On the Variability of the BSD and MIT Licenses

Trevor Maryka, Daniel German, Germán Poo-Caamaño

On the Variability of the BSD and MIT Licenses

Trevor Maryka and Daniel M. German and Germán Poo-Caamaño

University of Victoria, Canada
{tsmaryka,dmg,gpoo}@uvic.ca

Abstract. The MIT/X11 and the BSD are two of the most important family of Free and Open Source (FOSS) licenses. Because these licenses are to be inserted into the files that use it, and because they are expected to be changed by those who use them, their text has suffered alterations over time. Some of this variability is the result of licenses containing template fields which allow the license to be customized to include information such as the copyright holder name. Other variability can be attributed to changes in spelling, punctuation, and adding or removing conditions. This study empirically evaluated the extent that the BSD and MIT/X11 family of licenses are varied, and the manner and frequency in which license texts vary from the original definition. The study found that the BSD family has little variability, with a significant proportion fitting the common standard. The MIT/X11 family of licenses exhibited significantly more variation, with a higher propensity to customize the license text. In addition, the MIT/X11 license has spawned several specialized variants which likely constitute different legal meanings. Based on these findings, recommendations are proposed on what variability needs to be accommodated by the Software Package Data Exchange (SPDX) which is in the process of standardizing the allowed variability of both licenses.

1 Introduction

There exist many open sources licenses. The Open Source Initiate[1] (OSI), responsible for the definition of open source, has approved 70 licenses[2] The Software Package Data eXchange[3] (SPDX), a consortium of non-profit and profit organizations that attempt to standardize licensing information across parties lists 306 different licenses[4].

Some of the most important open source licenses are the family of BSD licenses, and the family of MIT licenses. These licenses comprise a very large portion of open source licensed software; in a study it was found that 9.1% of Debian applications were licensed under BSD or MIT licenses [2]. Furthermore, these licenses, which are known as Academic [6] are of particular interest because they allow unlimited use the software with very few restrictions[5].

[1] http://osi.org

[2] http://opensource.org/licenses/alphabetical

[3] https://spdx.org

[4] https://spdx.org/licenses

[5] The original BSD 4-clause license is an exception to this rule, unless the copyright owner is the University of California; for details, see https://spdx.org/licenses/BSD-4-Clause-UC

Both the BSD and MIT licenses are template licenses, because they have to be modified by the licensee to properly use them. The expected way these licenses are used is, first, by embedding the text of the license into each file; and second, by modifying the name of the copyright owner and related information in the file. Unfortunately, users of the license often further modify the text of license. It is a common practice of users of these licenses to replace some generic references to the copyright holders with (e.g. "the name of the organization" replaced with the name of the owner "ACME LTD"). In fact, the modifications made by users to the original BSD 4-Clauses license resulted into the BSD 3-Clauses and BSD 2-Clauses license (by dropping clauses from the original license).

This paper describes a empirical study of how the BSD and MIT licenses have been modified by its users. The contributions of this study include:: *a*) An empirical study documenting how the MIT and BSD are modified in practice in source code of Debian software packages; *b*) Analysis of this variability., and *c*) Recommendations for the SPDX Group on how to address the variability of these licenses in their future templates.. These results are being used by the SPDX Group to improve the templates of these licenses to match the way they are used in practice (without altering their legal meaning) and developers of tools that perform license identification.

2 Background and Related work

The Software Package Description Group is a consortium of for-profit and non-for-profit companies created under the auspices of the Linux Foundation. One of its primary objectives is the creation of the SPDX Standard [7,5]. The intent of the standard is to help in documenting and exchanging the license and copyright information of software components. The current version of the standard (v1.2) describes an easily parsable format to use to document what files are part of a component, their licenses and copyright owners, and the effective license of the component [7,5]. The SPDX Standard is expected to facilitate the exchange of this information and the creation of tools for software license compliance around it. This requires clear guidelines of how to identify and document licenses, specially Free and Open Source (FOSS) licenses. While most research has concentrated on the licenses documented by the Open Source Initiative (the so called 70 OSI-approved licenses), there exist many more licenses. As mentioned above, SPDX currently lists 306 licenses and has a mechanism to submit new licenses for consideration[6].

