

Why CC BY-SA is Unsuitable

This document attempts to put forward the case for moving away from Creative Commons' BY-SA license for Open Data. This is not the case *for* ODbL, only *against* CC BY-SA. Some of the cases also apply to other attribution/share-alike licenses.

Audience

This document is intended for "knowledgeable laymen"; those who understand the general concepts behind copyright law, the nature of the OSM data and the uses of it in the community. I have tried to cite as many sources as possible so that those more knowledgeable can seek them out for themselves, rather than rely on my excerpts.

Lack of Copyright Protection

An important case from U.S. law is *Fiest v. Rural*, which established that facts are not copyrightable. From Wikipedia [1]:

It is a long-standing principle of United States copyright law that "information" is not copyrightable, [Justice] O'Connor notes, but "collections" of information can be. Rural claimed a collection copyright in its directory. The court clarified that the intent of copyright law was not, as claimed by Rural and some lower courts, to reward the efforts of persons collecting information, but rather "to promote the Progress of Science and useful Arts" ([U.S. Const. 1.8.8](#)), that is, to encourage creative expression.

This is in contrast to the U.K., and some other jurisdictions, in which the "sweat of the brow" doctrine does reward the efforts of persons collecting information.

The ruling suggests that many parts of OSM data are not copyrightable; for example names, reference codes and house numbers. O'Connor's verdict contains:

The sine qua non of copyright is originality. To qualify for copyright protection, a work must be original to the author. [...] Original, as the term is used in copyright, means only that the work was independently created by the author (as opposed to copied from other works), and that it possesses at least some minimal degree of creativity. [...] To be sure, the requisite level of creativity is extremely low; even a slight amount will suffice. The vast majority of works make the grade quite easily, as they possess some creative spark, "no matter how crude, humble or obvious" it might be.

OSM meets the originality requirement, as long as it isn't copied from other works and it might be argued that the selection of appropriate tags, or the alignment of geometry to GPS traces or aerial imagery requires a "minimal degree of creativity". This is picked up by the Science Commons FAQ [2]:

In the United States, data will be protected by copyright only if they express creativity. Some databases will satisfy this condition, such as a database containing poetry or a wiki containing prose. Many databases, however, contain factual information that may have taken a great deal of effort to gather, such as the results of a series of complicated and creative experiments. Nonetheless, that information is not protected by copyright and cannot be licensed under the terms of a Creative Commons license.

And from [3]:

The merger doctrine in copyright states that if an idea and the expression of the idea are so tied together that the idea and its expression are one - there is only one conceivable way or a drastically limited number of ways to express and embody the idea in a work - then the expression of the idea is uncopyrightable because ideas may not be copyrighted.

Replacing "idea" with "fact" we can conclude that if there is a canonical form for tags or geometry then we will approach uncopyrightability. Given the "on the ground rule" [4] and "verifiability" rule [5] imply that the representation of a real-world feature in OSM should be independent of the contributor, it isn't clear that there is more than "a drastically limited number" of expressions.

Note that although *Mason v. Montgomery* [6] recognises the copyrightability of the maps themselves;

The protection that each map receives extends only to its original expression, and neither the facts nor the idea embodied in the maps is protected.

The requirement for creativity seems to be linked to the pictorial display of the information and the skill and judgement used, rather than the selection of sources of facts;

Because Mason's maps possess sufficient creativity in both the selection, coordination, and arrangement of the facts that they depict, and as in the pictorial, graphic nature of the way that they do so, we find no error in the district court's determination that Mason's maps are original.

Version 3.0 of the CC BY-SA license [10] defines "work" in section 1.h as:

"Work" means the literary and/or artistic work offered under the terms of this License...

This is followed by a non-exclusive list of things considered "works", which includes maps, but does not include the data underlying the maps to the extent such data is not protected by copyright. While it is possible that the definition includes the data without explicitly mentioning it, it is also possible that data would not fall under the definition of "work", meaning that CC BY-SA would protect the appearance and rendering of OSM data, but not the substance - the data.

CC BY-SA licenses are licenses in copyright only and were created with copyright specifically in mind. OSM's factual data is not copyrightable in most jurisdictions, except with respect to some "thin" copyright in the structure or organization of the databases that is not particularly valuable to the project.

Conclusion: It is quite likely that OSM data is not protected by U.S. (and other jurisdictions') copyright laws. This means it is also quite likely that CC BY-SA, which relies on copyright in the data, not the data collection, does not protect OSM data. Consequently, a different type of license and/or agreement would better protect the data we care about.

