

AUTHORSHIP IN WIKIPEDIA: LEGAL REQUIREMENTS, COMMUNITY OPINIONS, AND TECHNICAL BOUNDARIES*

by

THOMAS ROESSING**

One aim of the Wikipedia project is to provide mankind with free knowledge. This goal requires a set of licensing models for text and images. Community opinions differ on how free content should be, what the best licensing model is to achieve the objective of free content and on how Wikipedia should deal with commercial use of its contents. This paper explains the licensing problem and discusses community opinions on the 2009 license migration of the online encyclopedia project.

KEYWORDS

Wikipedia, intellectual property, free knowledge, licensing

* * *

This paper addresses the licensing of the free online encyclopedia Wikipedia. Although it is concerned with legal aspects of licensing, the main perspective of the analysis is a social-psychological one. After a description of Wikipedia's licensing system (for text and images) community opinions on consequences of licensing for free content and its contributors are analysed.

1. LICENSING FREE INFORMATION

Similar to free¹ software licensing free information licensing aims at making and keeping the licensed content available to the general public (O' Sullivan 2009: 71). The licensor grants the licensee the freedom to make further use of

* Full Paper: http://www.ifp.uni-mainz.de/files/roessing_wikipedia_and_public_opinion_political_communication.pdf

** roessing@uni-mainz.de

¹ Just for clarification: „Free“ refers to freedom of further use and does not necessarily mean „free of charge“.

the contributed content within the boundaries of the chosen free licensing model. From its beginning in 2001 until 2009 the Wikipedia project published its contents under the GNU Free Documentation License (GFDL)². This licence is a fine example of how free licensing works, what typical boundaries for further use are, and why the Wikimedia Foundation (the organisation which runs Wikipedia) finally decided to change licensing to a certain kind of Creative Commons license.

The further user of content published under the GFDL is granted to “copy and distribute the Document in any medium, either commercially or non-commercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies”, and that no other conditions whatsoever are added to those of the License (GFDL, Section 2). Detailed regulations of the licence text address

- Secondary sections, that may be declared as invariant
- Invariant sections that are forbidden except for front-cover texts or back-cover texts
- Transparent (machine readable) text
- Several details of legal nature.

The aim of these regulations is “to make a manual, textbook, or other functional and useful document “free” in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially.” (GFDL, Section 0). The fact that the GFDL was originally developed for manuals and textbooks accompanying free software is the reason why the license requires the further user to reproduce the complete licence text in all copies. It is no problem for the publisher of a handbook to add a few pages to this kind of document. But it is a problem, if small parts of text or images that have been licensed under the GFDL are to be re-published; therefore, the requirement to publish the license text was a problem for further users of Wikipedia’s content.

For further use of a small piece of GFDL-licensed text (e. g. the introductory words of <http://cs.wikipedia.org/wiki/Brno> for a small promotion flyer) the complete GFDL license text would have to be added to each copy. Even a single sheet version using a very small font size still requires an additional letter-size page. Thus, for example, printing postcards of GFDL-licensed pictures is frankly impossible. With this problem in mind, the Wikimedia foundation decided in 2009 to switch to the Creative Com-

² <http://www.gnu.org/copyleft/fdl.html> [2009-12-16]

mons licensing system.³ Creative Commons (CC) offers a variety of licensing options:

- *By*. The author(s) must be named
- *Nc*. No commercial use allowed
- *Nd*. No derivatives allowed: the work must not be changed
- *Sa*. Share alike: The work can only be further used when licensed under the same conditions.

