Proceedings of the Weizenbaum Conference 2022:
Practicing Sovereignty. Interventions for Open Digital Futures

REUSE SOFTWARE

MAKING COPYRIGHT AND LICENSING COMPLIANCE EASIER FOR EVERYONE

Lasota, Lucas
Humboldt University of Berlin
Free Software Foundation Europe
Berlin, Germany
lucas.lasota@hu-berlin.de

KEYWORDS

software license; free and open source software; copyright; compliance

DOI: 10.34669/wi.cp/4.6

ABSTRACT

Best practices for displaying data and metadata pertaining to software licensing and copyright are currently unharmonized. The multiple competing licensing requirements for communicating the chosen license of a software project and its copyright holders increase the compliance burden on project maintainers, especially for smaller free and open source (FOSS) ones. The “REUSE Software” initiative aims to remediate this situation by defining a set of easy-to-implement best practices for declaring copyright and licensing in an unambiguous, human- and machine-readable way, so that the information is preserved when the file is copied and reused by third parties. REUSE specifications facilitate management policies for digital commons, improving data and metadata communication for individuals, communities, governments, and businesses.

1 INTRODUCTION[8]

Digital transformation necessarily involves copyright and licenses, because software, the backbone element of digital technologies, is regulated by copyright.[9] The re-usability of software should be authorized by licenses or statutory copyright limitations or exceptions.[10] The multiple competing requirements for communicating the chosen license and the copyright holders increase the compliance burden on project maintainers, especially for smaller free and open source software (FOSS)[11] ones. The "REUSE Software" initiative[12] defines best practices for declaring copyright and licensing in an unambiguous, human- and machine-readable way, so that the information is preserved when the file is copied and reused by third parties. REUSE specifications aim to facilitate and improve management policies for the digital commons, improving data and metadata communication for individuals, communities, governments, and businesses.

2 CHALLENGES FOR COMPLIANCE

The more external components a software code encompasses, the harder it is to keep an overview of the copyright holders and their licensing choices. Since FOSS licenses are public documents that are shared openly, often by millions of users worldwide, their implementation generally does not involve negotiation among the parties. Therefore, proper information regarding the governing license is crucial to avoid legal (Synopsys, 2019) and security risks (Haddad, 2018). This is especially problematic for FOSS projects, as large public code repositories mean a decreased number of licensed repositories (Balter, 2015). Moreover, license proliferation fragments the requirements for copyright and license notices (OSI, 2006). Software projects incorporating content elements—as text, images, and videos—face an additional layer of complexity with content licensing compliance.[13]

How copyright and license information should be displayed depends on copyright law and license requirements. Especially important are notices for reciprocal licenses (also known as copyleft), as they require the derivative work to be licensed under the same licensing terms, which directly impacts license compatibility (Ku Wei Bin, Lasota & Jaeger, 2022). Although FOSS licenses

[8] This paper does not necessarily reflect the views of any organization the author may represent. The author thanks Richard Schmeidler and the reviewers for the proofreading and comments on the text.
[9] For the European Union, see Art. 1(1) of Software Directive (2009/24/EC) from 23 April of 2009.
[10] Regarding statutory exceptions, see Blázquez, Cappello & Valais, 2017.
[11] The definitions of free and open source software are taken respectively from the Free Software Foundation and Open Source Initiative. See Ku Wei Bin, Lasota & Jaeger, 2022, p. 10.
[12] See the project's web portal. Available at: https://reuse.software/ Retrieved on 30.06.22.
[13] See, for instance, the Creative Commons recommendations for applying a license to creative works. Available at: https://wiki.creativecommons.org/wiki/Marking_your_work_with_a_CC_license Retrieved on 30.06.22.

in general provide information on how the license notices should be applied, the vastly diverse recommendations remain unharmonized.

3 REUSE: SETTING HARMONIZED BEST PRACTICES

The REUSE best practices enable humans and machines alike to add and read data and metadata regarding licenses and copyright notices. They intend to relieve the license compliance burden for software projects and improve standardization for data and metadata transfer. This is relevant for:

- Individual developers, because it provides them a precise and easy-to-implement way to apply correct terms of license and copyright notices.
- Digital communities, because it improves how data and metadata for software re-usability is communicated.
- Academia, because it improves the re-usability of software in a safe and clear way in research projects.
- The public sector, because it fosters best practices for dealing with license and copyright notices, improves interoperability among agencies, and encourages open government.
- Commercial entities, because it allows them to optimize their software bill of materials and simplify development workflow.

