LINUXCON containercon CHINA 中国 CLOUDOPEN

THINK OPEN
开放性思维

License Information Management: Zephyr Case Study

Kate Stewart & Steve Winslow
Linux Foundation

Opening Up Your Source Code

So you picked a license...

Opening Up Your Source Code

So you picked a license...

Apache License
Version 2.0, January 2004
http://www.apache.org/licenses/

TERMS AND CONDITIONS FOR USE, REPRODUC

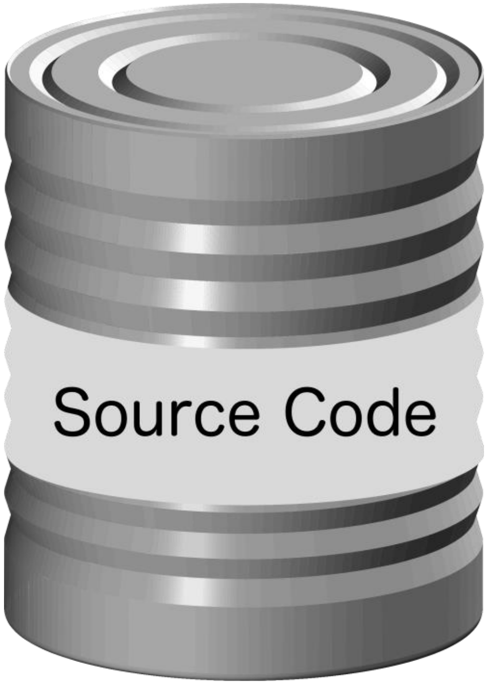
1. Definitions. "License" shall mean the terms and conditions for use, reproduction, and distribution as defined by Sections 1 through 9 of ...

LICENSE.txt

...now what?

”Confetti Girl” image by Scout; used under CC0-1.0; https://openclipart.org/detail/232158/confetti-girl

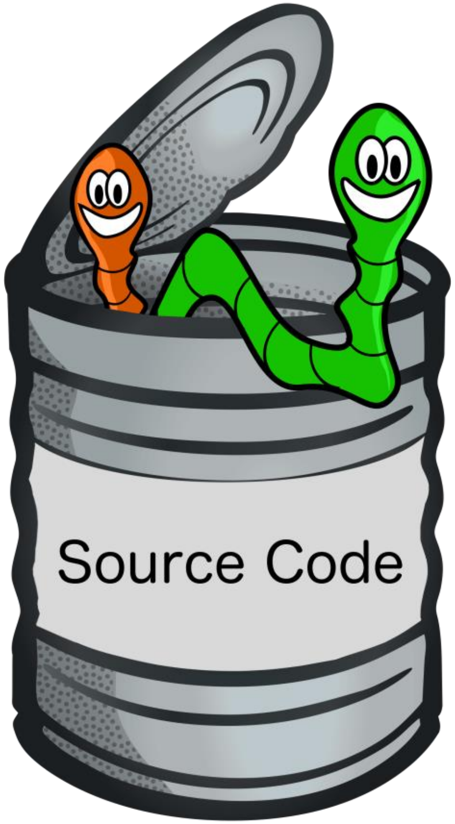
Opening Up Your Source Code



What licenses are already inside your source code?

”Tin can” image by jhnri4; used under CC0-1.0; https://openclipart.org/detail/194577/tin-can

Opening Up Your Source Code

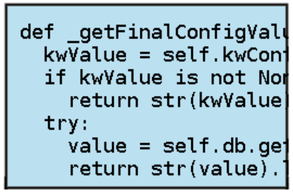


What licenses are already inside your source code?

(potentially more than you expected)

“Worm” image by neo1012; used under CC0-1.0; https://openclipart.org/detail/82867/worm
“Inchworm” image by artonymous; used under CC0-1.0; https://openclipart.org/detail/177554/inchworm
”can - coloured” image by frankes; used under CC0-1.0; https://openclipart.org/detail/246208/can-coloured

Opening Up Your Source Code



```
def _getFinalConfigValu
  kwValue = self.kwConf
  if kwValue is not Non
    return str(kwValue)
  try:
    value = self.db.get
    return str(value).
```

An existing code base might contain:

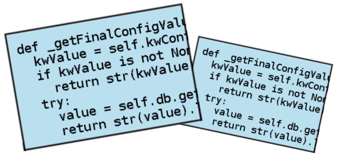
- your own code

Opening Up Your Source Code

An existing code base might contain:

- your own code
- third-party proprietary code

EULA



```
def _getFinalConfigVal
  kwValue = self.kwCon
  if kwValue is not No
    return str(kwValue
  try:
    value = self.db.ge
    return str(value).
```

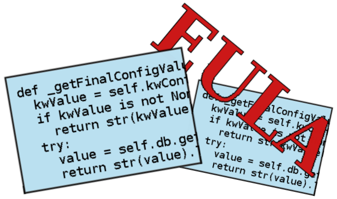
```
def _getFinalConfigVal
  kwValue = self.kwCon
  if kwValue is not No
    return str(kwValue
  try:
    value = self.db.ge
    return str(value).
```

Opening Up Your Source Code

GPL-2.0
and
MPL-1.1
and
Ms-RL
and
...

