How to curate license information for



Caren Kresse Open Source Automation Development Lab (OSADL) eG





Why re-using compliance data?

- The success of FOSS is partly due to the fact that development resources and efforts are reduced by reusing existing software components.
- Some of the reduction is canceled out by the effort required to scan and clear FOSS for license compliant use.
- There are many FOSS components that are deployed unmodified by a large number of users.
- Why not **share compliance tasks** in the same way as the software development?





The OSADL project: OSSelot The Open Source Curation Database

https://www.osselot.org https://github.com/Open-Source-Compliance/package-analysis

- Contains license and copyright analysis results for various packages (> 340 (>200 unique) packages and 1.5 Mio. files):
 - **Readme with metadata** of the package, *e.g.* download location, purl, reviews, comments.
 - **SPDX** Tag:Value, JSON, RDF-XML and YAML files with concluded licenses, copyright notices and comments on decisions per file.
 - Disclosure document with aggregated license texts, copyright notices and acknowledgments.





Liability and trust

Supply of legal information

Reuse of legal information









Liability and trust

Limiting liability

- Licensing to give maximal rights
 (e.g. CCO-1.0). Then, gift
 regulations apply and liability
 applies only for willful intent and
 gross negligence.
- There are no known lawsuits
 pursuant to supply of incorrect
 legal information for FOSS so
 that reservations have generally
 lessened.

Establishing trust

- Data are curated **conscientiously** and **diligently**.
- Compliance artifacts are created by named persons with expertise.
- There is a transparent review process.
- An organization stands with its reputation for the quality of the project.





Liability and trust

Limiting liability

- Licensing to give maximal rights
 (e.g. CCO-1.0). Then, gift
 regulations apply and liability
 applies only for willful intent and
 gross negligence.
- There are no known lawsuits
 pursuant to supply of incorrect
 legal information for FOSS so
 that reservations have generally
 lessened.

Establishing trust

- Data are curated conscientiously and diligently.
- Compliance artifacts are created by named persons with expertise.
- There is a transparent review process.
- An organization stands with its reputation for the quality of the project.





OSSelot curation guidelines: General

- Scanning and curating currently with **Fossology** with integrated **Scancode**.
- Download of source code **as upstream as possible**, e.g. directly from the project page.
- Curating license findings and copyright findings manually.





OSSelot curation guidelines: Copyrights

- Findings that were **incorrectly identified** as a copyright statement (*e.g.* license texts, code, *etc.*) are removed.
- Content that is not part of the copyright notice (*e.g.* formatting signs, license notices, comments on content, code, *etc.*) is removed from the copyright notice.
- If the source code tree contains an **AUTHORS** file, the content of this is given as value to the SPDX tag "PackageCopyrightText".





Curating copyright notices

FOSSology Findings

ScanCode Findings

User Findings

Activated FOSSology Findings statements:

Show 25 ✓ entries			
FOSSology Findings			
copyright::year_of(\$0); print <<"EOF";	deactivated [Undo]		
COPYRIGHT_YEAR'} if defined \$ENV{'OSSL_COPYRIGHT_YEAR'};	deactivated [Undo]		
copyright the software, and 2) offer you this license which gives you legal permission to copy, distribute and/or modify the software.			
Copyright .* The OpenSSL Project Authors/;	×		
Copyright text was shrunk to a boilerplate that points to the license shared" builds are now the default when possible Added support for "pipelining" Added the AFALG engine New threading API implemented Support for ChaCha20 and Poly1305 added to libcrypto and libssl Support for extended master secret CCM ciphersuites Reworked test suite, now based on perl, Test::Harness and Test::More Most* libcrypto and libssl public structures were made opaque, including: BIGNUM and associated types, EC_KEY and EC_KEY_METHOD, DH and DH_METHOD, DSA and DSA_METHOD, RSA and RSA_METHOD, BIO and BIO_METHOD, EVP_MD_CTX, EVP_MD, EVP_CIPHER_CTX, EVP_CIPHER_CTX, EVP_CIPHER, EVP_PKEY and associated types, HMAC_CTX, X509, X509_CRL, X509_OBJECT, X509_STORE_CTX, X509_STORE, X509_LOOKUP, X509_LOOKUP_METHOD libssl internal structures made opaque SSLv2 support removed Kerberos ciphersuite support removed RC4 removed from DEFAULT ciphersuites in libs			
1 copyright: stringz "Montgomery multiplication for IA-64, CRYPTOGAMS by <appro\@openssl.org>"</appro\@openssl.org>			
1 Copyright Siemens AG 2020			
Copyright Siemens AG 2019-2020	×		
CONVIGINT DURDOSS IT GOES DOT DECESSARILY LIST EVERYONE WHO DAS CONTRIDUTED CODE			
Copyright Patrick Powell 1995 This code is based on code written by Patrick Powell <pre><pre>papowell@astart.com></pre></pre>	×		
copyright, patent, trademark, and attribution notices from the Source form of the Work, excluding those notices that do not pertain to any part of the Derivative Works; and	deactivated [Undo]		
	FOSSology Findings copyright::year_of(\$0); print <<"EOF"; COPYRIGHT_YEAR} if defined \$ENV{OSSL_COPYRIGHT_YEAR}; copyright the software, and 2) offer you this license which gives you legal permission to copy, distribute and/or modify the software. Copyright the software, and 2) offer you this license which gives you legal permission to copy, distribute and/or modify the software. Copyright text was shrunk to a boilerplate that points to the license shared" builds are now the default when possible Added support for "pipelining" Added the AFALG engine New threading API implemented Support for ChaCha20 and Poly1305 added to liborypto and libss! Support for extended master secret CCM ciphersuites Reworked test suite, now based on perl, Test::Harness and Test::More Most* liborypto and libss! public structures were made opaque, including: BIGNUM and associated types, EC_KEY and EC_KEY_METHOD, DH and DH_METHOD, DSA and DSA_METHOD, BIO and BIO_METHOD, EVP_MD_CTX, EVP_CIPHER_CTX,	FOSSology Findings copyright::year_of(\$0); print <<"EOF"; deactivated [Undo] COPYRIGHT_YEAR} if defined \$ENV(*OSSL_COPYRIGHT_YEAR*); copyright the software, and 2) offer you this license which gives you legal permission to copy, distribute and/or modify the software. Copyright the software, and 2) offer you this license which gives you legal permission to copy, distribute and/or modify the software. Copyright text was shrunk to a boilerplate that points to the license shared" builds are now the default when possible Added support for "pipelining" Added the AFALG engine New threading API implemented Support for Chacha20 and Poly1305 added to libcrypto and libss! Support for extended master secret CCM ciphersuites Reworked test suite, now based on perf. Test:Harness and Test:More Most* libcrypto and libss! Support for extended master secret CCM ciphersuites Reworked test suite, now based on perf. Test:Harness and Test:More Most* libcrypto and libss! Support for extended master secret CCM ciphersuites Reworked test suite, now based on perf. Test:Harness and Test:More Most* libcrypto and libss! Dublic structures were made appaque, including: BIGNUM and associated types, EC KEY and associated types, EC KEY and the sociated types, EC KEY and associated types, EM MAC CTX, X509, X509 CAL, X509, S509 CRL, X509 STORE, X509 STORE, X509 LOOKUP, X509 LOOK	

