Užití veřejných licencí ke zpřístupnění databáze metadat

Reinöhl, Michal
2014

7th Conference on Grey Literature and Repositories: proceedings [online]. Prague: National Library of Technology, 2014 [cit. 2014-12-15]. Available from: http://nusl.techlib.cz/index.php/Sborniky. ISSN 2336-5021.

Using public licenses for publishing a database of metadata

Michal Reinöhl

michal@reinohl.cz

Masaryk University, Faculty of Law, Czech Republic

Abstract

The paper focuses its attention on the legal and practical possibilities when utilizing public licenses as a tool in making metadata available to the wider public in a form that is as open as possible, as well as on the obstacles that need to be overcome in doing so, in particular with regard to the rights of third parties. The paper will draw on the experience gained from publishing the database of NRGL metadata and will present generally applicable conclusions based on this.

Keywords

Public Licenses, Metadata, Legal Aspects, Databases, Open Data

7th Conference on Grey Literature and Repositories: proceedings [online]. Prague: National Library of Technology, 2014 [cit. 2014-12-15]. Available from: http://nusl.techlib.cz/index.php/Sborniky. ISSN 2336-5021.

Introduction

Public licenses are a wonderful tool, which can help spread and share data freely or with minimal restrictions. However, before publishing a database of metadata under the terms of a public license, one should consider a number of issues, some of which may not be immediately obvious, while others may seem discouraging at first glance. These issues deserve to be given some attention and that is the purpose of this paper.

If we intend to publish such a database, we should first consider which rights we hold in relation to the database and how we can dispose with these rights. Then, we should consider which other parties might also hold rights relating to our database or its contents, and the parties on whose rights we might infringe by our actions. Finally, we should select the public license that best fits our needs.

Legal protection of databases

In Czech law, a database is defined in the Copyright Act (Act no. 121/2000 Coll., on Copyright and Rights Related to Copyright[1]) as "*a collection of independent works, data, or other items arranged in a systematic or methodical manner and individually accessible by electronic or other means, irrespective of the form of the expression thereof.*"2

Such a database can primarily be protected against unlawful use by means of copyright and the Sui Generis Database Right (SGDR). Both of these protection regimes are set out in the Copyright Act.

A database can also be protected as a trade secret, through unfair competition legislation[3] or by means of contractual conditions[4] etc. However, these forms of protection are not so relevant to the object of this paper.

Copyright

Based on the Berne Convention for the Protection of Literary and Artistic Works and the Database Directive (Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases), the Copyright Act provides copyright protection to any database, "*which by the way of the selection or arrangement of its content is the author's own intellectual creation, and in which the individual parts are arranged in a systematic or methodical way and are individually accessible by electronic or other means.*"[5]

---

[1]An English translation of an outdated version of the Act can be found on the website of the Ministry of Culture: http://www.mkcr.cz/assets/autorske-pravo/12-AZ_2006_v_AJ.pdf

[2]Art. 88 of the Copyright Act

[3]CONNELLY KOHUTOVÁ, Radka. Databáze ve věku informační společnosti a jejich právní ochrana. Vyd. 1. V Praze: C.H. Beck, 2013, s. 129. ISBN 978-80-7400-493-3.

[4]Although some contractual restrictions on the use of a database might not be permitted under EU law, even in relation to databases which are not protected by copyright or the SGDR – the CJEU is to decide on this issue in case C-30/14

CJEU case on 'screen-scraping' has potential to affect business models, says expert. In: *Out-Law.com* [online]. 2014 [cit. 2014-11-20]. Available from: http://www.out-law.com/en/articles/2014/november/cjeu-case-on-screen-scraping-has-potential-to-affect-business-models-says-expert/

[5]Art. 2(2) of the Copyright Act; see also Art. 2(5) of the Berne Convention for the Protection of Literary and Artistic Works and Art. 3(1) of the Database Directive

This means that in order to enjoy copyright protection, databases only need to meet a lower standard of originality than other works, which need to constitute "*a unique outcome of the creative activity of the author.*"[6]

However, some amount of originality and creative work is required even in databases. A database of metadata will usually be very much determined by technical considerations, rules and constraints, and therefore will not be protected by copyright.[7]

As far as the contents are concerned, metadata mostly only involve just facts as such, which are not protected by copyright.[8] There may, however, be some copyright-protected works contained among the data, such as abstracts in bibliographical data. The consequences of this will be discussed later on.