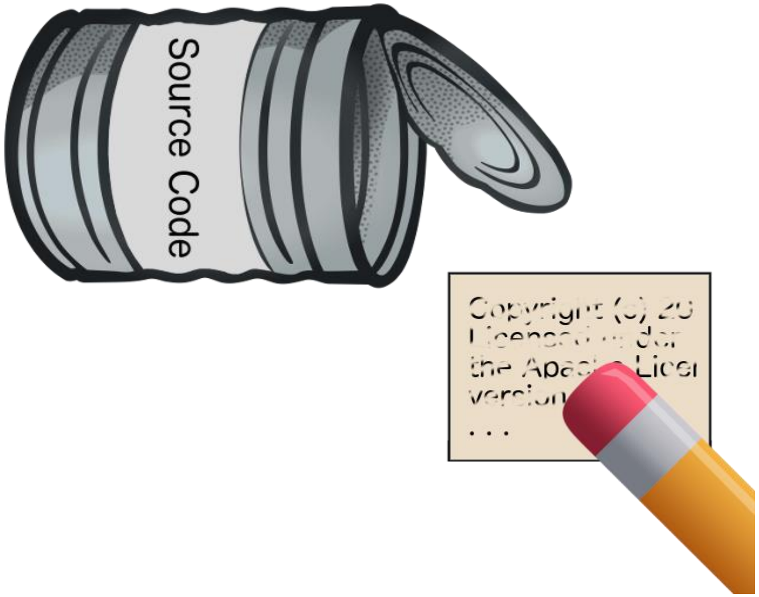
An existing code base might contain:

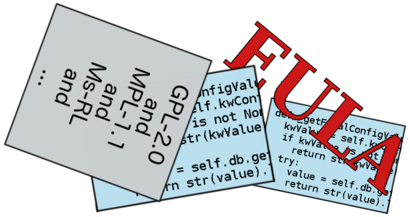
- your own code
- third-party proprietary code
- incompatible open source licenses



```
def _getFinalConfigVal
  kwValue = self.kwConf
  if kwValue is not Non
    return str(kwValue
  try:
    value = self.db.get
    return str(value).
```

Opening Up Your Source Code



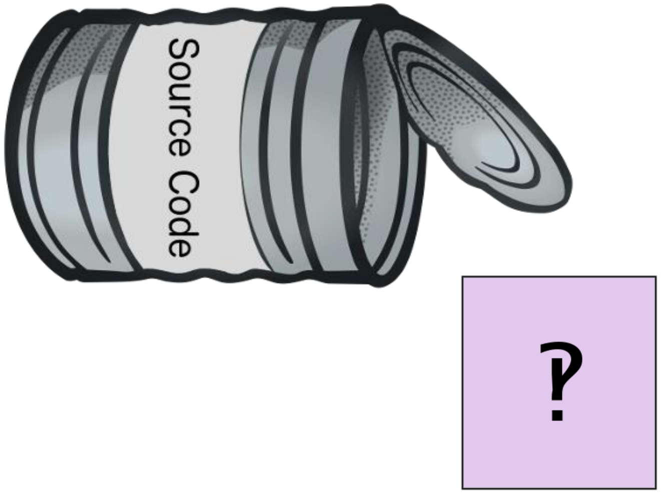


An existing code base might contain:

- your own code
- third-party proprietary code
- incompatible open source licenses
- missing open source licenses

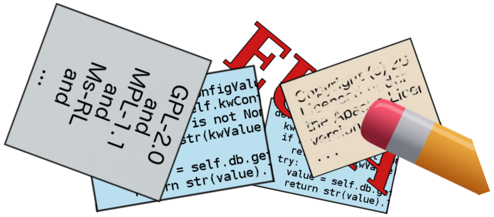
Pencil image by TheUjulala; used under CC0-1.0
https://pixabay.com/en/pencil-pen-orange-red-eraser-190586

Opening Up Your Source Code



An existing code base might contain:

- your own code
- third-party proprietary code
- incompatible open source licenses
- missing open source licenses
- puzzling license statements



Opening Up Your Source Code

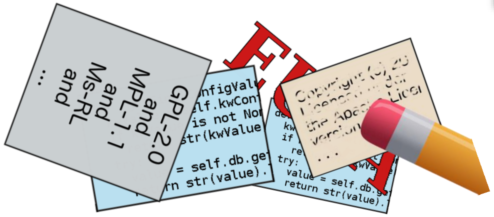
?

An existing code base might contain:

- your own code
- third-party proprietary code
- incompatible open source licenses

“See LICENSE in LICENSE”
(with no LICENSE file in repo)

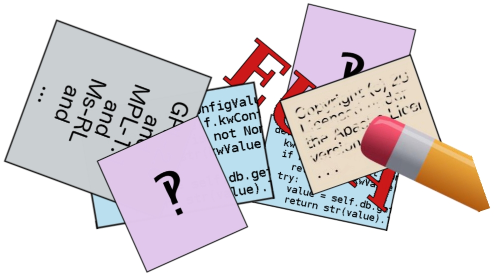
“Licensed under the **Creative Commons Attribution 4.0** International License, titled **CC-BY-SA-4.0**”



Opening Up Your Source Code

An existing code base might contain:

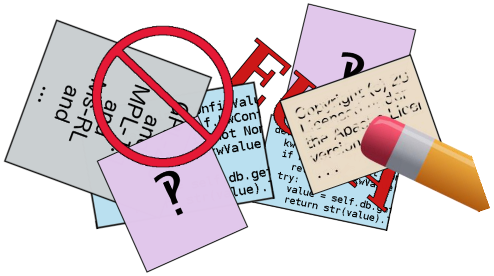
- your own code
- third-party proprietary code
- incompatible open source licenses
- missing open source licenses
- puzzling license statements
- your own confidentiality notices



Opening Up Your Source Code

An existing code base might contain:

- your own code
- third-party proprietary code
- incompatible open source licenses
- missing open source licenses
- puzzling license statements
- your own confidentiality notices
- code with snarky licenses