Six commonly used licensing models result from these four options:

- *CC-by*: The author(s) must be named (every other kind of further use like changing or commercial use are liable to the will of the further user).⁴
- *CC-by-sa*: Author(s) must be named and copies and derivatives must be published under the same conditions.⁵
- *CC-by-nd*: Author(s) must be named, the work must not be modified.⁶
- *CC-by-nc*: Authors must be named, the work must not be commercially distributed.⁷
- *CC-by-nc-sa*: Authors must be named, no commercial use, share alike.⁸
- *CC-by-nc-nd*: Authors must be named, no commercial use, no derivatives.⁹

There are also some newer licensing models dealing with remixes of free music.¹⁰

The Wikimedia foundation – that is the organisation that runs Wikipedia – chose *Creative Commons attribution share alike* (CC-by-sa) as the new licensing model for Wikipedia, because it is very similar to the GFDL but without some of its restrictions, especially the requirement to print the full text of the licence with each copy. CC-by-sa grants the rights (CC-by-sa, Section 3)

³ http://meta.wikimedia.org/w/index.php?title=Licensing_update/Timeline&oldid=1558317

⁴ <http://creativecommons.org/licenses/by/3.0/legalcode> [2009-12-17]

⁵ <http://creativecommons.org/licenses/by-sa/3.0/legalcode> [2009-12-17]

⁶ <http://creativecommons.org/licenses/by-nd/3.0/legalcode> [2009-12-17]

⁷ <http://creativecommons.org/licenses/by-nc/3.0/legalcode> [2009-12-17]

⁸ <http://creativecommons.org/licenses/by-nc-sa/3.0/legalcode> [2009-12-17]

⁹ <http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode> [2009-12-17]

¹⁰ E. g. <http://creativecommons.org/licenses/sampling+/1.0/legalcode> [2009-12-17]

- “to Reproduce the Work, to incorporate the Work into one or more Collections, and to Reproduce the Work as incorporated in the Collections”
- “to create and Reproduce Adaptations [...]”
- “to Distribute and Publicly Perform the Work including as incorporated in Collections”
- “to Distribute and Publicly Perform Adaptations”

The corresponding requirements are put down in section 4 of the licence (selection):

- Copies and derivatives must be published under the same licence (by-sa), the further user “must include a copy of, or the Uniform Resource Identifier (URI) for, this License with every copy of the Work You Distribute or Publicly Perform.” (CC-by-sa, Section 4)
- In case of redistribution or performance of the work, the further user must “keep intact all copyright notices for the Work and provide, reasonable to the medium or means You are utilizing: (i) the name of the Original Author (or pseudonym, if applicable) if supplied [...]”

Obviously, CC-by-sa is very similar to the original GFDL license with the important difference that the Creative Commons license offers the option to provide only a Web link to the license text instead of printing the whole license.¹¹

2. WIKIPEDIA’S LICENSING SHIFT

To prepare the shift from GFDL to CC-by-sa, Wikimedia set up a committee of nine international users¹² and a timeline beginning with some software preparations and ending with the implementation of the new license in June 2009.¹³ The proceedings were discussed and documented on *Meta* (meta.wikimedia.org, a website dedicated to the organisation and coordination of Wikimedia’s projects). Oppositional arguments were collected and discussed on a special subpage.¹⁴ Besides questions of reasonable means of attribution and meanings of legal terms the following objection is discussed:

¹¹ See also http://meta.wikimedia.org/w/index.php?title=Licensing_update/License_comparison&oldid=1497651

¹² http://meta.wikimedia.org/w/index.php?title=Licensing_update/Committee&oldid=1496672

¹³ http://meta.wikimedia.org/w/index.php?title=Licensing_update/Timeline&oldid=1558317

¹⁴ http://meta.wikimedia.org/wiki/Licensing_update/Questions_and_Answers/Oppositional_arguments

User *Wnt* argues that the new license could foster the misuse of Wikipedia's contents by "Google-spammers". The user refers to the use of free content to lure internet users to commercial web sites either flooding users with ads or leading them astray. Obviously, some forms of further commercial use are not really accepted by some members of the community.