Curating copyright notices per file

Software Heritage | License Browser | File Browser • Conf | Info | View | Licenses | Copyright/Email/Url/Author | ECC | keyword | Bucket | Spasht • Hex | Text | Formation

Hide Legend

- * Original BSD copyright notice is retained at the end of this file.
- * Copyright (c) 1989, 1991, 1993, 1994
- The Regents of the University of California. All rights reserved.
- * Copyright (c) 1997-2005 Herbert Xu <<mark>herbert@gondor.apana.org.au</mark>> * was re-ported from NetBSD and debianized.
- * rewrite arith.y to micro stack based cryptic algorithm by
- * Copyright (c) 2001 Aaron Lehmann <aaronl@vitelus.com>
- Modified by Paul Mundt <lethal@linux-sh.org> (c) 2004 to support
 dynamic variables.
- * Modified by Vladimir Oleynik <dzo@simtreas.ru> (c) 2001-2005 to be
- used in busybox and size optimizations,
- * rewrote arith (see notes to this), added locale support,
- rewrote dynamic variables.
- * Licensed under GPLv2 or later, see file LICENSE in this source tree.
- /* Copyright (c) 2001 Aaron Lehmann <aaronl@vitelus.com>
- * Permission is hereby granted, free of charge, to any person obtaining
- * a copy of this software and associated documentation files (the
- * "Software"), to deal in the Software without restriction, including

Legend
file text
copyright remark
URL
e-mail address
author or maintainer

Count 🔻	FOSSology	
2	Copyright (c) 2001 Aaron Lehmann <aaronl@vitelus.com></aaronl@vitelus.com>	×
2	Copyright (c) 1989, 1991, 1993, 1994 The Regents of the University of California. All rights reserved.	×
1	Modified by Paul Mundt < lethal@linux-sh.org > (c) 2004 to support dynamic variables.	×
1	copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution. 3. Neither the name of the University nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.	
1	copyright notice, this list of conditions and the following disclaimer.	×
1	Copyright (c) 1997-2005 Herbert Xu <herbert@gondor.apana.org.au> was re-ported from NetBSD and debianized.</herbert@gondor.apana.org.au>	×
1	(C) 2003 Vladimir Oleynik <dzo@simtreas.ru></dzo@simtreas.ru>	×
1	(c) 2001-2005 to be used in busybox and size optimizations, rewrote arith (see notes to this), added locale support, rewrote dynamic variables.	×

Showing 1 to 8 of 8 entries

Page

Curating copyright notices per file

Software Heritage | License Browser | File Browser • Conf | Info | View | Licenses | Copyright/Email/Url/Author | ECC | keyword | Bucket | Spasht • Hex | Text | Formation

Hide Legend	Count _	FOSSology	
* Original BCD converient nation is retained at the and of this file	2	Copyright (c) 2001 Aaron Lehmann <aaronl@vitelus.com></aaronl@vitelus.com>	×
* Original BSD copyright notice is retained at the end of this file. * Copyright (c) 1989, 1991, 1993, 1994	2	Copyright (c) 1989, 1991, 1993, 1994 The Regents of the University of California. All rights reserved.	×
* The Regents of the University of California. All rights reserved.	1	Modified by Paul Mundt <lethal@linux-sh.org> (c) 2004 to support dynamic variables.</lethal@linux-sh.org>	×
* Copyright (c) 1997-2005 Herbert Xu <herbert@gondor.apana.org.au> * was re-ported from NetBSD and debianized. * rewrite arith.y to micro stack based cryptic algorithm by * Copyright (c) 2001 Aaron Lehmann <aaronl@vitelus.com> * * Modified by Paul Mundt <lethal@linux-sh.org> (c) 2004 to support</lethal@linux-sh.org></aaronl@vitelus.com></herbert@gondor.apana.org.au>	1	copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution. 3. Neither the name of the University nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.	×
* Modified by Vladimir Oleynik <dzo@simtreas.ru> (c) 2001-2005</dzo@simtreas.ru>	to be 1	copyright notice, this list of conditions and the following disclaimer.	×
* used in busybox and size optimizations,* rewrote arith (see notes to this), added locale support,		Copyright (c) 1997-2005 Herbert Xu <herbert@gondor.apana.org.au> was re-ported from NetBSD and debianized.</herbert@gondor.apana.org.au>	×
* rewrote dynamic variables.	1	(C) 2003 Vladimir Oleynik <dzo@simtreas.ru></dzo@simtreas.ru>	×

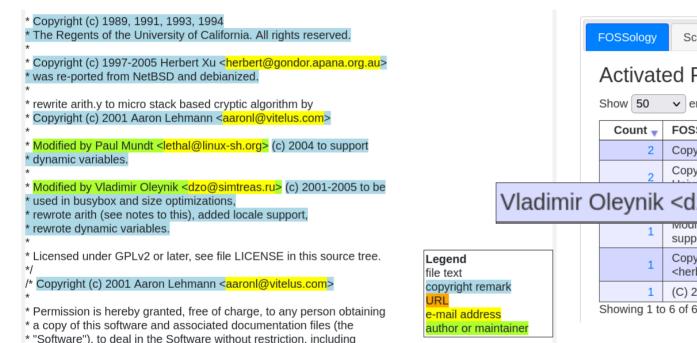
/* Copyright (c) 2001 Aaron Lehmann <aaronl@vitelus.com>

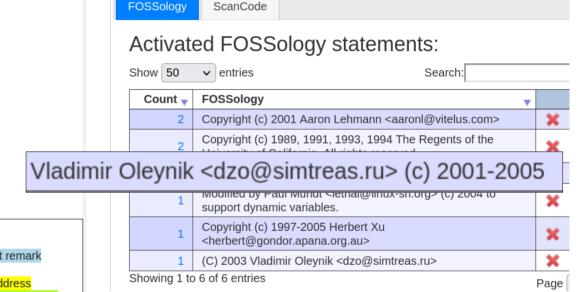
* Permission is hereby granted, free of charge, to any person obtaining * a copy of this software and associated documentation files (the

* "Software"), to deal in the Software without restriction, including

(c) 2001-2005 to be used in busybox and size optimizations, rewrote arith (see notes to this), added locale support, rewrote dynamic variables.