Sui Generis Database Right

The Sui Generis Database Right (SGDR) is a separate protection regime introduced by the Database Directive. According to the Copyright Act, the SGDR *"pertain to the maker of the database, provided that the formation, verification or presentation of the content of the database represents an investment, which is substantial in terms of quality or quantity, irrespective of whether the database or the contents thereof are subject to copyright protection or any other type of protection."*[9] As the SGDR is independent of copyright, both rights can apply to the same database simultaneously and each can even be held by a different person.[10]

The law requires the investment to be qualitatively and/or quantitatively substantial – that means that the investment need not be quantifiable (like e.g. a monetary investment) and as such a substantial investment of effort will be sufficient to warrant protection.[11] [12] This protection is therefore independent of creativity or originality, but it protects the (mostly

[6]Art. 2(1) of the Copyright Act

[7]Judgment of the Court (Third Chamber) of 1 March 2012, Football Dataco Ltd and Others v Yahoo! UK Ltd and Others. Case C-604/10 paras 37 – 39: "*the notion of the author's own intellectual creation refers to the criterion of originality (...) that criterion of originality is satisfied when, through the selection or arrangement of the data which it contains, its author expresses his creative ability in an original manner by making free and creative choices (...) and thus stamps his 'personal touch' (...) By contrast, that criterion is not satisfied when the setting up of the database is dictated by technical considerations, rules or constraints which leave no room for creative freedom.*"

[8]Art. 2(6) of the Copyright Act

[9]Art. 88a(1) of the Copyright Act

[10]This will mostly occur because copyright is held by the author who has to be a natural person, as opposed to the SGDR, which is held by the maker which can be a legal entity

[11]Recital 40 of the Database Directive

[12]Caution! An investment in the creation of the content is not relevant for SGDR protection – the aforementioned *obtaining* must be interpreted as the finding, acquisition and collection of existing items:

Judgment of the Court of European Union of 9 November 2004, Fixtures Marketing Ltd v Oy Veikkaus Ab. Case C−46/02.: "*The expression 'investment in ... the obtaining ... of the contents' of a database in Article 7(1) of Directive 96/9 on the legal protection of databases must be understood to refer to investment in the creation of that database. It thus refers to the resources used to seek out existing materials and collect them in the database but does not cover the resources used for the creation of materials which make up the contents of a database.*"

economic) investment in its creation.[13] These characteristics make SGDR protection somewhat similar to the Anglo-American *sweat of the brow* doctrine[14] (which is no longer used).[15]

If a database is protected by SGDR, then "*the maker of the database shall have the right to extraction*[16] *or re-utilisation*[17] *of the entire content of the database or of its part substantial in terms of quality or quantity, and the right to grant to another person the authorisation to execute such a right.*"[18]

In principle, the copying (extraction) and/or making public (re-utilization) of a substantial part of a database protected by the SGDR is only allowed with the permission of the maker or when a gratuitous compulsory license applies, such as personal use or use for (non-gainful) scientific or educational purposes.[19] Note that the "*Repeated and systematic extraction or re-utilization of insubstantial parts of the content of the database*"[20] is also disallowed. It is important to understand, however, that the SGDR does not protect the individual contents themselves, but rather the database as a whole.