However, the Meta is observed only by highly involved users who are interested in the organisation of the project. To gauge the opinion of ordinary contributors it is necessary to have a look into the discussions within the different language versions of Wikipedia. The English discussion took place on various discussion pages (talk) of pages related to copyright-policy.¹⁵ Generally, a recurring argument takes place about the ban on non-commercial licenses, i. e. licenses admitting further use if it is non-commercial or for educational purposes (or for Wikipedia) only. Some users have difficulties to accept that material licensed for Wikipedia is – within the boundaries of the license and copyright laws – free for commercial use. This licensing policy (for text and images or files alike) goes back to a decision made by Wikipedia's founder Jimmy Wales from 2004.¹⁶

User Muhammad is quoted with his opinion on commercial use of Wikipedia's content:¹⁷

"Having recently invested quite a lot in photographic gear, I am reluctant to give away my images for free. I have no problem with wikipedia or any other non-commercial organization using them, however I would not like commercial institutions to benefit at my expense. I therefore suggest wikipedia adopt a non commercial license such as CC-NC . This will not violate wikis principles of free knowledge and will convince photographers to release higher resolution pictures as well. It might also convince professional photographers to release their work knowing that their work will not be used for commercial means and that they will still be able to make a living."

Another opinion is stated by user *Cavebear42* on the talk page of the site where users discuss especially good images ("featured images"). This user introduces the aim to create free content into the discussion about commercial and non-commercial use of Wikipedia's contents:¹⁸

"Because we really want to use it [a good non-free image] isn't a justification for infringing copyright. the point of this project (the wiki) is to create

¹⁵ E. g. http://en.wikipedia.org/wiki/Wikipedia_talk:Copyrights/Archive_13

¹⁶ <http://lists.wikimedia.org/pipermail/wikien-l/2004-April/012156.html>

¹⁷ http://en.wikipedia.org/w/index.php?title=User:Tillman/Non-commercial_image_policy&oldid=276352518

¹⁸ http://en.wikipedia.org/w/index.php?title=Wikipedia_talk:Featured_picture_candidates/Archive_2&oldid=133606717. Italics in original, squared brackets by the author of the present paper.

an Free content encyclopedia. It is important that *'when free content is modified, expanded, or incorporated within another work, the resulting work must be legally distributable'*. This does not mean that we can use things which are legal for use here and only here."

The two quotes illustrate a dilemma of the licensing policy for free content: Does a ban on commercial further use foster or inhibit the generation of free information? With the further use of Wikipedia's content being much easier after the migration from the GFDL to CC-by-sa this problem becomes especially obtrusive. Thus, the shift from GFDL 1.2 (including the shift to a later version of GFDL) to GFDL 1.3 (including the possibility to migrate to Creative commons) instigated a fierce discussion in the German language version of Wikipedia, addressing text licensing as well as that of images.

3. COMMUNITY OPINIONS ON COMMERCIAL USE OF FREE TEXT CONTENTS

In April 2009, an unofficial referendum on the license migration took place in the German language version of Wikipedia.¹⁹ Forty-one users voted in favour of the license shift, 36 opposed retroactive changes of material that was published under a certain licence before. Despite the fact that Voting in Wikipedia are usually public, open, and due to the lack of a defined basic population in combination with self-selection of voters not representative, at least some conclusions on opinions of engaged and interested users of Wikipedia can be drawn (Roessing 2004). Two positions become obvious in the comments, which accompany some of the votes: Some people argue that the restrictions of the GFDL contradict the idea of free knowledge; the restrictions make further use so hard that the knowledge is not really free. Most opponents of the relicensing of Wikipedia's content argue that retroactive license changes are at least illegitimate, maybe even illegal. Some opponents of the new licensing model argue that the shift endangers free knowledge by making further use easier.²⁰ This dilemma becomes much more obvious when the arguments of the broader discussion are analysed (for approach and methodology of Wikipedia analysis see Roessing 2007; 2009).