Curating copyright notices per file









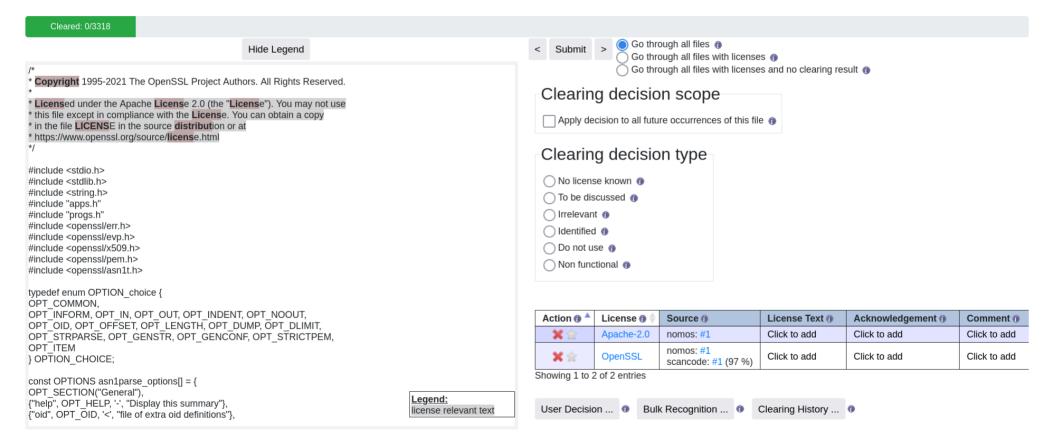
OSSelot curation guidelines: Licenses (1)

- Review is done on **file level**, *i.e.* every file in the source code tree for which at least one scanner found a result is analyzed and **confirmed or corrected**.
 - → individually or via Bulk Recognition