Because the SGDR was a completely new concept when it was introduced by the Database Directive, a lot of uncertainty arose around it. The ECJ has played an instrumental role in shaping and clarifying the scope and boundaries of the SGDR. Its most important decisions primarily concerned the term *database* and the s*ubstantial investment* required to obtain SGDR protection,[21] and the meanings of the terms *extraction*[22] and *re-utilization*.[23]

[13]Recitals 39 – 40 of the Database Directive

[14]LEŠČINSKÝ, Jan. Právní ochrana databází. Vyd. 1. Praha: C.H. Beck, 2003, s. 33. ISBN 8071798339.

[15]The doctrine was definitively rejected in the US by the Supreme Court in *Feist Publications, Inc., v. Rural Telephone Service Co.*, 499 U.S. 340 (1991)

[16]Art. 90(2) of the Copyright Act defines *extraction* as "*the permanent or temporary transfer of all the content of a database or a substantial part thereof to another medium by whatever means or in whatever form*"

[17]Art. 90(3) of the Copyright Act defines *re-utilisation* as "*any form of making available to the public all the content of a database or a substantial part thereof by the distribution of copies, by rental or by the on-line or any other forms of transmission.*"

[18]Art. 90(1) of the Copyright Act

[19]See Art. 92 of the Copyright Act

[20]Art. 90(5) of the Copyright Act

[21]Like the Fixtures Marketing Cases (C-46/02, C-338/02, C-444/02) and the British Horseracing Board Case (C-203/02)

[22]Judgment of the Court (Fourth Chamber) of 9 October 2008, Directmedia Publishing GmbH v Albert-Ludwigs-Universität Freiburg. Case C-304/07.

Judgment of the Court (Fourth Chamber) of 5 March 2009, Apis-Hristovich EOOD v Lakorda AD. Case C-545/07.

[23]Judgment of the Court (Third Chamber) of 18 October 2012, Football Dataco v. Sportradar. Case C-173/11

Judgment of the Court (Fifth Chamber) of 19 December 2013, Innoweb BV v Wegener ICT Media BV and Wegener Mediaventions BV. Case C-202/12.

Although the SGDR can be transferred, protected databases are more commonly made available via a contractual license[24] and there are standard public licenses suitable for this purpose, which we will discuss later.

Rights of third parties

Database rights

If we want to incorporate someone else's database (we will call it the External Database) or a part thereof into our database, we need to be careful not to infringe on the maker's SGDR or the author's copyright. Since this paper focuses on situations where we will probably not encounter any copyrightable databases, we will only consider the SGDR, but the considerations for copyright are similar to a certain degree.

External Database not published under a public license

We need to determine the following with regard to what has been discussed in the previous chapter:

1. Is the External Database protected by the SGDR? If so, then:
2. Will we include a substantial part of the External Database? If so:
3. Do we have a license to extract and re-utilize a substantial part of the External Database? If so:
4. Will the External Database remain relatively intact when included in our database? If so:
5. Does the license permit us to sub-license to others? Is the scope of the license broad enough?

1. Is the External Database protected by the SGDR?

Question no. 1 can be answered by applying the definition mentioned earlier – we must assess whether the creation of the External Database has required a substantial investment.

2. Will we include a substantial part of the External Database?

There is no exact delimitation of what constitutes a substantial part of a database, but when deciding on the answer to question no. 2, we should remember that a part can be substantial in terms of quantity or quality, where quantity refers simply to the volume of the contents and quality refers to the scale of the investment when obtaining, verifying and presenting the contents (sometimes a quantitatively small part may represent a large part of the investment).[25]

3. Do we have a license to extract and re-utilize a substantial part of the External Database?

Question no. 3 is rather straightforward – in order to include a substantial part of the External Database in our database, we need to extract that part, for which we need a license (or the maker of the External Database could transfer the SGDR to us, but that is not very likely).

[24]CONNELLY KOHUTOVÁ, Radka. Databáze ve věku informační společnosti a jejich právní ochrana. Vyd. 1. V Praze: C.H. Beck, 2013, s. 55. ISBN 978-80-7400-493-3.