User *Andreas E. Kemper* writes on April 15, 2009 (translation by the author of the present paper [TRG]):²¹ "The Wikimedia Foundation has an interest in easier dissemination of Wikipedia's contents. Usufructuaries are the (commercial) further users [...]". This user proposes to establish a special

¹⁹ http://de.wikipedia.org/w/index.php?title=Wikipedia_Diskussion:Lizenz%C3%A4nderung&oldid=61218362

²⁰ http://de.wikipedia.org/w/index.php?title=Wikipedia_Diskussion:Lizenz%C3%A4nderung&oldid=61218362

interest group to advocate the interest of Wikipedia's authors²² – including better credits to the articles' authors. Other users adopted the idea and organised a survey among authors on how to give credit to the authors.²³ The idea itself is consistent with German copyright Law (§ 32 UrhG) which demands reasonable remuneration for authors if they license their work to third parties (Klett/Sonntag/Wilske 2008). In the way of the discussion of the proposals to name authors directly in the article some users argued that the drawbacks of the GFDL as well as the complication of further use by the obligation to give prominently placed credit to authors could be a prophylaxis against commercial use of Wikipedia. User *Goldzahn* writes on April 27, 2009 (translation by TRG):²⁴ "I have already been thinking about this, and it is why I am not really happy about the license migration. The unhandy GFDL was some kind of precaution against total commercialisation." Noticeably, opposition against commercial use is not (or at least not in all cases) motivated by blatant anti-capitalism. Some of the authors argue that they want to make further use difficult to protect free knowledge. In their opinion, commercial use of free content is exploitation of unpaid volunteers and makes free knowledge subject to the control of big media companies.²⁵ Others, like user *Sebmol* on April 28, 2009, note that commercial use is an intended part of the idea of free knowledge: "That everyone is free to use Wikipedia's contents for every (and 'every' is including 'every commercial') purpose, is the basic idea of what we are doing here. Commercial use is neither an accident, nor an unintended side effect, it is part of the mission".²⁶

However, the proposed requirement of naming the authors directly in the article soon reaches its limits:

²¹ http://de.wikipedia.org/w/index.php?title=Wikipedia_Diskussion:Lizenz%C3%A4nderung&oldid=61218362, original: "Die Wikimedia-Foundation hat ein Interesse daran, dass die Inhalte von Wikipedia einfacher Verbreitung finden. Nutznießer sind die (kommerziellen) Weiternutzer [...]"

²² http://de.wikipedia.org/w/index.php?title=Benutzer_Diskussion:Andreas_E._Kemper/IG_der_Wikipedia-AutorInnen&oldid=62500221

²³ http://de.wikipedia.org/w/index.php?title=Wikipedia:Umfragen/Autorennennung_am_Artikel&oldid=59843542

²⁴ http://de.wikipedia.org/w/index.php?title=Wikipedia_Diskussion:Umfragen/Autorennennung_am_Artikel&oldid=62467780, original: „Solche Gedanken sind mir auch schon durch den Kopf gegangen, weshalb ich den Lizenzwechsel nicht wirklich mit Freude sehe. Die unpraktische GFDL war auf jeden Fall ein gewisser Schutz vor der totalen Kommerzialisierung.“

²⁵ http://de.wikipedia.org/w/index.php?title=Wikipedia_Diskussion:Umfragen/Autorennennung_am_Artikel&oldid=62467780#Kommerzialisierung_durch_die_Hintert.C3.BCr.3F

²⁶ http://de.wikipedia.org/w/index.php?title=Wikipedia_Diskussion:Umfragen/Autorennennung_am_Artikel&oldid=62467780, original: "Dass jeder die Inhalte der Wikipedia für jeden (und "jeden" schließt "jeden kommerziellen" ein) Zweck verwenden kann, ist Grundgedanke dessen, was wir hier tun. Die kommerzielle Verwendung der Inhalte ist weder Unfall noch unerwollte Nebenwirkung, sie ist Teil der Mission."

1. At least German law on intellectual property requires that a contribution of an author exceeds a threshold of originality (derived from § 2 UrhG). Minor contributions are in the public domain anyway, and, therefore, further users can only be required to give credit to those authors with major contributions.
2. Unfortunately, there is no computer software yet that could determine if a contribution exceeds the level of originality (adding to that is the problem that the definition of this level is not very exact). Who should determine which of the many (dozens, hundreds, in some cases thousands) of Wikipedia's editors deserve to be named as authors?
3. Naming a huge number of accounts (in many cases funny or provocative nicknames) may be practicable on a computer screen, but it is complicated, ill-favoured, and costly for printed articles.