License identification of source code is challenging [3,2]. In FOSS each file is expected to contain a comment (usually at the top) that documents how it is licensed. We will refer to it as the *license statement* of the file. The license statement of a file is expected to include who the copyright owner is and how the file is licensed. There are two methods in which a file documents its license: by-inclusion and by-reference. By-inclusion refers to a license that it used by including its text in the license statement of the file. For example, files that are licensed using the BSD and MIT licenses usually include the full text of the license in their license statements.

[6] http://spdx.org/spdx-license-list/request-new-license

On the other hand, by-reference corresponds to licenses that are not included in the license statement of the file; instead, a link to the license is provided. For example, the Apache-2.0[7] is expected to be used by-reference; the Appendix *How to apply the Apache License to your work* indicates how to add such reference (e.g. by including the following text: *Licensed under the Apache License, Version 2.0 (the "License"); [...] You may obtain a copy of the License at: http://www.apache.org/licenses/LICENSE-2.0*).

Many licenses that are used by-inclusion are expected to be customized by their users. For example, the original BSD license (BSD-4-Clause-UC[8]—which has the Regents of the University of California as its copyright owner) was converted into a template license where the text of the copyright owner is expected to be filled by the user (thus becoming the BSD-4-Clause). For example, the following paragraph illustrates the variable text (between { }, bold use for emphasis) that should be replaced when the license is used[9].

> IN NO EVENT SHALL **{{COPYRIGHT HOLDER}}** BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES

By-inclusion template licenses should be modified only in the form indicated. Unfortunately it is often the case that users of the licenses modify the text further. As described in [2], it was found that licensors often change the spelling of licenses (English vs British), change grammar (add or remove punctuation to clarify intent), and in some cases, add or remove extra clauses of known licenses to create new licenses.

Some of these changes result in licenses that are so frequently used that they deserve to receive a name. As mentioned above, the original BSD-4-Clause-UC was converted to a template license and called BSD-4-Clause (the user must *fill* the template with the name of the corresponding copyright owner, replacing the "Regents of the University of California" found in the original BSD-4-Clause-UC). The BSD-4-Clause is shown in Figure 1 (as documented by SPDX).

In other cases clauses were removed, resulting into the BSD-3-Clause and the BSD-2-Clause. In the case of the BSD-2-Clause, when the copyright owner is the NetBSD Foundation, Inc. the license is known as BSD-2-Clause-NetBSD (this is no longer a template license). Apache-1.0 is a derivative of the BSD-4-Clause: it reused its four clauses and the liability and warranty terms; furthermore, among many other changes, the Apache-1.0 fixed a spelling mistake of the BSD-4-Clause ("EXPRESSED" instead of "EXPRESS") and added "ITS" in "OR ITS CONTRIBUTORS".

These types of license customizations have created a problem for SPDX. The goal of SPDX is to document licenses and their use. This includes documenting the templates and how they are expected to be modified by their users. Unfortunately the original SPDX templates of these licenses have used different ways to document this variability. For example, SPDX documents the BSD-4-Clause as a template license, as shown in Figure 1. At it can be seen SPDX uses both <> and {{}} to document the sections that must be modified (shown in bold). Even though a template might not allow, a license is frequently modified. For example, in the BSD licenses it is common to see

[7] In this paper we refer to FOSS licenses by their the SPDX standardized names.

[8] For the remaining of this paper, we will refer to licenses by their SPDX name, see spdx.org/licenses/

[9] http://spdx.org/licenses/BSD-4-Clause

```
1  Copyright (c) <year>, <copyright holder>
2  All rights reserved.
3  Redistribution and use in source and binary forms, with or without modification,
4  are permitted provided that the following conditions are met:
5  1) Redistributions of source code must retain the above copyright notice, this
6     list of conditions and the following disclaimer.
7  2) Redistributions in binary form must reproduce the above copyright notice, this
8     list of conditions and the following disclaimer in the documentation and/or
9      other materials provided with the distribution.
10 3) All advertising materials mentioning features or use of this software must
11    display the following acknowledgement:
12      This product includes software developed by the organization.
13 4) Neither the name of the organization nor the names of its contributors may be
14    used to endorse or promote products derived from this software without specific
15    prior written permission.
16 THIS SOFTWARE IS PROVIDED BY COPYRIGHT HOLDER ''AS IS'' AND ANY EXPRESS OR IMPLIED
17 WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF
18 MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT
19 SHALL COPYRIGHT HOLDER BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL,
20 SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO,
21 PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR
22 BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN
23 CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN
24 ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED
25 OF THE POSSIBILITY OF SUCH DAMAGE.
```

Fig. 1. SPDX specification of the BSD-4 Clauses. The variable sections are depicted in bold.