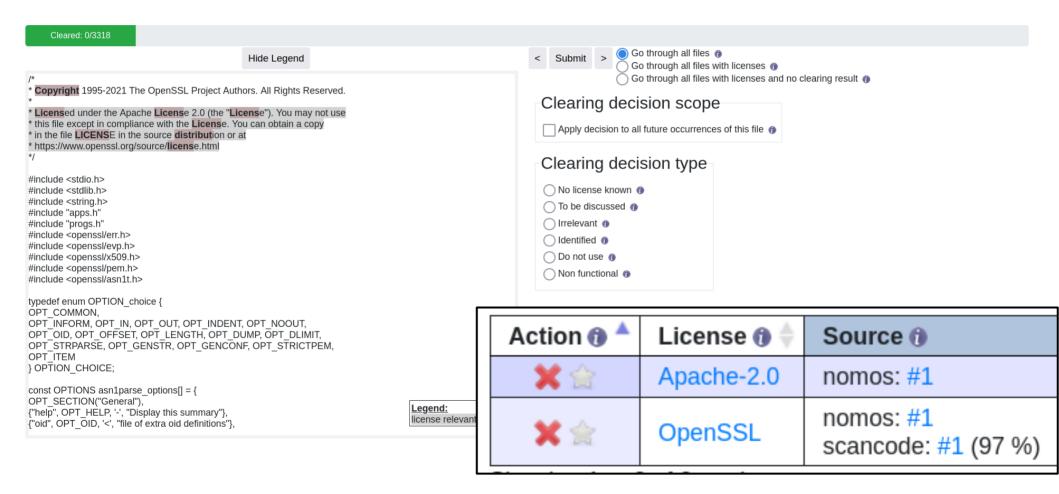




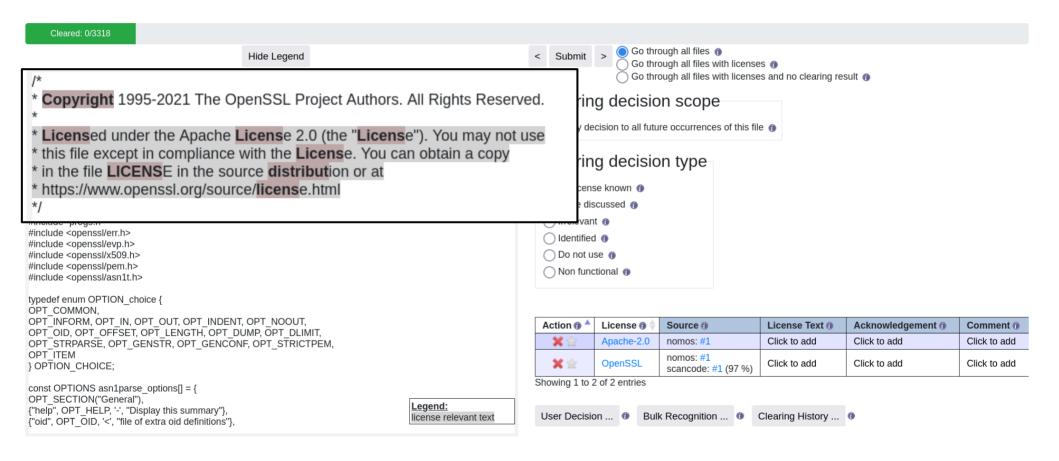
Curation



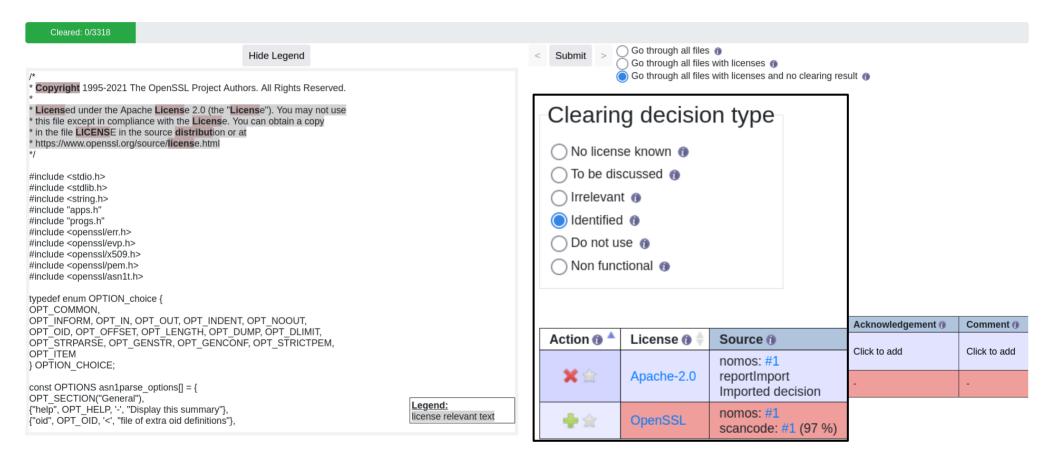
Curation: Scanner findings



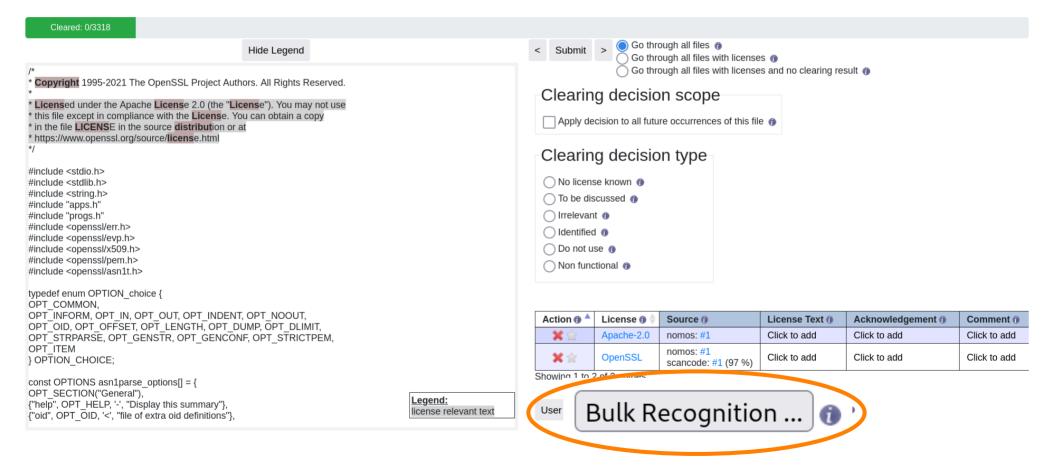
Curation: License relevant text



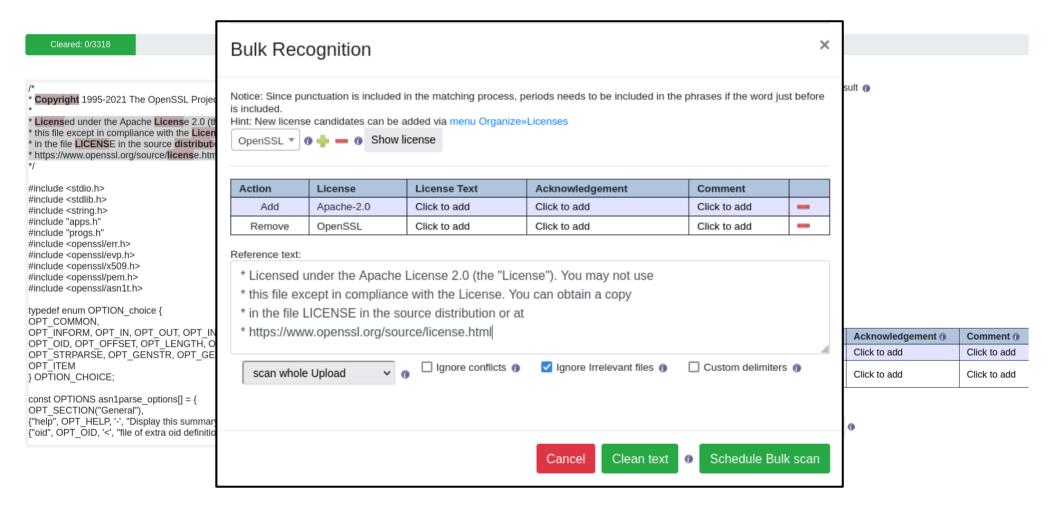
Curation: Decision



Curation: Bulk Recognition



Curation: Bulk Recognition



OSSelot curation guidelines: Main license

- Only if there is a main license for the package, e.g. LICENSE / COPYING / README file in root directory → given as value to the SPDX tag
 "PackageLicenseDeclared".
- This license is not automatically valid for all files and there is usually more than this one license in a package.

busybox-1.36.1.tar.bz2/busybox-1.36.1.tar/busybox-1.36.1/LICENSE

Software Heritage | License Browser | File Browser • Conf | Info | View | Licenses | Copyright/Email/Url/Author | ECC | keyword | Bucket | Spasht • Hex | Text | For

Cleared: 0/839

Hide Legend

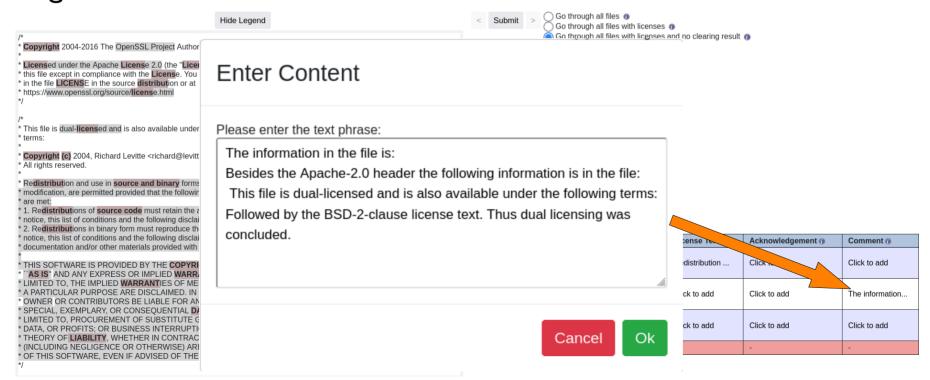
--- A note on GPL versions

BusyBox is **distribut**ed under version 2 of the General Public **Licens**e (included

in its entirety, below). Version 2 is the only version of this **licens**e which this version of BusyBox (or modified versions derived from this one) may be **distribut**ed under.



• In case a license conclusion is not obvious, the decision is explained in the "Comments" section which maps to the SPDX tag "LicenseComments".



• License comments are made with the following heuristic:

```
The information in the file is:
"[Quote licensing information in the source code file]"
[Give reason for conclusion] Therefore, [license] is concluded.
```





• License comments are made with the following heuristic:

```
The information in the file is:
"[Quote licensing information in the source code file]"
[Give reason for conclusion] Therefore, [license] is concluded.
```

• Example 1: No version of the license is given

```
The information in the file is:
"This file is GPL'd."
As no version of the GPL is given, GPL-1.0-or-later is concluded.
```





• License comments are made with the following heuristic:

```
The information in the file is:
"[Quote licensing information in the source code file]"
[Give reason for conclusion] Therefore, [license] is concluded.
```

Example 2: URL for license text

```
The information in the file is:
"This file is licensed under License A. You can find the license text at https://www.LicenseTextOfLicenseA.com."
The URL contains the license text of License A, therefore License A is concluded. The information was retrieved on [date].
```





Common corrections of scanner findings (1)

 Not a license: The scanner concludes a license from an expression in a file that is not actually a license expression at all.

```
typedef int32_t lng;

typedef ush Pos;
typedef unsigned IPos;
/* A Pos is an index in the character window. We use short instead of int to
* save space in the various tables. IPos is used only for parameter passing.
*/

enum {
WINDOW_SIZE = 2 * WSIZE,

Legend:
license relevant text
```

Action 🕡 ^	License 🐧 🔷	Source ()
₽	Zlib	scancode: #1 (100 %)
-	GPL-1.0+	scancode: #1 (78 %)
	GPL-2.0+	scancode: #1 (100 %) nomos: #1

Showing 1 to 3 of 3 entries





Common corrections of scanner findings (2)

• **Not the file's license:** The scanner concludes a license from a license expression that is part of the file's content but not the license of the file itself.

