectively authorised to use the partners’ databases as intended.

The fact that often at least a general indication of the “Open Access ideals” behind the database is given does not help much. To state that a database is available in “the sense of the Open Access” movement offers an idea of what the purpose of the developers is; it could give guidance to a court in the case of litigation, but not the legal certainty necessary in these cases to avoid possible liability. The use of specific references, such as to a Creative Commons License, is once again not sufficient. On the contrary, as we have seen, the reference is often made to the wrong licence, in a way that will offer no possibility of different interpretation. The databases are often made available under a licence that reserves the SGDR to its maker; therefore any unauthorised use – such as acts of data mining – is in violation of the terms of use and will give rise to liability. The common-sense

\textsuperscript{399} Thanks to Jochen Schirrwagen, Maarten Hoogerwerf and Johanna McEntyre for facilitating such list.

\textsuperscript{400} For more details about Open Access see above Chapter 4.1.
argument that, within the OpenAIREplus project, the owner of the database will never pursue such a course is a weak one, since it is precisely when rights are allocated with uncertainty, when the owners of such rights change, or when the project's dimensions or initial conditions are subject to significant changes, that the importance of having drafted the proper contractual agreement and allocated property rights with precision becomes fundamental.

The situation is serious but not irremediable. It is advisable to urgently change the terms of use of the databases of OpenAIREplus and its partners to include a reference to the correct licences. We have analysed some of the best examples of Open Access licences and we have also stated our preference, for the reasons explained above, for the upcoming version 4.0 of the Creative Commons licences. Until the final version 4.0 is available, version 3.0 ported to EU legislation dealing with the SGDR is an acceptable solution. In such latter case, however, one has to bear in mind that the SGDR is waived, therefore no conditions, attribution, Share Alike, or any other condition will be applicable to the reuse of the database. Such reuse will, nonetheless, be legitimate.

Additionally, it is fundamental that the databases used by OpenAIREplus are made available under such licences in their entirety, therefore not only the data but the databases themselves. Only in these circumstances will activities such as data mining of the entire databases and reproduction of their contents be in accordance with the licences employed.

One of the objectives of OpenAIREplus is to become a reference model and infrastructure for the European scientific community and for European society in general, and in view of this, OpenAIREplus cannot be limited to a predetermined number of repositories. However, the inclusion of any new repository will have to be undertaken following the instructions given above, and we would strongly recommend that dedicated personnel with specific legal expertise in the field of copyright and related rights and in Open Access models be employed for this purpose. Only a case-by-case analysis of the newly added repositories can guarantee the avoidance of legal risks in including the wrong repositories. With the back-up of this study, this is, however, a task that would be both feasible and gratifying.