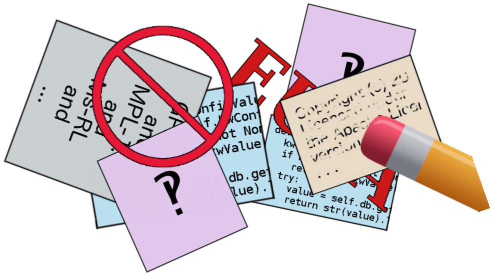
Opening Up Your Source Code

An existing code base might contain:

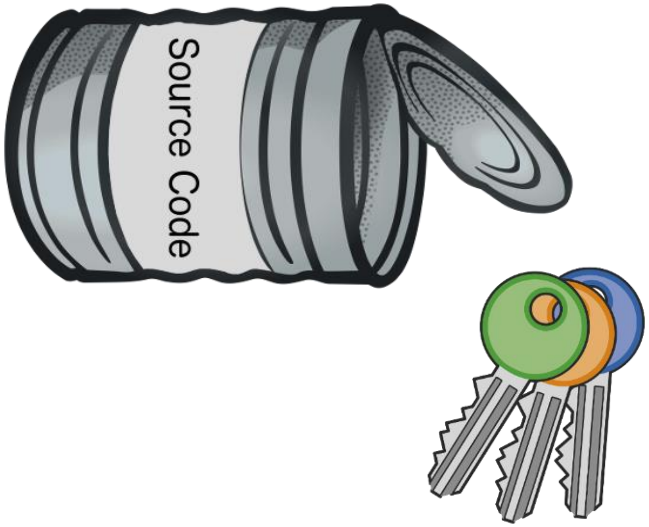
- your own code
- third-party proprietary code
- incompatible open source licenses

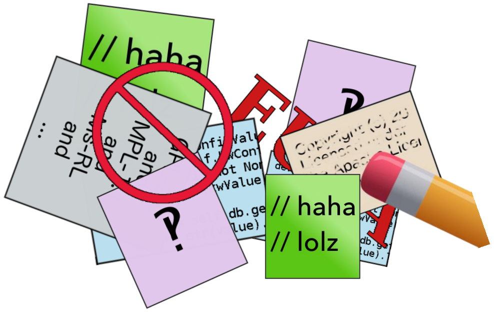
“This is free software; you can redistribute it and/or modify it under the terms of the BSD License. **Use by owners of Che Guevarra parafernalia is prohibited, where possible, and highly discouraged elsewhere.**”

// haha
// lolz



Opening Up Your Source Code



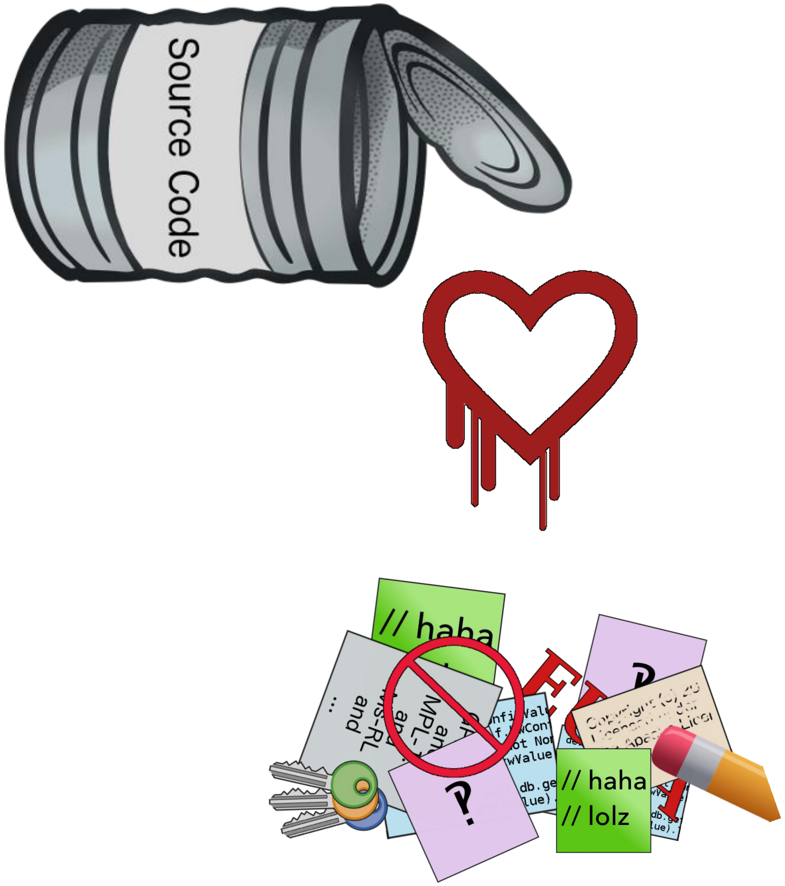


An existing code base might contain:

- your own code
- third-party proprietary code
- incompatible open source licenses
- missing open source licenses
- puzzling license statements
- your own confidentiality notices
- code with snarky licenses
- code with secret keys or passwords

”Cles de serrure – lock keys” image by enolynn; used under CC0-1.0
https://openclipart.org/detail/190821/cles-de-serrure-lock-keys

Opening Up Your Source Code

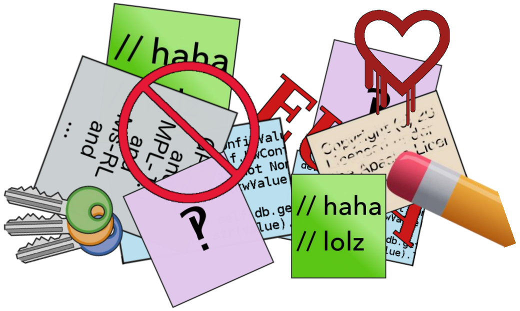


An existing code base might contain:

- your own code
- third-party proprietary code
- incompatible open source licenses
- missing open source licenses
- puzzling license statements
- your own confidentiality notices
- code with snarky licenses
- code with secret keys or passwords
- code with security vulnerabilities

Heartbleed logo image by Synopsys, Inc.; used under CC0-1.0
http://heartbleed.com/

Opening Up Your Source Code



An existing code base might contain:

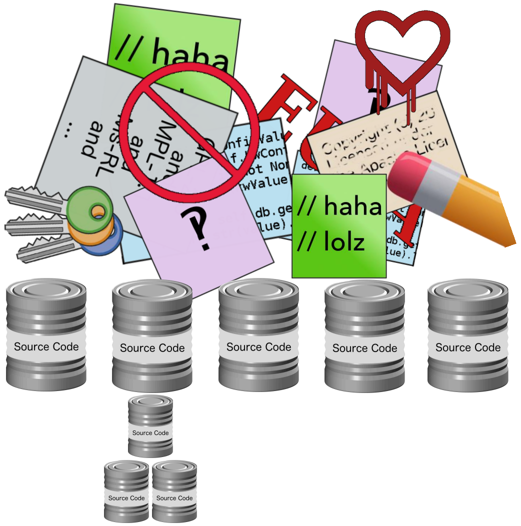
- your own code
- third-party proprietary code
- incompatible open source licenses
- missing open source licenses
- puzzling license statements
- your own confidentiality notices
- code with snarky licenses
- code with secret keys or passwords
- code with security vulnerabilities

Opening Up Your Source Code

An existing code base might contain:

- your own code
- third-party proprietary code
- incompatible open source licenses
- missing open source licenses
- puzzling license statements
- your own confidentiality notices
- code with snarky licenses
- code with secret keys or passwords
- code with security vulnerabilities
- dependencies with any of the above

Opening Up Your Source Code



An existing code base might contain:

- your own code
- third-party proprietary code
- incompatible open source licenses
- missing open source licenses
- puzzling license statements
- your own confidentiality notices
- code with snarky licenses
- code with secret keys or passwords
- code with security vulnerabilities
- (sub)dependencies with any of the above

Tempting Response: Ignore it

Drop in a LICENSE.txt file and declare yourself done

License Management

General process:

- Identify licenses
- Address incompatibilities
- Address compliance
- Communicate licenses

License Management

General process:

- Identify licenses
- Address incompatibilities
- Address compliance
- Communicate licenses

Focusing on these two in this talk

(getting these right enables meaningful conversations about the other two)

License information can be managed!

This is not an insurmountable challenge

Tackling it benefits projects and benefits the whole ecosystem
(and not just by making lawyers happier!)

Avoid making “perfect” the enemy of “better”

There are gaps in today’s tooling but there is also forward progress

Identifying Licenses

Goal 1: Determine which licenses are relevant to your project

Goal 2: Do so in an automated, scalable way

Identifying Licenses

Different types of scans:

- license scanning
- code scanning
- dependency scanning

Scanning Tools

Quick and dirty; no tooling needed

Look for relevant words / fragments:

- “licen”
- “redist”
- “copyright”
- common license fragments: “bsd”, “gpl”, “general public”, “cddl”, ...

Manual searches

```
grep -nri
```

(or your favorite command line args)

```
Ctrl-F
```

(or your favorite editor’s equivalent)

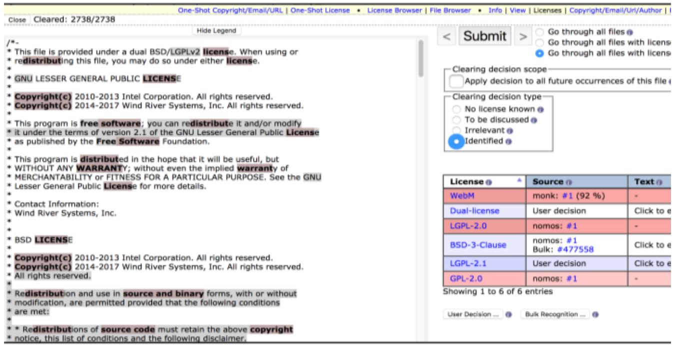
Scanning Tools

FOSSology is used to scan a codebase for licenses

Performs textual analysis and regular expression scanning to identify likely license notices and references

Supplemented with manual review to remove false positives and investigate unusual findings

fossology


One-Shot Copyright/Email/URL | One-Shot License • License Browser | File Browser • Info | View | Licenses | Copyright/Email/Url/Author |
Close
Hide Legend
< Submit >

Cleared: 2738/2738

```
/*-
* This file is provided under a dual BSD/LGPLv2 license. When using or
* redistributing this file, you may do so under either license.
*
* GNU LESSER GENERAL PUBLIC LICENSE
*
* Copyright(c) 2010-2013 Intel Corporation. All rights reserved.
* Copyright(c) 2014-2017 Wind River Systems, Inc. All rights reserved.
*
* This program is free software; you can redistribute it and/or modify
* it under the terms of version 2.1 of the GNU Lesser General Public License
* as published by the Free Software Foundation.
*
* This program is distributed in the hope that it will be useful, but
* WITHOUT ANY WARRANTY; without even the implied warranty of
* MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU
* Lesser General Public License for more details.
*
* Contact Information:
* Wind River Systems, Inc.
*
*
* BSD LICENSE
*
* Copyright(c) 2010-2013 Intel Corporation. All rights reserved.
* Copyright(c) 2014-2017 Wind River Systems, Inc. All rights reserved.
* All rights reserved.
*
* Redistribution and use in source and binary forms, with or without
* modification, are permitted provided that the following conditions
* are met:
*
* * Redistributions of source code must retain the above copyright
* notice, this list of conditions and the following disclaimer.
```