COPYRIGHT HOLDER replaced with the actual name of the copyright holder (line 16 in Figure 1.

In SPDX Version 1.2 this variability has been accounted for via a template language. It documents the variable sections of a license using regular expressions. It indicates where and how a license can be modified and still be considered the same license. SPDX is now in the stage of updating the templates for these licenses[10]. The challenge is not longer how to document the variability, but what variability should be documented as permitted without changing the meaning of the license.

3 Research Questions

The two most common template licenses in open source are the family of BSD licenses (BSD-4-Clause, BSD-3-Clause, and BSD-2-Clause, etc.) and the family of MIT licenses (MIT, X11, MIT-CMU, MIT-advertising, etc). In Debian 5.0 these licenses account for more than 5% of all the files, and more than 8% of the applications that in which all their source files were licensed under the same license [2].

This study attempts to answer two fundamental questions for these two families of licenses. Firstly, what is the degree of variation of a license? In other words, if we imagine the original license text as the root of a tree, what branches of variation in the text have been created? Secondly, what is the frequency that variation occurs at? In particular, can we find the common patterns of variability used, and separate these from the less commonly used variants?

[10] See https://github.com/dmgerman/spdxTemplates for the current state of these templates (not yet approved by SPDX); for example the template for the BSD-4-Clause can be found at https://github.com/dmgerman/spdxTemplates/blob/master/bsd-4-clause/bsd-4-clause.txt

4 Methodology

The subject of this study is the Debian Linux Distribution, version 6. Ninka was run on all the files of this distribution, encompassing more than 1.3 million source files contained in 10,014 projects. Ninka identified 42,653 licenses in the BSD family, and 28,205 licenses in the MIT/X11 family. The variability analysis of these licenses worked in a top-down manner, by first identifying the largest, most commonly used variations, working down to smaller, less commonly used variations to the point where further analysis of variability would be negligible.

Ninka uses a pipeline architecture, and divides license identification into several stages. Two of them were of value to this research: sentence matching, and license matching. Sentence matching corresponds to the process of matching a given sentence to a known valid licensing sentence (irrespective of the license it belongs to). Even though the entire license might not match a known license, one of more of its sentences might. This allows to identify sentences of BSD and MIT licenses in cases where the entire license didn't match a known license. When Ninka matches a sentence, it outputs the known sentence, and what (if any) known variability it exhibited.

Enumeration and analysis of variability was done in a top down manner, in three levels. Initially the exact license that Ninka identified, such as the strict SPDX BSD2 vs. the less strict BSD2 (as named by Ninka), was enumerated. This level will be designated "license level variability". Secondly, the variability of the construction of sentence tokens that comprise a license was enumerated. An example of this would be a non-SPDX license, such as BSD2, must have one or more individual sentences that are not the strict version. Multiple combinations of strict and non-strict sentences are possible. The combination of sentence tokens can be referred to as the "token signature", and this level of variability will be designated "token signature variability". Finally, the content of each variable section within a sentence token was extracted from the sentence token file and enumerated. This level or variability will be designate

5 Results

5.1 The BSD Family

We use a top-down approach to describe the variability of the BSD family of licenses. Figure 2 depicts this variability.

BSD-4 Ninka identified 3,251 licenses in Debian 6 as variants of the BSD-4-Clauses (8% of all BSDs). Of them, 1,887 licenses were identified as exactly the SPDX BSD-4-Clause. The rest 1,370 showed mostly small variations. The most common of these changes (1,179–86%) were due to changes in the non-endorsement clause statement (Lines 13-15 in Figure 1): 457 licenses replaced "nor the names of its contributors" with specific names (e.g. "the University nor of the Laboratory") or remove it altogether. Another 220 licenses alter alter the text "contributors" to "co contributors" or "co-contributors". A further 398 licenses had the "Neither" removed from this clause. A very small number licenses (75) modified the Clause-3 by replacing the word "by"

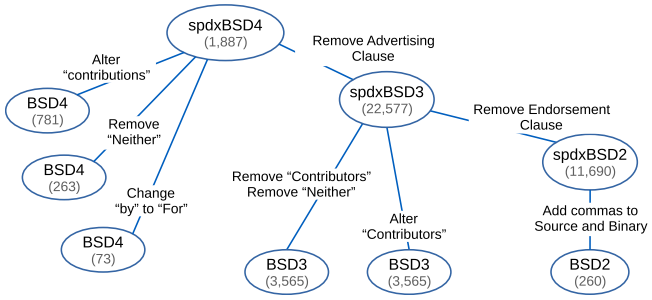

Fig. 2. BSD Variability Tree

with “for” (Lines 11-13 in Figure 1). All other variability of the BSD-4 license were negligible.