This can be disabled by turning off FEATURE_SHOW_SCRIPT in the configuration, though it won't prevent a determined user from extracting the **source code**.

It can be argued that embedded scripts are linked into the BusyBox binary and are therefore not **subject to** the 'mere aggregation' exception in the GPL. If this is the case embedded scripts should have a **licenc**e compatible with BusyBox's GPL v2-only **licenc**e.

Lamandi
<u>Legend:</u>
license relevant text
licerise relevant text

Action 🕡 ^	License 🐧 🔷	Source ()
<u>∲</u>	GPL-1.0+	scancode: #1 (85 %)
₽ 🚖	GPL-2.0	scancode: #1 (99 %) nomos: #1

Showing 1 to 2 of 2 entries





Common corrections of scanner findings (3)

• License text: Files that contain only a license text (e.g. COPYING) are concluded by the scanners to be licensed under the respective license. This is usually not correct. Most license texts are not explicitly licensed. The GNU licenses contain a license statement for the license text itself which is concluded for these cases.

GNU GENERAL PUBLIC LICENS E
Version 2, June 1991
Copyright (C) 1989, 1991 Free Software Foundation, Inc.
51 Franklin St, Fifth Floor, Boston, MA 02110-1301 USA
Everyone is permitted to copy and distribute verbatim copies
of this licens e document, but changing it is not allowed.
Preamble

Action 🐧 ^	License 🕦	Source (1)
₽	GPL-2.0-or-later WITH Autoconf-exception	monk: #1 (98 %)
X 🚖	License-of-GNU-Licenses	User decision
₽	GPL-2.0+	monk: #1 (99 %)
⊕ ☆	GPL-2.0	nomos: #1 scancode: #1 (100 %)

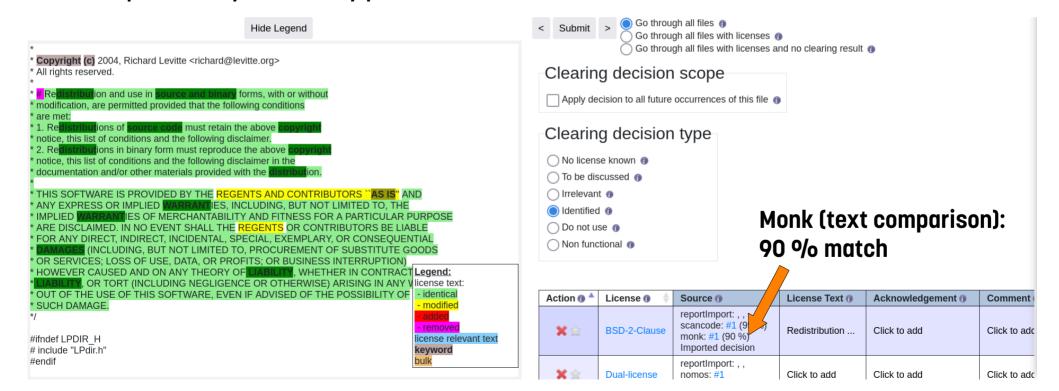
Showing 1 to 4 of 4 entries





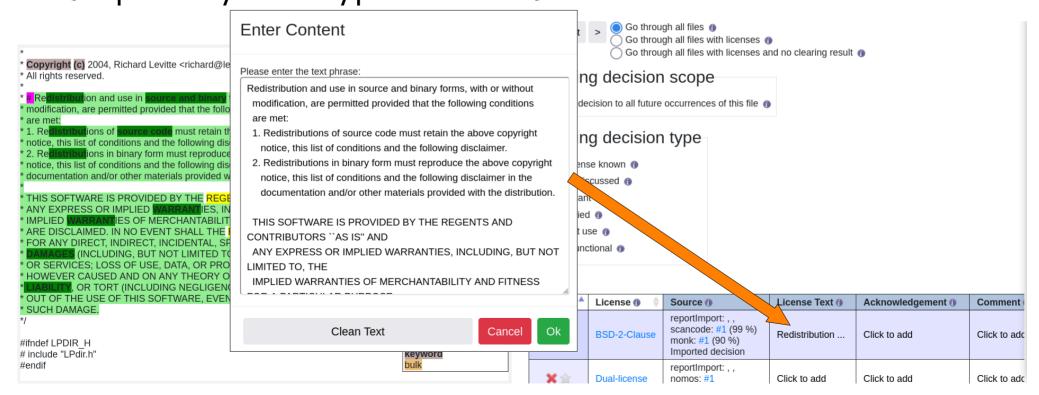
Common corrections of scanner findings (4)

• **Generic license texts:** If an individual text differs from the generic text, the individual license text is provided (especially BSD-type licenses).



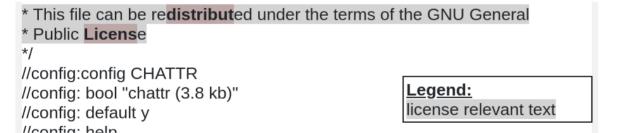
Common corrections of scanner findings (4)

• **Generic license texts:** If an individual text differs from the generic text, the individual license text is provided (especially BSD-type licenses).



Common corrections of scanner findings (5)

• Imprecise finding: The scanner finding might be imprecise, e.g. w.r.t. to the version of a license. This is concluded e.g. as v1.0-or-later.



Action 🚯 ^	License 🕦 🔷	Source ()
X	GPL-1.0+	scancode: #1 (100 %)
⊕ ☆	GPL	nomos: #1

Showing 1 to 2 of 2 entries





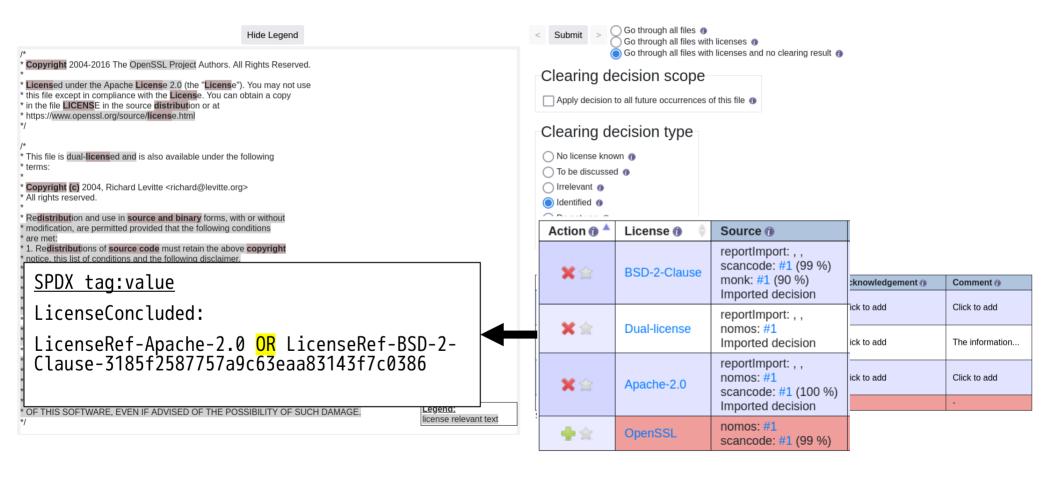
Common corrections of scanner findings (6)