Combining CC BY-SA with Other Data

When CC BY-SA OSM data is used to render a map (e.g: tiles) then that map is also CC BY-SA licensed. Therefore, combining these rendered items with other data would require that the other data is compatible with, or can be released under, CC BY-SA. This makes it very difficult, or impossible, to use certain data sources in rendered maps.

It is also unclear whether that means that OSM data is allowed to be combined with proprietary data, even when no improvements are made to the map. For example; Partitioning OSM data using polygons from a commercial data set (e.g: ONS "super output areas" [7]) for the purposes of analysis and statistical reporting.

Share-Alike

One of the Use-Cases for the new license, entitled "OSM in Google Map Maker" [8], says;

We would like to avoid someone like Google loading the whole of OSM into their Map Maker system, where Google then lay claim to any further improvements made by users. It is ok for them to load OSM, but improvements must then be shared back.

In such a case, Google would be required to distribute the rendered tiles under CC BY-SA, but they would be free to continue to use and improve the data without releasing it. This is a case which the community does not want to happen, yet it may be possible under CC BY-SA.

While a rendered map must be distributed under CC BY-SA, there is no requirement in that license to release the data from which the map was rendered. This allows a situation in which OSM data can be improved using some additional data source or effort, the result rendered as tiles distributed under CC BY-SA, but the improved data may not need to be made available.

Attribution

CC BY-SA 2.0 [9] has the following provision for attribution:

If you distribute, publicly display, publicly perform, or publicly digitally perform the Work or any Derivative Works or Collective Works, You must keep intact all copyright notices for the Work and give the Original Author credit reasonable to the medium or means You are utilizing by conveying the name (or pseudonym if applicable) of the Original Author if supplied; the title of the Work if supplied; to the extent reasonably practicable, the Uniform Resource Identifier, if any, that Licensor specifies to be associated with the Work, unless such URI does not refer to the copyright notice or licensing information for the Work; and in the case of a Derivative Work, a credit identifying the use of the Work in the Derivative Work (e.g., "French translation of the Work by Original Author," or "Screenplay based on original Work by Original Author").

When CC BY-SA is used for user contributions as well as data distribution then the license could be interpreted as requiring all contributors names (or pseudonyms) at a minimum. This is clearly practically unworkable and has been widely ignored in the OSM community for many years. A publisher from outside the community may feel that OSM data is unsuitable for print use, as it might leave them liable for copyright infringement if any one member of the community decided to enforce the license. This is, presumably, not the intention of the community.

Uncertainty and doubt over extent of derived work

There have been a number of cases where companies have looked to use OSM or OSM-derived information, but their lawyers have recommended against use of our maps [11]. Their concerns were due to the unclear boundary between collective and derived works; does the derived work extend only to the base map, the base map plus any overlays, or to the work around it such as a news show or a book? Whilst a community guideline exists on our wiki, the fact that there are tens of thousands of copyright holders means that statements made on the wiki or by OSMF cannot be authoritative, leaving the possibility that any one contributor may claim that there is a failure to properly comply with CC BY-SA. This is a risk that many companies find off-putting.

References

- [1] Feist v. Rural, http://en.wikipedia.org/wiki/Feist_v._Rural
- [2] Science Commons FAQ, <http://sciencecommons.org/resources/faq/databases#canicc>
- [3] Copyright, Originality, and the End of the Scenes a Faire and Merger Doctrines for Visual Works, Michael D. Murray, 2006, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=900148
- [4] http://wiki.openstreetmap.org/wiki/Disputes#On_the_Ground_Rule
- [5] <http://wiki.openstreetmap.org/wiki/Verifiability>
- [6] Mason v. Montgomery Data, <http://www.coolcopyright.com/cases/fulltext/masonmontgomerytext.htm>
- [7] <http://www.ons.gov.uk/about-statistics/geography/products/geog-products-area/names-codes/soa/index.html>
- [8] http://wiki.openstreetmap.org/index.php/Brief_for_proposed_OSM_licence#OSM_in_Google_Map_Maker
- [9] Creative Commons Attribution Share-Alike 2.0, <http://creativecommons.org/licenses/by-sa/2.0/legalcode>

[10] Creative Commons Attribution Share-Alike 3.0, <http://creativecommons.org/licenses/by-sa/3.0/legalcode>

[11] NPE maps FAQ, <http://www.npemap.org.uk/FAQ.html#openstreetmap>