[25]See Judgment of the Court (Grand Chamber) of 9 November 2004, The British Horseracing Board Ltd and Others v William Hill Organization Ltd. Case C-203/02

4. Will the External Database remain relatively intact when included in our database?

As we mentioned before, the SGDR does not protect the contents of the database themselves, but rather the "databaseness" of the whole. The purpose of question no. 4 is basically to determine whether publishing our database would effectively allow the public to access the External Database (or a substantial part thereof) or just its contents.

If we take a substantial part of the External Database and append it to our database without any significant modifications, the maker of the External Database will still hold the SGDR to that part of our database. This means that in order to be able to publish our database, we will need a license to re-utilize the External Database. If, on the other hand, we take the data from the External Database, but rearrange it, select only some of it, combine it with different data and then use it in our database, the maker of the External Database will hold no SGDR relating to our database. Given that these examples obviously merely constitute the ends of the spectrum of possible situations, it will often be difficult to determine whether there are third-party SGDR pertaining to our database.[26] It is therefore advisable to be careful when publishing databases that include third-party data and to obtain the appropriate licenses when in doubt.

5. Does the license permit us to sub-license to others? Is the scope of the license broad enough?

Let us say that we have included a substantial part of the External Database in a mostly intact form in our database. Even if the maker of the External Database does grant us a license to extract and re-utilize the External Database, this does not yet mean that we can publish our database under the terms of a public license. According to Article 2363 of the Civil Code (Act no. 89/2012 Coll., Civil Code): "*The licensee may grant an authorisation under the licence in full or in part to a third party only where the license agreement so stipulates.*" This means that the license has to expressly state that we can sublicense the right to extract and re-utilize the External Database to others.

Another important issue to consider is the fact that a public license usually offers a very broad set of permissions. Remembering the *nemo plus iuris* rule (i.e. that no one can transfer more rights than he himself has), we need to make sure that the license to the External Database grants us all the permissions that we intend to offer others through a public license. Standard public licenses apply worldwide, for the whole term of SGDR protection and, barring some express exemptions, without restrictions with regard to the ways of use or quantity.

According to Czech law, if a license does not state any limitations pertaining to the ways of use, the quantity, the territory or the time, these limitations will be determined according to the purpose of the license agreement and, if the scope of these limitations is not otherwise implied by the purpose of the agreement, it is deemed to be as follows:

"*a) The territorial scope of the licence is limited to the territory of the Czech Republic;*

b) The time scope of the licence is limited to the normal time for the given type of work and way of use, however not longer than one year from the date on which the licence was granted; where the work ought to be delivered only after the granting of the licence, the term shall be one year from the date of delivery;

[26]Caution! Remember that regardless of the answer to this question, we still need a license to extract a substantial part of the External Database in the first place

c) The quantity scope of the licence is limited to the normal quantity for the given type of work and way of use."[27]

As we can see, if the license to extract and re-utilize the External Database does not expressly specify its scope, we will have to deduce it from the purpose of the agreement. If it isn't clear that the purpose of the agreement was to allow us to publish our database under a public license, it may be difficult to argue that the purpose of the agreement implies a broad enough scope for the license. Therefore, if we are able to set or negotiate the terms of the license, we should aim to have the license either expressly stipulate that it allows us to publish our database under the terms of a specific public license, or expressly set out its scope at least at the extent of the public license we intend to use.

External Database published under a public license

Using a substantial part of an External Database licensed under a public license is relatively simple (the purpose of public licensing is, after all, to make this simple), but there are still some things that we should bear in mind. Firstly, we need to determine whether there is a limitation to the public license that would prevent us from utilizing the External Database (such as a No Derivatives limitation, clashing Share-Alike clauses, a Non-Commercial condition if we want to use our database commercially, or any other restrictions like technical requirements on derivative databases etc.). As public licenses generally do not restrict multiple licensing, we could try to obtain an individual license from the maker, if the terms of the public license would prevent us from making use of the External Database. We also need to be sure to attribute properly, while remembering that the External Database might contain parts of other databases, so there might be more than one database maker to give attribution to – it is important to find and carefully read the attribution notices.