○ Go through all files
○ Go through all files with licens
● Go through all files with licens

Clearing decision scope
□ Apply decision to all future occurrences of this file

Clearing decision type
○ No license known
○ To be discussed
○ Irrelevant
● Identified

License	Source	Text
WebM	monk: #1 (92 %)	-
Dual-license	User decision	Click to e
LGPL-2.0	nomos: #1	-
BSD-3-Clause	nomos: #1 Bulk: #477558	Click to e
LGPL-2.1	User decision	Click to e
GPL-2.0	nomos: #1	-

Showing 1 to 6 of 6 entries

Scanning Tools

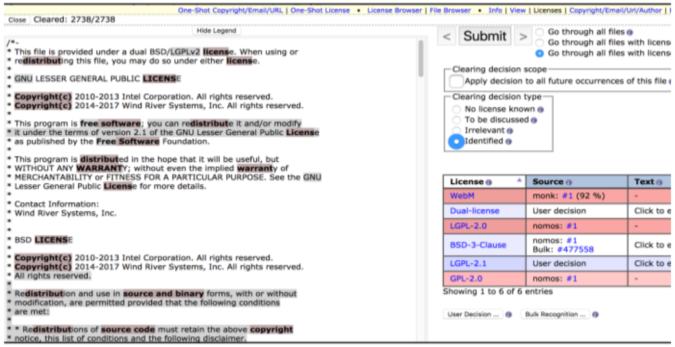
Version 3.3 released in May 2018

Since 3.2 it includes:

- SPDX file imports
- Obligation analysis and summaries

https://www.fossology.org/

https://github.com/fossology/fossology


One-Shot Copyright/Email/URL | One-Shot License • License Browser | File Browser • Info | View | Licenses | Copyright/Email/Url/Author |
Close
Hide Legend
<
Submit
>
User Decision ...
Bulk Recognition ...

Cleared: 2738/2738

```
/*-
* This file is provided under a dual BSD/LGPLv2 license. When using or
* redistributing this file, you may do so under either license.
*
* GNU LESSER GENERAL PUBLIC LICENSE
*
* Copyright(c) 2010-2013 Intel Corporation. All rights reserved.
* Copyright(c) 2014-2017 Wind River Systems, Inc. All rights reserved.
*
* This program is free software; you can redistribute it and/or modify
* it under the terms of version 2.1 of the GNU Lesser General Public License
* as published by the Free Software Foundation.
*
* This program is distributed in the hope that it will be useful, but
* WITHOUT ANY WARRANTY; without even the implied warranty of
* MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU
* Lesser General Public License for more details.
*
* Contact Information:
* Wind River Systems, Inc.
*
*
* BSD LICENSE
*
* Copyright(c) 2010-2013 Intel Corporation. All rights reserved.
* Copyright(c) 2014-2017 Wind River Systems, Inc. All rights reserved.
* All rights reserved.
*
* Redistribution and use in source and binary forms, with or without
* modification, are permitted provided that the following conditions
* are met:
*
* * Redistributions of source code must retain the above copyright
* notice, this list of conditions and the following disclaimer.
```

○ Go through all files
○ Go through all files with licens
● Go through all files with licens

Clearing decision scope

□ Apply decision to all future occurrences of this file

Clearing decision type

○ No license known
○ To be discussed
○ Irrelevant
● Identified

License	Source	Text
WebM	monk: #1 (92 %)	-
Dual-license	User decision	Click to e
LGPL-2.0	nomos: #1	-
BSD-3-Clause	nomos: #1 Bulk: #477558	Click to e
LGPL-2.1	User decision	Click to e
GPL-2.0	nomos: #1	-

Showing 1 to 6 of 6 entries

Scanning Tools

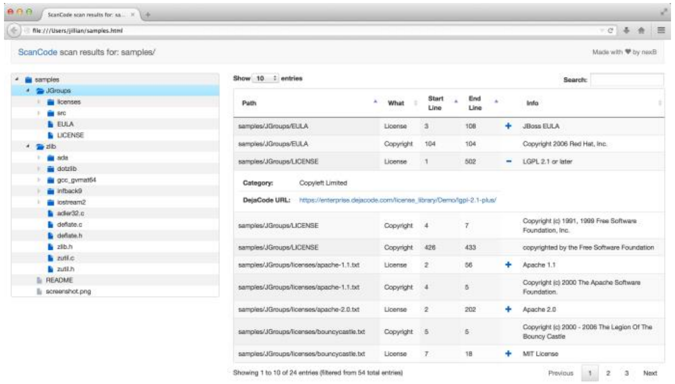
ScanCode Toolkit

by nexB

From ScanCode's README:

ScanCode is a suite of utilities used to scan a codebase for license, copyright, package manifests and dependencies and other interesting information that can be discovered in source and binary code files.

https://github.com/nexB/scancode-toolkit


ScanCode scan results for: sa... ×
file:///Users/jillian/samples.html
Made with ♥ by nexB
Show 10 entries
Search:

ScanCode scan results for: samples/

- samples
  - JGroups
    - licenses
    - src
    - EULA
    - LICENSE
  - zlib
    - ada
    - dotzlib
    - gcc_gvmat64
    - infback9
    - iostream2
    - adler32.c
    - deflate.c
    - deflate.h
    - zlib.h
    - zutil.c
    - zutil.h
  - README
  - screenshot.png