• **Dual licensing:** If the context requires to chose one specific license, this choice must be noted. In any case, all applicable licenses are concluded. Dual licensing with more than two licenses requires manual post-processing.





Dual licensing



Dual licensing (more than 2 licenses)

```
* Copyright (c) 1986-2003 may safely be consumed by a BSD or GPL license.

* Written by: Vladimir Oleynik <dzo@simtreas.ru>

* Used ideas:

* Adam Rogoyski <rogoyski@cs.utexas.edu>

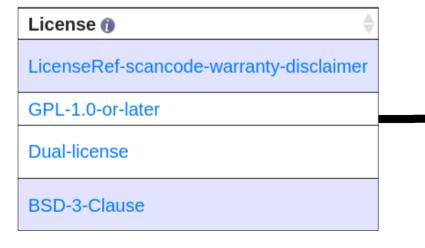
* Dave Cinege <dcinege@psychosis.com>

* Jakub Jelinek (c) 1995

* Erik Andersen <andersen@codepoet.org> (Majorly adjusted for busybox)

* This code is 'as is' with no warranty.

*/
```



SPDX tag:value Post-Processing

LicenseConcluded:

LicenseRef-scancode-warranty-discaimer AND LicenseRef-GPL-1.0-or-later AND LicenseRef-BSD-3-Clause AND LicenseRef-Dual-license

Dual licensing (more than 2 licenses)

```
* Copyright (c) 1986-2003 may safely be consumed by a BSD or GPL license.

* Written by: Vladimir Oleynik <dzo@simtreas.ru>

* Used ideas:

* Adam Rogoyski <rogoyski@cs.utexas.edu>

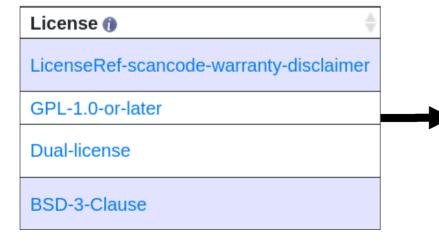
* Dave Cinege <dcinege@psychosis.com>

* Jakub Jelinek (c) 1995

* Erik Andersen <andersen@codepoet.org> (Majorly adjusted for busybox)

* This code is 'as is' with no warranty.

*/
```



SPDX tag:value Post-Processing

LicenseConcluded:

LicenseRef-scancode-warranty-discaimer AND (LicenseRef-GPL-1.0-or-later OR LicenseRef-BSD-3-Clause)

Common corrections of scanner findings (7)

• **License exceptions:** Fossology notes the license and the exception as separate findings. This is corrected to one finding using the SPDX license expression [License] WITH [exception], e.g. GPL-2.0-or-later WITH GCC-exception-2.0.

This program is **free software**; you can re**distribut**e it and/or modify it under the terms of the GNU General Public **Licens**e as published by the **Free Software** Foundation; either version 2, or (at your option) any later version.

This program is **distribut**ed in the hope that it will be useful, but WITHOUT ANY **WARRANT**Y; without even the implied **warrant**y of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public **Licens**e for more details.

You should have received a copy of the GNU General Public **Licens**e along with this program; if not, write to the **Free Software** Foundation, Inc., 59 Temple Place - Suite 330, Boston, MA 02111-1307, USA. */

/* As a special exception, when this file is copied by Bison into a Bison output file, you may use that output file **without restrict**ion **Legend**:

This special exception was added by the Free Software Foundallicense relevant text

in version 1.24 of Bison. */



Action 🕡 ^	License (1)	Source (1)
₽ 😭	LicenseRef-scancode-bison-exception-2.0	scancode: #1 (95 %)
♣ 🚖	Bison-exception	nomos: #1
	GPL-2.0+	nomos: #1 scancode: #1 (95 %)
♣ 🚖	GPL-2.0	scancode: #1 (100 %)
X 🚖	GPL-2.0-or-later-WITH-Bison-exception	User decision

Showing 1 to 5 of 5 entries

Common corrections of scanner findings (8)

• External references: If a file does not contain the name or text of a license but references an external resource, such as a COPYRIGHT file in the root directory or a URL, the external reference is checked, the detected license is concluded and the process is documented as a LicenseComment (in case of a URL, the date of access is noted).

```
# This program is free software; you can redistribute it and/or modify it
# under the same terms as Perl itself.
use Getopt::Long qw(:config no_auto_abbrev no_ignore_case);
use File::Find;
use strict;

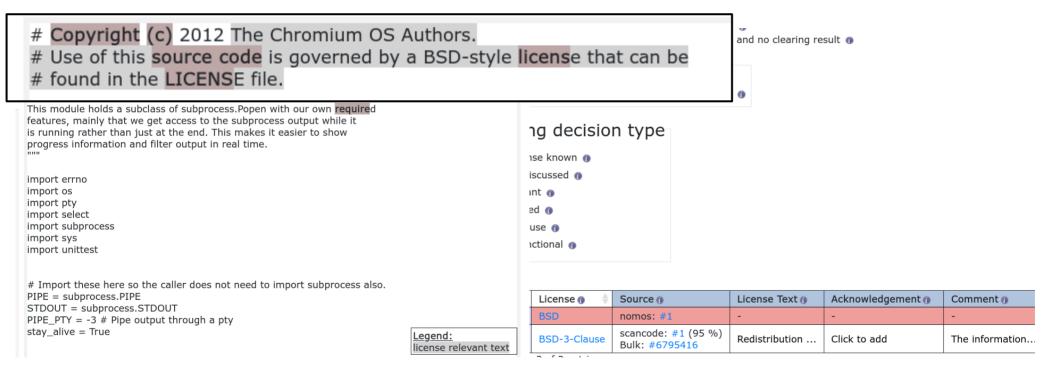
# Set up some default values

my $kdir="""
```

Action 🕡 Å	License 🐧	Source (1)
X ≘	GPL-1.0+	nomos scancode: #1 (100 %)
X≘	Artistic-1.0-Perl	nomos: #1 scancode: #1 (100 %)
X ☆	Dual-license	nomos

Showing 1 to 3 of 3 entries

And now: The unexpected (1)







And now: The unexpected (1)

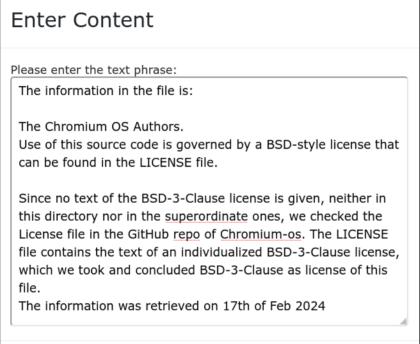
Copyright (c) 2012 The Chromium OS Authors
Use of this source code is governed by a BSD# found in the LICENSE file.