Personal Data

According to the Personal Data Protection Act (Act no. 101/2000 Coll., on the Protection of Personal Data and on Amendment to Some Acts) and the Data Protection Directive (Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data), personal data is "*any information relating to an identified or identifiable natural person ('data subject')*."[28] This means, for example, that a person's name in connection with any other piece of information about them is deemed to be personal data and therefore some metadata, like bibliographic data, will contain personal data.

Storing personal data, or performing any other operations upon it (i.e. processing personal data[29]), is usually only allowed with the data subject's consent. However, the Personal Data Protection Act provides some exceptions where consent is not needed. Given that we will mostly be processing personal data which is publicly available on the Internet, we will be able to make use of the exception[30] for processing lawfully published data.[31] If we were to process

[27]Art. 2376 of the Civil Code, in connection with Art. 2389

[28]Art. 4(a) of the Personal Data Protection Act and Art. 2(a) of the Data Protection Directive

[29]Art. 4(e) of the Personal Data Protection Act

[30]Art. 5(2)(d) in connection with Art. 4(l) of the Personal Data Protection Act

[31]See NONNEMANN, František. Další zpracování oprávněně zveřejněných osobních údajů. Právní rozhledy: časopis pro všechna právní odvětví. Praha: C. H. Beck, 2011, č. 18.

data that was not published, we would almost certainly still be processing the data for the purposes of legitimate interests which also does not require the data subjects' consent. In that case, however, we would need to inform the data subjects of the fact that their personal data was being processed.[32]

There are also other requirements set by the Personal Data Protection Act which aim to ensure that all processing is fair, lawful and not excessive, but considering the minimal scope of personal data processing that is required for maintaining and publishing a database of metadata, they will not be an issue for us.

Copyright in contents

If the metadata in our database contains copyright-protected works, such as abstracts in bibliographical databases, we need to consider whether we have the right to use the works ourselves (especially the right to communicate them to the public[33]) and to license this right to others in order to allow for the good re-usability of our database.

These works will often be subject to the Employee Work regime set out in Article 58 of the Copyright Act. In such cases, if we are the employer, we will usually exercise the author's economic rights to the work and should therefore have no problems with it. If we do not exercise these rights, we should try to obtain a very broad license to these works (similarly to the issue described in 3.1., question 5). If we do not succeed in obtaining an appropriate license to some works, we should consider removing such works from our database altogether.

Contracts and policies

Finally, we should consider whether there are any contractual or policy limitations to how we can license our database. For example, we may have received some external funding, which requires the results to be published in a certain way, or there might be guidelines on publishing set by institutional/company policy.

Selecting the right public license

We need to select the appropriate public license based on the limitations that we have discussed in previous chapters and on the purpose we aim to achieve by licensing our database under a public license. An analysis of the available options would be beyond the scope of this paper, but we would probably choose from the range of Creative Commons licenses version 4.0[34], the Open Data Commons licenses[35] or one of the public domain dedication tools such as CC0[36] or ODC PDDL[37]. If none of these licenses suit our needs and limitations, we can consider

[32]Art. 11(5) in connection with Art. 5(2)(e) of the Personal Data Protection Act

[33]Art. 18 of the Copyright Act

[34]About The Licenses. [online]. [cit. 2014-11-20]. Available from: http://creativecommons.org/licenses/

[35]Licenses | Open Data Commons. [online]. [cit. 2014-11-20]. Available from: http://opendatacommons.org/licenses/

[36]Creative Commons – CC0 1.0 Universal. [online]. [cit. 2014-11-20]. Dostupné z: http://creativecommons.org/publicdomain/zero/1.0/

[37]Open Data Commons Public Domain Dedication and License (PDDL) | Open Data Commons. [online]. [cit. 2014-11-20]. Available from: http://opendatacommons.org/licenses/pddl/

creating a bespoke license, but it is usually better to stick with the standardized licenses, if at all possible, especially for the sake of interoperability and ease of re-use.