Path	What	Start Line	End Line		Info
samples/JGroups/EULA	License	3	108	+	JBoss EULA
samples/JGroups/EULA	Copyright	104	104		Copyright 2006 Red Hat, Inc.
samples/JGroups/LICENSE	License	1	502	−	LGPL 2.1 or later
Category:	Copyleft Limited				
DejaCode URL:	https://enterprise.dejacode.com/license_library/Demo/lgpl-2.1-plus/				
samples/JGroups/LICENSE	Copyright	4	7		Copyright (c) 1991, 1999 Free Software Foundation, Inc.
samples/JGroups/LICENSE	Copyright	426	433		copyrighted by the Free Software Foundation
samples/JGroups/licenses/apache-1.1.txt	License	2	56	+	Apache 1.1
samples/JGroups/licenses/apache-1.1.txt	Copyright	4	5		Copyright (c) 2000 The Apache Software Foundation.
samples/JGroups/licenses/apache-2.0.txt	License	2	202	+	Apache 2.0
samples/JGroups/licenses/bouncycastle.txt	Copyright	5	5		Copyright (c) 2000 - 2006 The Legion Of The Bouncy Castle
samples/JGroups/licenses/bouncycastle.txt	License	7	18	+	MIT License

Showing 1 to 10 of 24 entries (filtered from 54 total entries)

ScanCode Toolkit screenshot Copyright (c) 2017 nexB Inc. and others; used under Apache-2.0
https://github.com/nexB/scancode-toolkit/blob/develop/samples/screenshot.png

Scanning tools

Various other scanning tools and services, including open source and proprietary / commercial options

Some include security vulnerability detection

Some include initial free tiers for open source projects
(read carefully how they define “open source” and “projects”)

Scanning tools

Keep in mind:

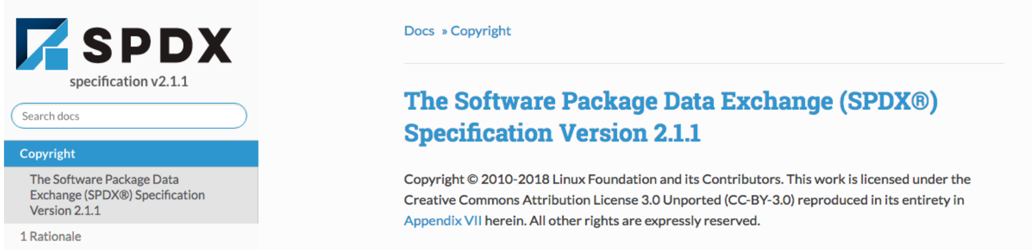
However automated the tooling is,
some manual review will likely be required

Communicating License Information

Goal 1: Let others know what licenses are relevant to your project

Goal 2: Do so in an automated, scalable way

Communicating License Information



SPDX specification v2.1.1

Search docs
Copyright
The Software Package Data Exchange (SPDX®) Specification Version 2.1.1
1 Rationale
Docs » Copyright

The Software Package Data Exchange (SPDX®) Specification Version 2.1.1

Copyright © 2010-2018 Linux Foundation and its Contributors. This work is licensed under the Creative Commons Attribution License 3.0 Unported (CC-BY-3.0) reproduced in its entirety in Appendix VII herein. All other rights are expressly reserved.

From the specification:

- The Software Package Data Exchange (SPDX®) specification is a standard format for communicating the components, licenses, and copyrights associated with software packages.

Current verson:
https://spdx.github.io/spdx-spec/

Prior Versions:
https://spdx.org/specifications

Communicating License Information

SPDX Documents comprise manifests of files from software packages

Includes checksum hashes per file, license information and other optional data

Two official formats:

- **XML** – easier for automated consumption
- **Tag-value** - easier for human consumption

Translation tools can convert to spreadsheets, JSON, YAML, XML etc., and next revision of spec (2.2) will make them official

SPDX

SPDX Documents

https://spdx.github.io/spdx-spec/

Communicating License Information

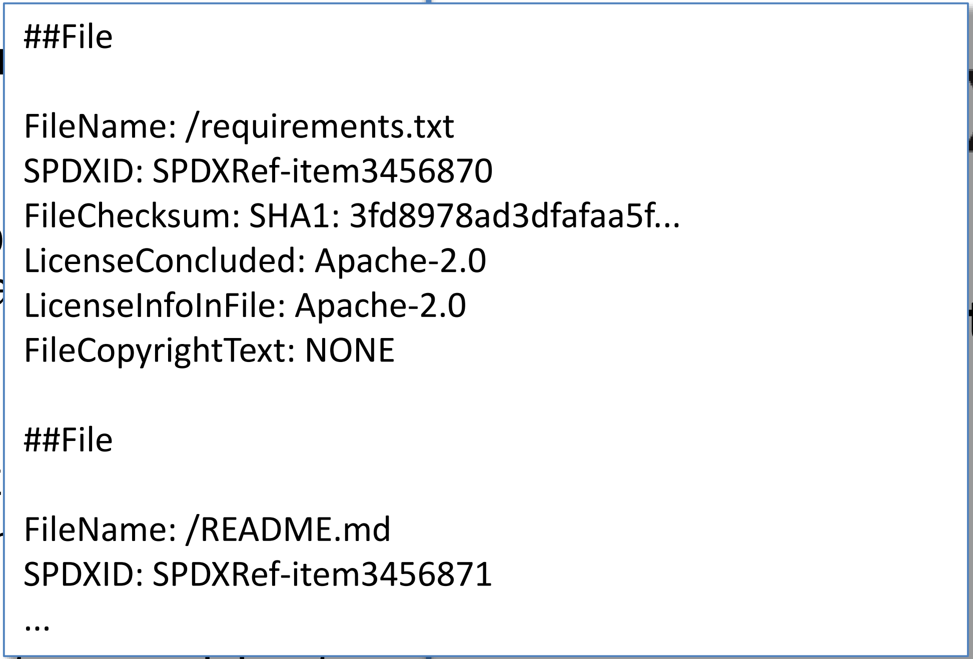
SPDX Documents comprise r
from software packages