4 REUSE SPECIFICATIONS

REUSE's core specifications are based on SPDX,[14] an open standard for communicating software bill of material information, including components, copyrights, licenses and security references. SPDX maintains a license list,[15] which defines standardized identifiers for a wide spectrum of commonly found licenses and exceptions used in FOSS, data, hardware, or documentation. SPDX was designed to provide "a common language and vocabulary to express security, licensing, and copyright information for products, components, packages, files and code snippets, enable tools to be created and facilitate the introduction of compliance automation." (Haddad, 2018, p. 154) The combination of these standards enables the REUSE's three-step compliance procedure:

[14] Software Package Data Exchange (SPDX) is an international ISO open standard (ISO/IEC 5962:2021) managed by the Linux Foundation.
[15] The SPDX License List includes a standardized short identifier, the full name, the license text, and a canonical permanent URL for each license and exception. Available at: https://spdx.org/licenses, retrieved on 30.06.22.

- Choosing and applying a license by downloading from the SPDX list the license text and information. This data should be stored in a *LICENSES* directory in the source code repository.
- *Providing to every single file a copyright and license header based on the SPDX standard:*
  - *# SPDX-FileCopyrightText: [year] [copyright holder] <[email address]>*
  - *# SPDX-License-Identifier: [identifier]*
- Confirming compliance with the REUSE toolkit,[16] which also is capable of automating the two previous steps.

5 RESOURCES AND SUPPORTERS

Software license compliance is a complex and vast area populated with a multitude of initiatives and tools to help compliance efforts. REUSE, which has a community-based approach, collaborates with several complementary projects[17], such as ClearlyDefined[18], OpenChain[19] and FOSSology[20]. In addition, for developers, there is a series of resources for easy engagement and adoptions, an open mailing list for discussion and deliberation, extensive FAQs, and a constantly updated toolkit with compliance tools, API checks, and provision for numerous CI/CD solutions.

Although it is not possible to know exactly the number of adopters, by February 2023, 1443 software repositories using the REUSE API are successfully implementing and following the best practices. REUSE had been adopted by the Linux Kernel, and several large companies. The specifications are also a central element in the compliance workflow for the European Commission's Next Generation Internet Initiative,[21] serving as consortium best practices for software and research projects developing human-centric technologies for the future of the Internet.

---

[16] See the tool's dedicated section on the REUSE website. Available at: https://reuse.software/faq/#tool, retrieved on 30.06.222.
[17] For an overview of complementary initiatives, see: https://reuse.software/comparison/, retrieved on 30.06.22.
[18] ClearlyDefined is an Open Source Initiative incubator project. The goals of the project are to collect and display meta and security information about a large number of software and data projects distributed on different package registries. It also motivates developers and curators to extend data about a project's licensing and copyright situation. REUSE in comparison concentrates on fixing the problem at the file level for individual projects. See: https://clearlydefined.io/about. Retrieved on 07.02.23.
[19] The OpenChain Project is focused on building trust in the free software supply chain. OpenChain focuses on making free software license compliance more transparent, predictable, and understandable for participants in the software supply chain. OpenChain recommends REUSE as one component to increase clarity of the licensing and copyright situation, but has higher requirements to achieve full conformance. See: https://www.openchainproject.org/, retrieved on 07.02.23.
[20] FOSSology is a toolkit for Free Software compliance, stores information in a database, and includes license, copyright and export scanners. It is more complex than REUSE and its helper tool and rather optimized for compliance officers and lawyers. REUSE instead intends to have all licensing and copyright information stored in or next to the source files to safeguard this information when reused elsewhere. See: https://www.fossology.org/about/, retrieved on 07.02.23.
[21] For a detailed overview of the initiative, see the NGI0 Zero website. Available at: https://www.ngi.eu/ngi-projects/ngi-zero/, retrieved on 30.06.22.

6 REFERENCES

1. Balter, Ben (2015). *Open source license usage on GitHub.com*. GitHub Blog. Retrieved on 26.06.2022. Available at: https://github.blog/2015-03-09-open-source-license-usage-on-github-com/
2. Blázquez, Cappello & Valais (2017). *Exceptions and limitations to copyright*. IRIS Plus, European Audiovisual Observatory, Strasbourg.
3. Haddad, Ibrahim (2018). *Open Source Compliance in the Enterprise*. 2nd ed. The Linux Foundation: San Francisco.
4. Ku Wei Bin, G., Lasota, L. and Jaeger, T. (2022). *Free and Open Source Software Licensing: Frequently Asked Questions - Next Generation Internet Legal To-Dos*. Berlin: Free Software Foundation Europe.
5. Open Source Initiative (2006). *Report of License Proliferation Committee and draft FAQ*. OSI Website. Retrieved on 26.06.2022. Available at: https://opensource.org/proliferation-report
6. Synopsys (2019). *Top open source licenses and legal risk for developers*. Retrieved on 30.06.22. Available at: https://www.synopsys.com/blogs/software-security/top-open-source-licenses/