Conclusion

There are a lot of legal obstacles to navigate before one can publish a database of metadata under a public license. It is important to remember that, while sometimes the publishing of the database is simple and without problems, there are often issues that need to be resolved first and in some cases it simply isn't possible at all. However, this should not be discouraging, because we should be able to identify the threats and choose the right approach when armed with some knowledge and a bit of foresight.

References

LEGISLATION

Act no. 121/2000 Coll., on Copyright and Rights Related to Copyright.

Act no. 89/2012 Coll., Civil Code.

Act no. 101/2000 Coll., on the Protection of Personal Data and on Amendment to Some Acts.

Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases.

Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data.

Berne Convention for the Protection of Literary and Artistic Works and Art.

JUDICIAL DECISIONS

Judgment of the Court of European Union of 9 November 2004, Fixtures Marketing Ltd v Oy Veikkaus Ab. Case C-46/02.

Judgment of the Court (Grand Chamber) of 9 November 2004, Fixtures Marketing Ltd v Organismos prognostikon agonon podosfairou AE (OPAP). Case C-444/02.

Judgment of the Court (Grand Chamber) of 9 November 2004, Fixtures Marketing Ltd v Svenska Spel AB. Case C-338/02.

Judgment of the Court (Grand Chamber) of 9 November 2004, The British Horseracing Board Ltd and Others v William Hill Organization Ltd. Case C-203/02.

Judgment of the Court (Fourth Chamber) of 9 October 2008, Directmedia Publishing GmbH v Albert-Ludwigs-Universität Freiburg. Case C-304/07.

Judgment of the Court (Fourth Chamber) of 5 March 2009, Apis-Hristovich EOOD v Lakorda AD. Case C-545/07.

Judgment of the Court (Third Chamber) of 1 March 2012, Football Dataco Ltd and Others v Yahoo! UK Ltd and Others. Case C-604/10.

Judgment of the Court (Third Chamber) of 18 October 2012, Football Dataco v. Sportradar. Case C-173/11.

Judgment of the Court (Fifth Chamber) of 19 December 2013, Innoweb BV v Wegener ICT Media BV and Wegener Mediaventions BV. Case C-202/12.

Feist Publications, Inc., v. Rural Telephone Service Co., 499 U.S. 340 (1991).

OTHER SOURCES

CONNELLY KOHUTOVÁ, Radka. Databáze ve věku informační společnosti a jejich právní ochrana. Vyd. 1. V Praze: C.H. Beck, 2013, xiv, 207 s. ISBN 978-80-7400-493-3.

LEŠČINSKÝ, Jan. Právní ochrana databází. Vyd. 1. Praha: C.H. Beck, 2003, xiv, 173 s. ISBN 80-717-9833-9.

NONNEMANN, František. Další zpracování oprávněně zveřejněných osobních údajů. Právní rozhledy: časopis pro všechna právní odvětví. Praha: C. H. Beck, 2011, č. 18.

CJEU case on 'screen-scraping' has potential to affect business models, says expert. In: *Out-Law.com* [online]. 2014 [cit. 2014-11-20]. Available from: http://www.out-law.com/en/articles/2014/november/cjeu-case-on-screen-scraping-has-potential-to-affect-business-models-says-expert/.

About The Licenses. [online]. [cit. 2014-11-20]. Available from: http://creativecommons.org/licenses/.

Licenses | Open Data Commons. [online]. [cit. 2014-11-20]. Available from: http://opendatacommons.org/licenses/.

Creative Commons – CC0 1.0 Universal. [online]. [cit. 2014-11-20]. Available from: http://creativecommons.org/publicdomain/zero/1.0/.

Open Data Commons Public Domain Dedication and License (PDDL) | Open Data Commons. [online]. [cit. 2014-11-20]. Available from: http://opendatacommons.org/licenses/pddl/.