Includes checksum hashes p
information and other optiona

Two official formats:

- **XML** – easier for automat
- **Tag-value** - easier for hur

Translation tools can convert to spreadsheets, JSON, YAML, XML etc., and next revision of spec (2.2) will make them official



```
##File

FileName: /requirements.txt
SPDXID: SPDXRef-item3456870
FileChecksum: SHA1: 3fd8978ad3dfafaa5f...
LicenseConcluded: Apache-2.0
LicenseInfoInFile: Apache-2.0
FileCopyrightText: NONE

##File

FileName: /README.md
SPDXID: SPDXRef-item3456871
...
```

X

ts

https://spdx.github.io/spdx-spec/

Communicating License Information

From the License List:

"...a list of commonly found licenses and exceptions used in free and open source and other collaborative software or documentation."

"The purpose of the SPDX License List is to enable easy and efficient identification of such licenses and exceptions in an SPDX document, in source files or elsewhere."

SPDX

SPDX License List

https://spdx.org/licenses

Communicating License Information

From the License List:

“...a list of commonly found licenses a
exceptions used in free and open sou
and other collaborative software or
documentation.”

“The purpose of the SPDX License Li
enable easy and efficient identificatior
such licenses and exceptions in an SPDX
document, in source files or elsewhere.”

Examples:

BSD-2-Clause
BSD-3-Clause
GPL-2.0-only
GPL-3.0-or-later
MIT
MPL-2.0
…

PDX

cense List

https://spdx.org/licenses

Communicating License Information

One-line comment in each source code file to unambiguously designate the applicable license(s)

Examples:

```
/* SPDX-License-Identifier: GPL-2.0-only */

// SPDX-License-Identifier: BSD-2-Clause OR MIT

# SPDX-License-Identifier: Apache-2.0 AND MIT
```

SPDX Short-Form IDs

Usage example:
https://www.kernel.org/doc/html/latest/process/license-rules.html

Communicating License Information

One-line commen
unambiguously d

Examples:

```
/* SPDX-License-Identifier: GPL-2.0-only */

// SPDX-License-Identifier: BSD-2-Clause OR MIT

# SPDX-License-Identifier: Apache-2.0 AND MIT
```

If a file’s license ID looks like this, maybe rethink that file’s structure….

GPL-3.0 AND GPL-2.0+ AND GPL-2.0 AND LGPL-2.1+ AND LGPL-2.1 AND MIT AND BSD-3-Clause AND (AFL-2.1+ OR BSD-3-Clause) AND (MIT OR LicenseRef-BSD OR LicenseRef-GPL) AND (MIT OR LicenseRef-GPL) AND (MPL-1.1 OR GPL-2.0 OR LGPL-2.1) AND LicenseRef-MIT-style

Usage example:
https://www.kernel.org/doc/html/latest/process/license-rules.html

Communicating License Information

The REUSE Initiative (from Free Software Foundation Europe) provides **best practices** in communicating license information for an entire package, and **tools** to assist in confirming compliance with those practices.

Includes recommendations for how and where to place copyright notices, license references and license texts

Makes use of SPDX short-form identifiers

REUSE SOFTWARE

REUSE Initiative

https://reuse.software

Related Suggestions

Contribution instructions for your project:

- Include a file (CONTRIBUTIONS.md) which explains that contributions are required to be made under the project's license
- In that file, also include:
  - the Developer Certificate of Origin (https://developercertificate.org/)
  - a statement that "Signed-off-by:" lines in commit messages signal an affirmation to the DCO

Related Suggestions

Location for third-party software:

- Whenever possible, where third-party software is included within your repository, keep it in a separate “third-party/” or “ext/” or similar folder
  - May already be a standard or semi-standard, e.g. “vendor/” folder for many Golang projects; “node_modules/” for NPM projects
- Helps flag to downstream users that licenses may differ
- Also provides a good place to focus when looking for security vulnerabilities in dependencies

Current Status and Gaps

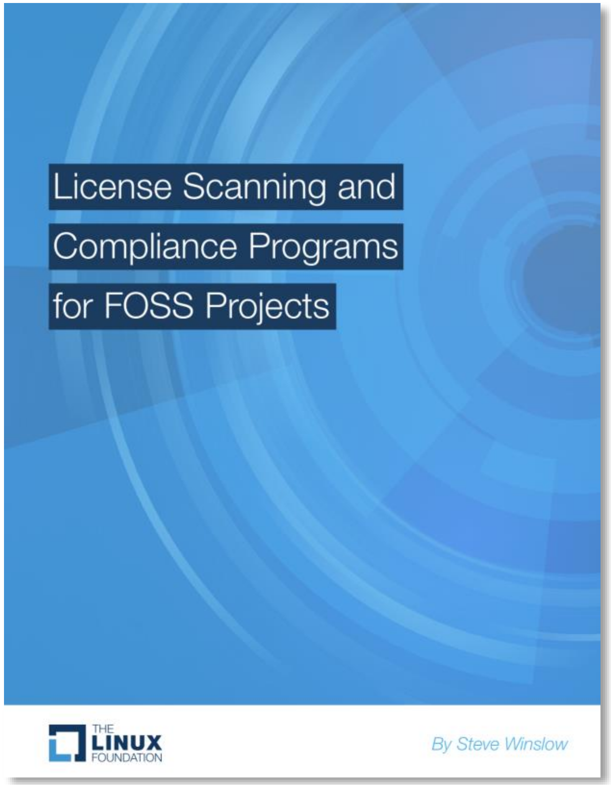
How well do all these pieces fit together?