4. COMMUNITY OPINIONS ON COMMERCIAL USE OF FREE IMAGES

The challenges of the relationship between free content and commercial use become most visible in the realm of image licensing. The problem is this. Version 1.2 of the GFDL contains a paragraph indicating that the licensor accepts later versions of the GFDL for the licensed content, too:

"Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation".²⁷ Version 1.3 contains, as already mentioned, the possibility to migrate to *Creative Commons attribution share alike* licensing: "An MMC [Massive Multiauthor Collaboration Site] is "eligible for relicensing" if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing."²⁸

The requirement to print the whole license constricts especially the further use of images, because postcards or pamphlets are usually small products with no opportunity to print a lengthy license text. Some photographers enjoyed this because they were unhappy with commercial use of

²⁷ <http://www.gnu.org/licenses/old-licenses/fdl-1.2.txt> [2010-01-10]

²⁸ <http://www.gnu.org/licenses/fdl-1.3.txt> [2010-01-13]

their images. When GFDL 1.3 came up, some of them started to re-license their images to “GFDL 1.2 only” in order to save the factual non-commercial status of their contributions to Wikipedia.²⁹ This sparked a fierce and in parts crude discussion among Wikipedians. Some quotes illustrate the different points of view (translation by TRG). User *Sebmol* writes on January 11, 2008:

“The question is not, if this license template [GFDL 1.2 only] should exist, but if such a licensing for Wikipedia can be regarded as ‘free content’. We must not forget that ‘free content’ is about providing content for everyone and for every purpose (leaving legal requirements aside). Attempts to constrict further use, e. g. to prevent commercial use, are not compatible with this approach. It is an open secret that many image uploaders use GFDL just because this license impedes further use. This practice, which contradicts the idea of free content, should not be fostered by the possibility of freezing the licensing to a certain version without ‘or later [version]’ passus.”³⁰

User *Eva K.* objects to such positions. On November 28, 2008, she states (translation by TRG) that she regards excessive freedom for further users as a

“burlesque of freedom, because of the idea behind it that products of mental creativity are fair game, open for everyone’s use like free beer. This impedes the creativity, because professionals like artists, musicians, photographers and authors are to be ripped of their living in favour of the close-fistedness of the masses and the greed of commercial exploiters [...] I do not want to throw the rights to my works to some perfect stranger exploiters – people that actually should be stopped by free content. But freedom cannot be achieved for free [...]”³¹

Both parties claim to think or act in favour of free content – either by fostering commercial use, or by impeding it. After Wikipedia’s migration to GFDL 1.3 and then to CC-by-sa, time will tell if this decision was right, and to what extent commercial use endangers the freedom of free content.

²⁹ E. g. http://de.wikipedia.org/w/index.php?title=Datei:Eschborner_Stuhl.jpg&diff=cur&oldid=24781369

³⁰ http://de.wikipedia.org/wiki/Wikipedia:Fragen_zur_Wikipedia/Archiv/2008/Woche_02#Festschreibung_von_Bildlizenzen_auf_GFDL_1.2, original: “Die Frage ist nicht so sehr, ob es diese Lizenzvorlage geben sollte, sondern ob solch eine Lizenzierung für die Wikipedia trotzdem als „freier Inhalt“ angesehen werden kann. Wir dürfen nicht vergessen, dass es bei „freien Inhalten“ darum geht, Inhalte für jeden und zu jedem Zweck zur Verfügung zu stellen (rechtlich bedingte Einschränkungen unberührt). Bestrebungen, diese Weiter-nutzung einzuschränken, um zum Beispiel kommerzielle Verwertungen zu verhindern, sind mit diesem Grundsatz nicht vereinbar. Es ist ein schlecht gehütetes Geheimnis, dass viele Bilduploader GFDL verwenden, eben weil die Lizenz es selbst recht schwierig macht, die Bilder weiterzuverwenden. Dieser, eigentlich dem Grundsatz der freien Inhalte diametral entgegengesetzten Praxis sollte nicht noch dadurch Unterstützung geboten werden, in dem eine Festschreibung auf eine bestimmte Version ohne „oder später“-Klausel ermöglicht wird.”