"""Subprocess execution

This module holds a subclass of subprocess.Popel features, mainly that we get access to the subprocess running rather than just at the end. This makes progress information and filter output in real time.

import errno import os import pty import select import subprocess import sys import unittest

Import these here so the caller does not need PIPE = subprocess.PIPE STDOUT = subprocess.STDOUT PIPE_PTY = -3 # Pipe output through a pty stay alive = True

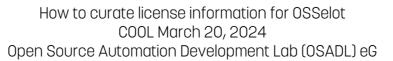


h licenses and no clearing result 🐽 of this file 🐽 Comment @ License Text @ Acknowledgen (95%)Redistribution ... Click to add The information...



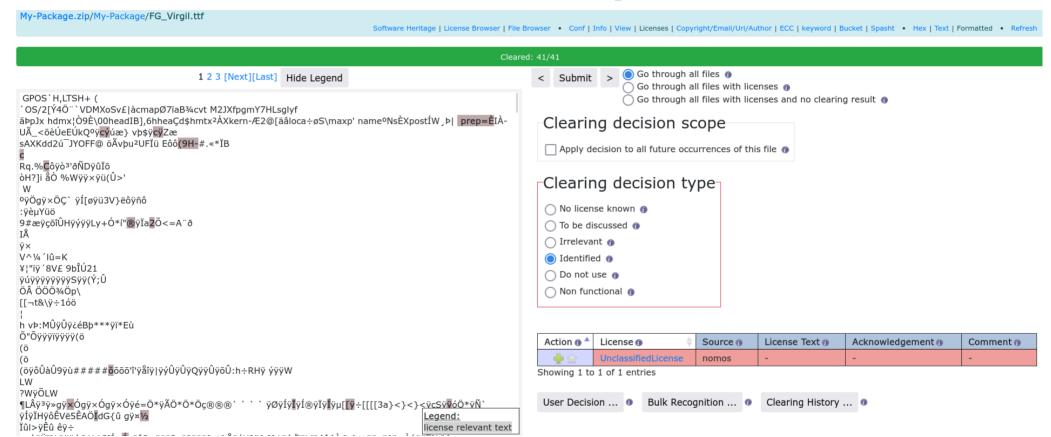








And now: The unexpected (2)





How to curate license information for OSSelot COOL March 20, 2024 Open Source Automation Development Lab (OSADL) eG



And now: The unexpected (2)

My-Package.zip/My-Pac FG_Virgil.ttf Copyright/Email/Url/Author | ECC | keyword | Bucket | Spasht • Hex | Text | Formatted • Refresh Font file Go through all files Go through all files with licenses 🐽 GPOS `H,LTSH+ (Go through all files with licenses and no clearing result 🚯 `OS/2[Y4Ö"`VDMXoSv£|àcmapØ7ïaB¾cvt M2JXfpgmY7HLsglyf āþpJx hdmx¦Ò9È\00headIB],6hheaCd\$hmtx²ÀXkern-Æ2@[äâloca÷øS\maxp' nameoNsÈXpostÍW b| prep=ÈIÀ-Clearing decision scope UÃ <ōèÚeEÚkO°ÿcýúæ} vb\$ÿcÿZæ sAXKdd2ú JYOFF@ öÃvbu2UFÏü Eôô(9H-#.«*ÏB Apply decision to all future occurrences of this file n Rg.%Côÿò³'ðÑDÿûÏō òH?1i åÒ %Wÿÿ×ÿü(Û> Clearing decision type °ÿÖgÿ×ÖC` ÿÍſøÿü3V}ëôÿñô No license known :ÿèuYüö 9#æÿcōîÛHÿýÿÿLy+Ó*í"®ÿÏa2Õ<=A¨ð To be discussed 🖳 Irrelevant 🐽 V^1/4 'Iû=K Identified 6 ¥!"ïÿ '8V£ 9bÎÚ21 Do not use not use ÿúÿÿÿÿÿÿÿÿSÿÿ(Ý;Û ÖÂ ÖÖÖ¾Öp\ Non functional n [[¬t&\v+1óö h vÞ:MÛÿÛÿ¿éBb***ÿï*Eù Õ"Õÿÿÿïÿÿÿÿ(ö License Text n Action 🙃 📤 License @ Source @ Acknowledgement (1) Comment @ nomos (öÿôÛàÛ9ÿù#####ōōōō'î'ÿåîÿIÿýÛÿÛÿOÿÿÛÿōÛ:h÷RHÿ ýÿÿW



ÿÍÿÏHÿôÊVë5ÊAÖIdG{û gÿ×1/2

¶LÁÿ³ÿ»gÿ×ógÿ×ógÿ×óÿé=Ö*ÿÃÖ*Ö*Ö¢®®®`````ÿØÿÍÿĬÿÍ®ÿÏÿĬÿµ[¡ÿ÷[[[[3a}<}<}<}<\sçvvőö*vñ`

LW ?WŸÕLW

Ïûl>ÿÊû êÿ÷

How to curate license information for OSSelot COOL March 20, 2024 Open Source Automation Development Lab (OSADL) eG

Legend:

license relevant text

Showing 1 to 1 of 1 entries

Bulk Recognition ... 0

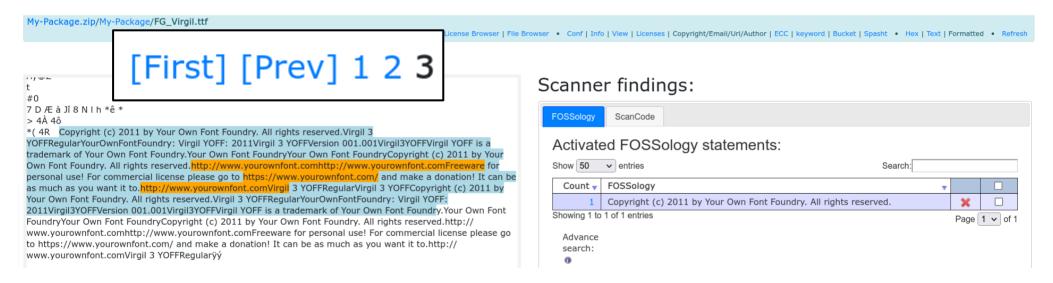
User Decision ... 0



Clearing History ... 0

And now: The unexpected (2)