BSD-3 Ninka identified 22,577 licenses as the SPDX BSD-3-Clauses (this was the most common). A further 4,822 had small changes that can be considered still within the spirit of the BSD. As with the BSD-4, most changes were due to to variations in the Non-Endorsement clause (Clause 3): 4,579 (91%). In 3,702 “Neither” was removed (in 96% of these cases the “contributors” were removed too). Another 887 licenses had the text “its contributors” changed to “his contributors”, “other contributors”, or “any contributors”. The remaining 61 licenses indicate specific contributors (as the BSD-4 mentioned above, e.g. “the University nor of the Laboratory”). The rest of the variants were negligible. Also, 887 (18%) non-endorsement clause changes are due to a change the text “nor the names of its contributors”. 826 (93%) of these use some small variant on the qualifier before “contributors” such as “his contributors”, “other contributors”, or “any contributors”. The remaining 61 cite specific contributors (“the University nor of the Laboratory”). All other variability of the BSD-3 license was ignored because its incidence represented less than 2%.

BSD-2 Ninka identified 11,690 licenses as SPDX BSD-2. We found very little variability in licenses that can still be considered BSD-2 (only 458 licenses). 295 had variability in the Redistributions of Source Condition and 260 in the Redistributions of Binaries Condition. In most cases the changes are minor: 287 added a comma after the word “conditions”, 268 added “above” before “copyright notice” (in both conditions). The remaining changes were negligible.

5.2 The MIT/X11 Family

The MIT/X11 family of licenses is considerably more fragmented the BSD family. Ninka divides the MIT/X11 family into eleven separate licenses, compared to the 3 documented by SPDX.

The most frequent license identified by Ninka is the MIT as documented by SPDX: 10,219 (36% of all MIT/X11 family). The tree of MIT/X11’s significant variability, and the frequency of each variant is show in Figure 2.

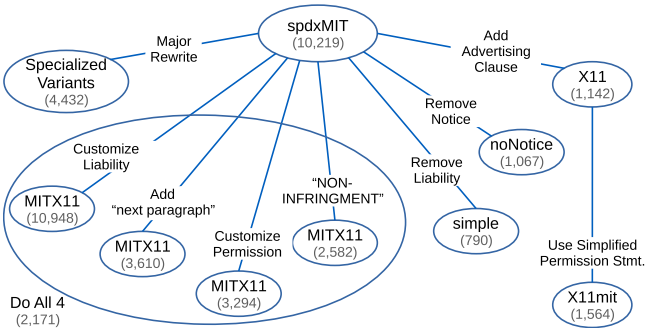

Fig. 3. MIT/X11 Variability Tree

Of the 17,986 (64%) licenses that are not the strict SPDX version, Ninka identifies five types of licenses which are closely derived from the SPDX definition: X11, X11mit, MITX11, MITX11noNotice, and MITX11simple. These non-specialized licenses account for 13,554 (75%) non-strict licenses.

Ninka also identifies five types of licenses which are much more specialized variants of the MIT/X11 license: MITold, BindMITX11Var, MIToldNoSellNoDoc, X11Festival, MITNoSellNoDocBSDvar (these are names that Ninka has given to each of these variants), and the SPDX MITNFA. 4,432 (25%) of the non-SPDX licenses are identified in this specialized category, and each license arguably represents a more significant re-writing of the strict license, rather than a smaller variation (with MITold likely being the a predecessor rather than a re-writing).

The non-specialized variant that is most closely derived from the SPDX version is the license Ninka identifies as MITX11. This license comprises 8,991 (66%) of the non-SPDX, non-specialized licenses. These MITX11 licenses are similar to the SPDX standard, but with at least one sentence changed.

Ninka identifies 1,142 (8%) licenses as X11 (Ninka's X11 is different from SPDX X11—the latter is not a template license and can only be used by the X11 Consortium). This X11 license is the SPDX MIT with an additional advertising material clause:

> Except as contained in this notice, the name of the **{{author}}** shall not be used in advertising or otherwise to promote the sale, use or other dealings in this Software without prior written authorization from **{{author}}**.