5. CONCLUSION AND OPEN QUESTIONS

Licensing a huge international encyclopedia is a complicated matter and the community behind the online encyclopedia is at least in parts divided over the best licensing policy (cf. e.g. the essay by Eric Möller [2006], deputy of the Wikimedia Foundation). The relicensing of Wikipedia by the Wikimedia foundation sparked considerable dispute, especially on attempts to impede commercial use of free content. Additionally, some general questions from the area of licensing and intellectual property are pending or will come up in the future as the Wikipedia project is growing and free content is further developed:

- What makes a contributor become an author?
- How does licensing influence authorship as such in “Massive Multiauthor Collaboration Sites”?
- How do different local copyright laws work together in a collaborative system of hundreds of thousands of authors? There are manifold licensing models, within Wikipedia especially for images and files; authors come from all over the world, some are anonymous represented only by the Internet-Protocol-address of their computers; Wikipedia’s content can be retrieved in almost every country in the world, but the Foundation that runs it is based in San Francisco...

It seems that licensing, free content and collective authorship are a major field of interest and research for legal experts as well as for communication researchers and sociologists.

³¹ http://de.wikipedia.org/w/index.php?title=Wikipedia_Diskussion:Meinungsbilder/GFDL_1.2-only_f%C3%BCr_Dateien&oldid=54291053 , original: „[Ich sehe Deine Ansicht eher als Karikatur von Freiheit an,] weil dahinter die Mentalität steht, daß Produkte geistiger Schaffenskraft Freiwillig seien und allen als Freibier zu Verfügung zu stehen haben. Sowas unterbindet aber nur die Schaffenskraft, weil insbesondere professionell tätige Personen wie Künstler, Musiker, Fotografen und Autoren um ihren Lebensunterhalt bestohlen werden sollen zugunsten des Geizes der Massen und der Habgier kommerzieller Ausschlachter[...]. [Nein, ich lebe nicht davon, aber] ich möchte mir die Rechte an meinen Werken wahren und nicht irgendwelchen hergelaufenen Ausschlachtern zum Fraß vorwerfen – Leute, denen eigentlich mit freien Inhalten ein wenig Einhalt geboten werden soll. Aber Freiheit gibt es auch nicht für geschenkt [...]

REFERENCES

- [1] Klett, A. R. , Sonntag, M. & Wilske, S., 2008, Intellectual property law in Germany. Protection, enforcement and dispute resolution, Beck, München.
- [2] O'Sullivan, Dan (2009): *Wikipedia. A new community of practice?* Farnham: Ashgate.
- [3] Möller, E, 2006, The Case for Free Use: Reasons Not to Use a Creative Commons – NC License, in *Open Source Jahrbuch 2006*. [http://www.opensourcejahrbuch.de/download/jb2006/chapter_06/osjb2006-06-02-en-moeller.pdf]
- [4] Roessing, T., 2004, Prevalence, Technology, Design and Use of Polling on Websites. Quality Criteria despite the Impossibility of Representativeness? Paper presented at the World Association of Public Opinion Research (WAPOR) Thematic Seminar on Quality Criteria in Survey Research V in Cadenabbia, Italy, June 24-26, 2004.
- [5] Roessing, T., 2007, *Wikipedia and Public Opinion*. Paper presented at the 60th Annual Conference of the World Association of Public Opinion Research (WAPOR) in Berlin, Germany, September 19 - 21, 2007. [Full Paper: http://www.ifp.uni-mainz.de/files/roessing_wikipedia_and_public_opinion.pdf]
- [6] Roessing, T., 2008, 'Opinion Formation in Wikipedia - Theory, Measurement, and Findings', *Masaryk University Journal of Law and Technology*, vol. 2, pp. 163-176.
- [7] Roessing, T., 2009, *Wikipedia and Public Opinion: The Online Encyclopedia as a Transmission Belt for Political Communication*. Paper presented at the Second Cairo International Conference on Public Opinion Polls "Public Opinion Polls in a Changing Society." Cairo, Egypt, November 9,2009