



Master's thesis

Master's Programme in Computer Science

Public Copyright Licenses in Software Engineering: A Multivocal Systematic Literature Review

Akira Taguchi

February 16, 2024

FACULTY OF SCIENCE
UNIVERSITY OF HELSINKI

Contact information

P. O. Box 68 (Pietari Kalmin katu 5)
00014 University of Helsinki, Finland

Email address: info@cs.helsinki.fi

URL: <http://www.cs.helsinki.fi/>

HELSINGIN YLIOPISTO – HELSINGFORS UNIVERSITET – UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty		Koulutusohjelma — Utbildningsprogram — Study programme	
Faculty of Science		Master's Programme in Computer Science	
Tekijä — Författare — Author			
Akira Taguchi			
Työn nimi — Arbetets titel — Title			
Public Copyright Licenses in Software Engineering: A Multivocal Systematic Literature Review			
Ohjaajat — Handledare — Supervisors			
Prof. Tomi Männistö			
Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages	
Master's thesis	February 16, 2024	8 pages, 1 appendix pages	
Tiivistelmä — Referat — Abstract			
<p>Public copyright licenses play a major part in software engineering. For example in open source there must be an appropriate public copyright license attached to the source code in order for open-source software to be freely available for possible modification and redistribution. Understanding public copyright licenses can be difficult. This could stem from the legal nature of the license texts and the large number of already-existing public copyright licenses. sub-problem here. thesis' contribution to solution here.</p> <p>Tell about the research method here.</p> <p>Tell about the results here.</p> <p>Tell about the discussion here.</p>			
<p>ACM Computing Classification System (CCS) Social and professional topics → Computing / technology policy → Intellectual property → Licensing</p>			
Avainsanat — Nyckelord — Keywords			
open source, free / libre software, copyright, proprietary software, copyleft, license			
Säilytyspaikka — Förvaringsställe — Where deposited			
Helsinki University Library			
Muita tietoja — övriga uppgifter — Additional information			
Software study track			

Acknowledgements

Contents

1	Introduction	1
1.1	Research goal, questions and contributions	2
1.2	Background and terminology of public copyright licenses	3
1.3	Thesis structure	3
2	Methods	4
3	Results	5
4	Discussion	6
5	Conclusions	7
	Bibliography	8
A	Table of software licenses	i

1 Introduction

Public copyright licenses play a major part in software engineering. For example in open source there must be an appropriate public copyright license attached to the source code in order for open-source software to be freely available for possible modification and redistribution. Because open source is central to software engineering the licenses enabling open source must also be considered important in the same context.

Public copyright license is defined by Wikipedia with the following words (Category:Public copyright licenses, 2012):

”A public copyright license is a copyright license where the licensees are not limited. Examples include free content, open content, Creative Commons, free software and open source licences.”

Understanding public copyright licenses can be difficult. This could stem from the legal nature of the license texts and the large number of already-existing public copyright licenses. The license texts usually favors correctness over the readability for the developer. This is because the license text has to act as a valid legal instrument otherwise it cannot be endorsed (Ferguson, 2006). The lack of understanding of public copyright licenses leaves too much room for interpretation. In June 21, 2023 IBM’s Red Hat seemingly violated a public copyright license, GNU General Public License version 2 or GPLv2 for short (Kuhn, 2023) (McGrath, 2023). This was an unpleasant surprise to the public since the project behind GPL, GNU initially attempted to ensure the users via the GPL have to the following three freedoms (GNU, 1996):

- Freedom 1: The freedom to study how the program works, and change it so it does your computing as you wish. Access to the source code is a precondition for this.
- Freedom2: The freedom to redistribute copies so you can help others
- Freedom 3: The freedom to distribute copies of your modified versions to others. By doing this you can give the whole community a chance to benefit from your changes. Access to the source code is a precondition for this.

Regardless, IBM’s Red Hat essentially rendered the previously public Red Hat Enterprise

Linux, or RHEL for short, proprietary software. If the licenses would be more easily understood the proprietarization of RHEL would have been less of a surprise to the users. On top of public copyright license details, software engineers in general have a tough time understanding the basic goals of public copyright licenses used in software engineering. In the instance of the RHEL incident it would not have been a big surprise to software engineers if they would have known about other licenses and what they try to achieve or how old is GPLv2 and why it has been succeeded by GPLv3.

This thesis' goal is to contribute into the solving these problems in a structured manner. First we state definitions and terminology used in the scope of this thesis. We go over the reasons why there does not exist consistent terminology in this area and why the conversely the definitions are the most stabile ones in this area. Second we take a deep dive into the multivocal literature through a systematic literature review. To make more information available, a mapping study connected to the terminology scope defined in the first step is needed. Third step will be to offer our own suggestions and basic knowledge for professionals and academics in the industry to enhance the understanding of public copyright licenses in software engineering. This is also the basis for future research and contributes to stablizing the terminology in the academic field.

1.1 Research goal, questions and contributions

The primary objective of this research

- RQ1: How have public software licenses changed throughout the years?
- RQ2: Why have public software licenses changed throughout the years?
- RQ3: What is the speed of change in software licenses throughout the years?
- RQ4: What are the limitations for a rapidly changing software license?
- RQ5: How can we make it easier for developers to understand various public software licenses?

1.2 Background and terminology of public copyright licenses

literature review here

1.3 Thesis structure

Write how this thesis is going to contribute to the aforementioned problems.

2 Methods

1. Develop Review Protocol
2. Validate Review Protocol
3. Identify Relevant Research
4. Select Primary Studies
5. Assess Study Quality

3 Results

7. Extract Required Data

8. Synthesize data

4 Discussion

5 Conclusions

Bibliography

Category:Public copyright licenses (2012). *Category:Public copyright licenses* — *Wikipedia, The Free Encyclopedia*. https://web.archive.org/web/20240131085240/https://en.wikipedia.org/wiki/Category:Public_copyright_licenses. Accessed: 2024 Jan 31.

Ferguson, D. (2006). “Syntar Errors: Why Version 3 of the GNU General Public License Needs Debugging”. In: *North Carolina Journal of Law & Technology* 7.2, pp. 397–420. URL: <https://heinonline.org/HOL/Page?handle=hein.journals/ncjl7&id=403>.

GNU (1996). *What is Free Software?* <https://web.archive.org/web/19980126185518/https://www.gnu.org/philosophy/free-sw.html>. Accessed: 2024 Feb 8.

Kuhn, B. M. (2023). *A Comprehensive Analysis of the GPL Issues With the Red Hat Enterprise Linux RHEL Business Model*. <https://web.archive.org/web/20240205080551/https://sfconservancy.org/blog/2023/jun/23/rhel-gpl-analysis/>. Accessed: 2024 Feb 5.

McGrath, M. (2023). *Furthering the evolution of CentOS Stream*. <https://web.archive.org/save/https://www.redhat.com/en/blog/furthering-evolution-centos-stream>. Accessed: 2024 Feb 5.

Appendix A Table of software licenses

List of FOSS licenses goes here.