Ninka's X11mit accounts for 1,564 (12%) licenses and it is a variation of Ninka's X11 license with an old style, simplified permission statement:

> Permission to use, copy, modify, distribute, and sell this software and its documentation for any purpose is hereby granted without fee, provided that the above copyright notice appear in all copies and that both that copyright notice and this permission notice appear in supporting documentation.

MITX11noNotice accounts for 1,067 (8%) licenses, and corresponds to the SPDX MIT license with the notice statement removed:

> The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

MITX11simple accounts for 790 (6%) licenses and is a variant of the SPDX MIT with the liability statement removed:

> IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE [...]

Variability in Sentences Within the MIT family of licenses, we found that many variants shared the same type of variability at the sentence level.

Liability Statement. The most commonly varied statement is the liability disclaimer sentence. 10,948 (80%) of the non-strict, non-specialized licenses alter the text "AUTHORS OR COPYRIGHT HOLDERS" to either cite specific organizations/authors (67%), or to vary it in some fashion (33%), such as changing it to "THE ABOVE COPYRIGHT HOLDERS".

Notice Statement. The second most commonly variant is the notice statement. 3,610 (26%) of the non-strict, non-specialized licenses vary the text of the notice statement:

> The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

These licenses add the text "(including the next paragraph)" after the text "permission notice", implying that the warranty statement must also be included.

Permission Statement. The third most common variant is in the permission statement. 3,294 (24%) of the non-strict, non-specialized licenses vary the permission statement *Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"),...* in the following manner: 2,445 (74%) of licenses split "sublicense" to "sub license"; 655 (20%) licenses add an "on" before "limitation"; 582 (18%) licenses add a ", distribute without modifications" after "distribute".

"As Is" Statement. Finally, the fourth most commonly varied statement in the MIT/X11 family is the "As Is" warranty disclaimer sentence. This variation simply changes the word "NONINFRINGMENT" to "NON-INFRIGNMENT" (adding a dash). 2,852 (21%) of non-specialized, non-strict licenses add this dash. A significant proportion (2,171) of non-strict, non-specialized licenses use the non-strict versions of the permission, notice, "As Is", and liability statements at the same time.

6 Analysis

The SPDX version of each BSD license is exactly equivalent to the BSD licenses found in Debian 6 for the vast majority (85%) of licensing statements. The most commonly changed sentence is the non-endorsement clause found in BSD-3 and BSD-4. It is likely that this statement is most commonly varied because it contains a template field. The requirement to customize the template field likely increases the propensity to make other changes to the sentence, such as the removal of the word "Neither" and the customization of the "nor the names of its contributors" phrase.

The MIT/X11 family of license exhibits significantly more variability, with the majority of licenses not exactly matching the SPDX version of the license. The most significant and common variation of the license family is for the licensor to treat the text "AUTHORS OR COPYRIGHT HOLDERS" as a template field, and to customize it. Licensors also add conditions and grants such as "(including the next paragraph)", or "distribute without modifications", which may or may not alter the legal meaning of the license. Additionally, licensors adjust the spelling of words such as "sublicense" and "noninfringment", by adding dashes or spaces.

7 Recommendations

The BSD family of licenses very closely matches the SPDX standard, so little if any changes are required to the SPDX license list to accommodate its variability. If the word "Neither" was considered optional in the non-endorsement clause, an additional 4,100 Debian 6 licenses could directly match the SPDX, but that only represents 1% of the BSD family so this is likely a negligible variation.

The MIT/X11 family of licenses has significantly more variability, so small changes to accommodate potentially varied license texts would be beneficial to the SPDX standard, as it would increase the number of licenses that can accurately be included by users of the standard.

MIT/X11 template. The most significant variation that SPDX could accommodate within the MIT/X11 family would be the alteration of the text "AUTHORS OR COPYRIGHT HOLDERS", in the liability statement, to be the template field "{{authors or copyright holders}}". This text has been treated as a de facto customizable field by licensors, despite the fact that this text was not intended to be customized. Accommodating this customizable field would increase the number of SPDX equivalent licenses in Debian 6 by at least 4,110, which is a 40% increase.