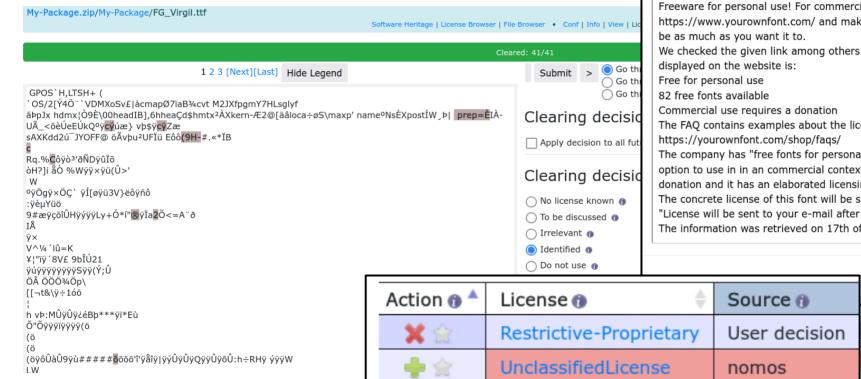
• When there is no license information, check copyright findings:

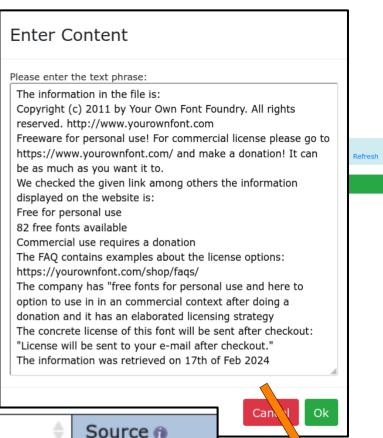






And now: The unexp







How to curate license information for OSSelot COOL March 20, 2024 Open Source Automation Development Lab (OSADL) eG



Comment @

The information...

nowledgement 🚯

to add

SPDX in OSSelot

• OSSelot SPDX reports serve as **stand-alone documents** to fulfill information obligations, *i.e.* **all license texts are included** and not only referenced.





SPDX tag:value report – Creation information

```
Curator's name
## Creation Information
Creator: Tool: spdx2
Creator: Person: Oliver Fendt
CreatorComment: <text>
This document was created using license information and a generator
from Fossology.
It contains the license and copyright analysis of OpenSSL 3.0.5
Please check "LicenseComments" for explanations of concluded licenses
</text>
Created: 2022-07-06T14:58:22Z
```

LicenseListVersion: 2.6

SPDX tag:value report – Package information

```
Only if there is a main license for the package (LICENSE/COPYING file
## Package Information
##-
                                                             in root directory)
PackageName: openssl-openssl-3.0.5.tar.gz
[...]
PackageChecksum: SHA1: edc3465a8a43ce580268e726b6f7b827f4a6261e
PackageChecksum: SHA256:
b6363cf1bca88f0a46a768883a225e644135432d6a51ab1c4660ab58af541 #7 the package contains
PackageChecksum: MD5: 22733b9187548b735201fd9f7aa12e71
                                                                       licensing, e.g. third
PackageLicenseConcluded: NOASSERTION
                                                                         party notices.
PackageLicenseDeclared: LicenseRef-Apache-2-0
PackageLicenseComments: <text> Additional information found in: [File name of third
party notices]
[Third party notices]
[...]</text>
```

SPDX tag:value report – File information

##File

Final license decision

```
FileName: openssl-3.0.5.tar.qz/openssl-3.0.5.tar/openssl-openssl-3.0.5/crypto/LPdir_wince.c
SPDXID: SPDXRef-item158856105
                                                                              If not obvious,
FileChecksum: SHA1: dc3e4bb9f2cf76426da9ad5dbc8ad4a2356c3359
                                                                          explanation of license
FileChecksum: SHA256: fd878a5b569cd41d63ba673420a4d95adfac9ad3048ea0fb4
                                                                                 decision
FileChecksum: MD5: 62fbea2db5fb486d537d869af119135b
LicenseConcluded: LicenseRef-Apache-2.0 OR LicenseRef-BSD-2 Lause-3185f2587757a9c63eaa83143f7c0386
LicenseComments: <text>The information in the file is:
Besides the Apache-2.0 header the following information is in the file:
This file is dual-licensed and is also available under the following terms:
Followed by the BSD-2-clause license text. Thus dual licensing was concluded. </text>
LicenseInfoInFile: LicenseRef-Apache-2.0-
                                                                          Scanner findings
LicenseInfoInFile: LicenseRef-OpenSSL
LicenseInfoInFile: LicenseRef-Dual-license
                                                                            Copyright notices
LicenseInfoInFile: LicenseRef-BSD-2-Clause REGENTS-AND-CONTRIBUTORS
FileCopyrightText: <text> Copyright 2004-2016 The OpenSSL Project Authors
```

Copyright (c) 2004, Richard Levitte <richard@levitte.org></text>

SPDX tag:value report – License information

```
##-----
## License Information
##-----
```

LicenseID: LicenseRef-Apache-2.0
LicenseName: Apache License 2.0

ExtractedText: <text> Apache License

Version 2.0, January 2004

http://www.apache.org/licenses/

Also include template licenses to make the SPDX report stand-alone.

Add "LicenseRef-" prefix for SPDX validity.

TERMS AND CONDITIONS FOR USE, REPRODUCTION, AND DISTRIBUTION

1. Definitions.

 $[\ldots]$

SPDX tag:value report – License information

LicenseID: LicenseRef-BSD-2-Clause-3185f2587757a9c63eaa83143f7c0386
LicenseName: BSD-2-Clause-3185f2587757a9c63eaa83143f7c0386

ExtractedText: <text> Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- 1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- 2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS
'AS IS'' AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT
LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR
A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT
OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL,
SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT
LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE,
DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY
THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT
(INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE
OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE. </text>

Conclusion

- The OSSelot project provides a publicly available database with curated compliance information licensed under CCO-1.0 for frequently used FOSS components.
- To maintain **trust** in the database, the data curation must follow an **established standard** and be conducted **diligently**.
- Some manual input and review is still indispensable, but:

The reuse of curated licensing and copyright information can tremendously reduce the time required to clear a software package!







- https://www.osselot.org is the official project page with news, explanations, example use cases, tools and to dos.
- https://wiki.osselot.org contains tools for automated use of the curation data and information on the curation process.
- The curated material and other project info is available via GitHub https://github.com/Open-Source-Compliance/package-analysis
- Contact: info@osselot.org

Questions, requests and contributions are very welcome!