...disparate tools; it’s a work in progress

Focus is now turning to developing centralizing tools to unify these different parts of the licensing story

- e.g. Quartermaster (http://qmstr.org/)

To Learn More…



License Scanning and Compliance Programs for FOSS Projects

THE LINUX FOUNDATION

By Steve Winslow

Free publication available from The Linux Foundation website:

https://www.linuxfoundation.org/publications/license-scanning-compliance-programs-foss-projects/

Now available in Chinese!

FOSS项目的
许可扫描及合规计划

THE LINUX FOUNDATION

作者：Steve Winslow

Free publication available from The Linux Foundation website:

https://www.linuxfoundation.org/publications/license-scanning-compliance-programs-foss-projects/

DOWNLOAD THE PAPER (CHINESE)

Case Study: Zephyr

Zephyr™

https://www.zephyrproject.org/

https://github.com/zephyrproject-rtos/zephyr

Case Study: Zephyr

The Zephyr project is Apache-2.0 licensed

The project leaders and developers have intentionally focused on improving management of the license information for their code

Case Study: Zephyr

Zephyr license processes:

- License review (in addition to code review) for all commits not fully under Apache-2.0
  - Currently a manual process
  - Would prefer to have checking IDs automatically

Case Study: Zephyr

Zephyr license processes:

- Each Zephyr source code file has a one-line SPDX-License-Identifier comment

```
/* SPDX-License-Identifier: Apache-2.0 */
```

Case Study: Zephyr

Zephyr license processes:

- Anything not under the project's Apache-2.0 license is in a separate "ext/" directory
  - Might not have SPDX-License-Identifier for these files
  - Keeping third party files unmodified makes it easier to refresh updates
  - Process for contributing is documented, and expectation that a README will provide appropriate licensing information as part of initial commit before it is accepted. Expectation is it will reflect any updated licensing. https://github.com/zephyrproject-rtos/zephyr/blob/master/doc/contribute/contribute_non-apache.rst

Case Study: Zephyr

Zephyr license details:

- Apache-2.0 license text in LICENSE file
- Details about choice of license, processes and use of DCO in CONTRIBUTING.rst file

Case Study: Zephyr

Zephyr license details:

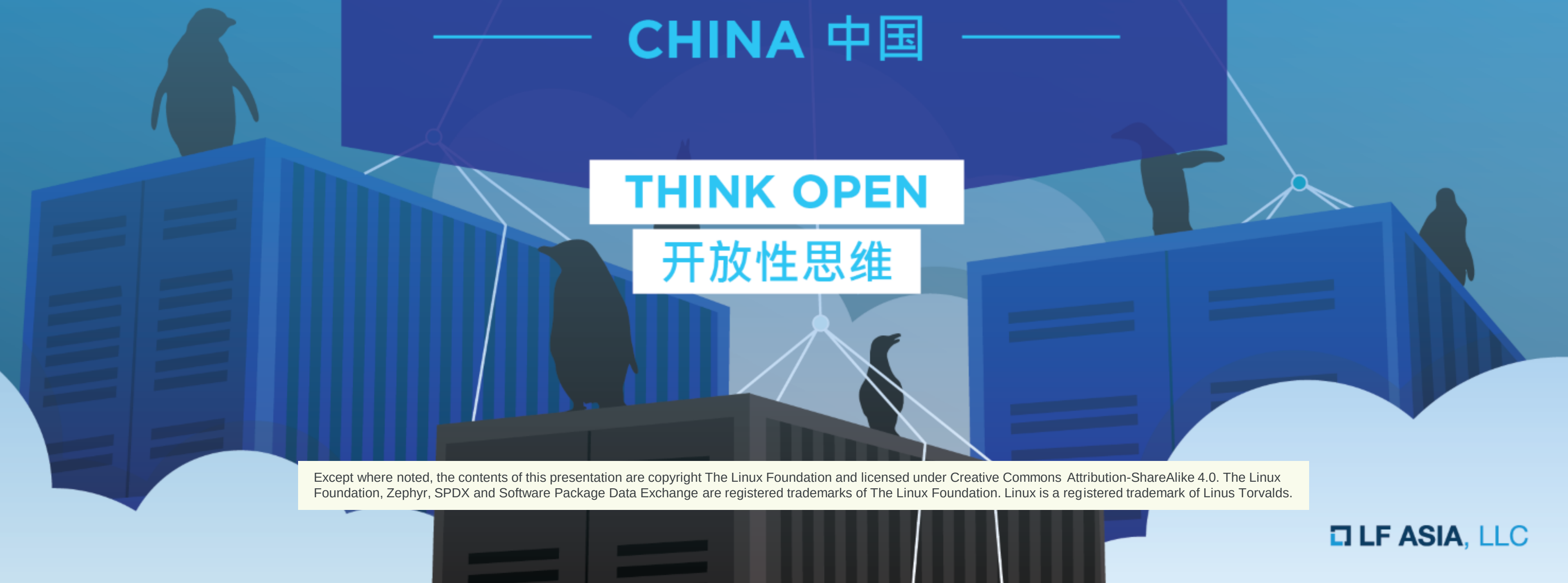
- Project page with clear details about non-Apache licenses in the codebase: http://docs.zephyrproject.org/LICENSING.html
- “SPDX-License-Identifiers” in all other files make it easy to auto-generate license details
- Will be generating .spdx file with first LTS release, and all releases after.

LINUXCON

containercon

CLOUDOPEN

CHINA 中国



THINK OPEN

开放性思维

LF ASIA, LLC