Equivalent Phrases. Much like the set of equivalent phrases used in the text normalization phase of Ninka, SPDX has published a small list of spellings which are considered legally equivalent [7]. The following recommended equivalent sets should be added to this list to accommodate variability found in BSD and MIT/X11 licenses:

- *"contributors"*, *"co-contributors"*, *"co contributors"*
- *"sublicense"*, *"sub-license"*, *"sub license"*
- *"noninfringement"*, *"non-infringement"*

The inclusion of these equivalent phrases would accommodate 5,517 occurrences of variability found. Note that this figure is greater than the number of individual licenses that can be accommodated, as the changes to "sublicense" and "noninfringment" may not be mutually exclusive.

MIT/X11 Additions. Adding the grant "distribute without modifications" to the list of permissions granted by the license may be legally redundant. If the licensee is permitted to redistribute a modified copy of software, then theoretically the licensee could redistribute a modified copy which is functionally equivalent to the original. As such, distributing the functionally equivalent modified copy may be equivalent to distributing the software without modifications, so the additional grant may be redundant. This implies that adding this clause constitute a license error and should be removed.

Derived Licenses. Some MIT/X11 licenses, like the specialized variants Ninka recognizes, and the variants that add the advertising clause, remove the notice clause, or replace the permission statement represent different legal meanings from the strict SPDX MIT license. These licenses need to be individually added to the SPDX license list to be accommodated by the standard. The process to add a license to the list is expensive [1], so these should be accommodated in a top down manner, with more frequently proposed/requested licenses added first. However, this is not currently the case. Licenses are currently considered for addition in the order in which they are proposed.

8 Threats To Validity

We identify the following threats to validity of this empirical study. Regarding External validity. The use of Debian 6 as a data source is likely representative of the BSD and MIT/X11 license families, as Debian offers a diverse and comprehensive view of FOSS landscape [4]. Additionally, the accuracy of Ninka would play a significant role in the external validity of this study, since Ninka has been externally verified to have a high accuracy of 96.6% [2].

Regarding Reliability validity. This study can be replicated. Ninka, the is freely available for anyone to use, including the source code files containing the sentence token expressions and token matching rules. Additionally, the scripts used to extract sentence token signatures and sentence token variability, and the spreadsheet output of each sentence token analyzed have been made available for replication at: http://turingmachine.org/2015/mit-bsd.

9 Conclusion

Licensors change the text of standard open source licenses for many purposes, including customizing the license with their specific organization's name, adding or removing conditions, and changing spelling or punctuation. Open source licensing standards like SPDX may be affected by the variability in licenses, as the variability may alter the legal meaning of licenses, creating legal issues in matching an altered license. In contrast, requiring an overly strict "perfect match" of open source licenses to the standard may result in the exclusion of many license texts with negligible variability. This paper presented an empirical study of the extent that the BSD and MIT/X11 family of licenses vary from their original definition. The BSD family of licenses closely match the existing SPDX templates, with little additional variability. The MIT/X11 family of licenses was found to be much more fragmented and heavily customized, including the creation of several specialized variants based from the original X11 license, customization of the text "authors or copyright holders", spelling alterations, and the adding and removing of conditions, grants and whole sentences. Small changes to the SPDX template for the MIT license, and to the SPDX list of equivalent words [7] would accommodate some the essential variation found within the license at a low cost.

References

1. German, D., Penta, M.D.: A method for open source license compliance of java applications. IEEE Software 29(3), 58–63 (2012)
2. German, D.M., Manabe, Y., Inoue, K.: A sentence-matching method for automatic license identification of source code files. In: 25nd IEEE/ACM International Conference on Automated Software Engineering (ASE 2010) (2010)
3. Gobeille, R.: The FOSSology project. In: MSR '08: Proceedings of the 2008 International Conference on Mining Software Repositories. pp. 47–50. ACM, New York, NY, USA (2008)
4. Gonzalez-Barahona, J.M., Robles, G., Michlmayr, M., Amor, J.J., German, D.M.: Macro-level software evolution: a case study of a large software compilation. Empirical Software Engineering 14(3), 262–285 (June 2009)
5. Lovejoy, J., Odence, P., Lamons, S.: Advancing the Software Package Data Exchange: An update. International Free and Open Source Software Law Review 2(2), 145–152 (2013)
6. Rosen, L.: Open Source Licensing: Software Freedom and Intellectual Property Law. Prentice Hall (2004)
7. Stewart, K., Odence, P., Rockett, E.: Software Package Data Exchange (SPDX) Specification. International Free and Open Source Software Law Review 2(2), 191–196 (2010)