Sharing Open Source License and Copyright Data with SPDX

Martin Michlmayr, Hewlett-Packard

tbm@hp.com
Agenda

- The Problem
- Overview of the SPDX Spec
- Tools to Help
- SPDX and Distributions
- Current Status and Future
Open Source Compliance: The Challenge

Companies combine OSS with other software

Creating an accurate bill of materials requires effort & research

Your code
OSS Package
OSS Package
Outsourced SW
3rd party SW

Software Bill of Materials (BOM)
Open Source Compliance: The Challenge

The effort is repeated at each step in the supply chain.
SPDX - A standard format for communicating the components, licenses and copyrights associated with a software package.

Key pillar in Linux Foundation’s Open Compliance Program.
What it Means to You

**Embedded & SW Supply Chains**
- Save Time/Money
- Better Compliance

**Open Source Developers**
- Help Users Comply With Your Licenses

**Consumers of SW & OSS**
- Understand Licensing of the Code You Use
Participation is from a range of organizations and across various roles.
Working Group Operation

- The working group runs similarly to an open source project without centralized constitution or bylaws
- Intellectual property contributed by participants members is covered under the Creative Commons license (CC-BY-3.0)

Structure
- General Meeting and mailing list
- Teams: Technical, Business, Legal

Very inclusive process
- Self-subscription for interested participants
- Those willing to “do” can influence direction
- Mailing lists, wiki, phone calls, BOFs...
- http://spdx.org
Specification Goals

- A file format for license and copyright information to accompany packages
  - Guiding Principle: Just the facts – no interpretations

- A standardized short form to refer to the official version of common licenses

- Benefits
  - Allows easy exchange of license information between companies reducing burden on both suppliers and consumers
  - Avoids due diligence redundancy where the same source code package is analyzed multiple times by different receivers
  - Provides a unified method for exchanging license information
The SPDX File

Info about source of this SPDX™ file

Associates SPDX™ file with a software package (tarball, zip, archive)

Text of licenses that are not in SPDX™ standard list

License associated with each file. Refers to std.license list on SPDX™ or a non-standard license included with the file

Log of 3rd party reviews

File is in RDF/XML form; can be converted to tag value or spreadsheet.
Analysis Information

- SPDX Version (used in creation of SPDX file)
- How this info was generated
  - Manual review (who, when)
  - Tool (id, version, when)
- Creator (who created the SPDX file)
- Creator License (license of data in file)
Package Information

- Formal Name of Package (Full name given by originator and version information)
- Package File Name (Name package obtained under (.tar, .rpm, etc.))
- Package Checksum (to unambiguously map file to a package)
- Package Download Location (download URL)
- Licensing for Package
  - Declared License- License that has been asserted for the package
  - Concluded License- License that Creator has concluded
- Copyright Text
- Description of Package (optional)
This section is for licenses not on the standard list.

Aim for ~90% coverage with standard short forms; no plan to be exhaustive

Background: ~20 licenses cover most code; OSI recognizes 67 licenses as open source
The SPDX List of “Standard Licenses”

SPDX license repo

<table>
<thead>
<tr>
<th>License Identifier</th>
<th>Recognized Exceptions</th>
<th>Full name of License</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFL-3.0</td>
<td></td>
<td>Academic Free License 3.0</td>
</tr>
<tr>
<td>ACPL-3.0</td>
<td></td>
<td>GNU Affero General Public License v3</td>
</tr>
<tr>
<td>APL</td>
<td></td>
<td>Adaptive Public License</td>
</tr>
<tr>
<td>ASL-2.0</td>
<td></td>
<td>Apache License, 2.0</td>
</tr>
<tr>
<td>APSL-2.0</td>
<td></td>
<td>Apple Public Source License 2.0</td>
</tr>
<tr>
<td>Artistic-2.0</td>
<td></td>
<td>Artistic license 2.0</td>
</tr>
<tr>
<td>AAL</td>
<td></td>
<td>Attribution Assurance License</td>
</tr>
<tr>
<td>BSD-4-Clause</td>
<td></td>
<td>BSD 4-clause “Original” or “Old” License</td>
</tr>
<tr>
<td>BSD-3-Clause</td>
<td></td>
<td>BSD 3-clause “New” or “Revised” License</td>
</tr>
<tr>
<td>BSD-2-Clause</td>
<td></td>
<td>BSD 2-clause “Simplified” or “FreeBSD” License</td>
</tr>
<tr>
<td>BSL-1.0</td>
<td></td>
<td>Boost Software License 1.0</td>
</tr>
<tr>
<td>CATOSL-1.1</td>
<td></td>
<td>Computer Associates Trusted Open Source License 1.1</td>
</tr>
<tr>
<td>CC-BY-1.0</td>
<td></td>
<td>Creative Commons Attribution 1.0</td>
</tr>
<tr>
<td>CC-BY-NC-1.0</td>
<td></td>
<td>Creative Commons Attribution Non Commercial 1.0</td>
</tr>
<tr>
<td>CC-BY-ND-1.0</td>
<td></td>
<td>Creative Commons Attribution No Derivatives 1.0</td>
</tr>
<tr>
<td>CC-BY-SA-1.0</td>
<td></td>
<td>Creative Commons Attribution Share Alike 1.0</td>
</tr>
<tr>
<td>CC-BY-NC-SA-1.0</td>
<td></td>
<td>Creative Commons Attribution Non Commercial Attribution Share Alike 1.0</td>
</tr>
<tr>
<td>CC-BY-2.0</td>
<td></td>
<td>Creative Commons Attribution 2.0</td>
</tr>
<tr>
<td>CC-BY-NC-2.0</td>
<td></td>
<td>Creative Commons Attribution Non Commercial 2.0</td>
</tr>
<tr>
<td>CC-BY-ND-2.0</td>
<td></td>
<td>Creative Commons Attribution No Derivatives 2.0</td>
</tr>
<tr>
<td>CC-BY-SA-2.0</td>
<td></td>
<td>Creative Commons Attribution Share Alike 2.0</td>
</tr>
<tr>
<td>CC-BY-NC-SA-2.0</td>
<td></td>
<td>Creative Commons Attribution Non Commercial Attribution Share Alike 2.0</td>
</tr>
</tbody>
</table>

List of most common licenses (100+)
Include common exceptions
Standardized license names
Exact text of licenses
Available on SPDX website – URLs won’t change

Copyright Linux Foundation 2011
File Information

- File Name
- File Type (source, binary, archive)
- File Check Sum
- Concluded License (license determined by SPDX file creator)
- Copyright Text
- Artifact Project Name (from which project it came)
Review

- Reviewer
- Review Date
- Review Comment
Objectives

- Reduce the effort of creating, consuming and validating SPDX documents
- Provide a translation from the technical document (e.g. RDF/XML or tag-value format) and a more readable format
- Provide a mechanism for validating SPDX documents
- Enable contributions and review of the tool implementation by the broader technical community through open source licensing
SPDX and Distros

- Distros do a lot of license review when adding a new open source package
- No unified format used by distros (e.g. Debian uses free-form debian/copyright; DEP5 proposal for machine-readable debian/copyright)
- SPDX (or derived format) would allow more collaboration between distros: would reduce work for everyone, make it easier to find potential problems, etc
- Distros are a good way of working with upstream: getting SPDX adopted by open source projects directly.
How to indicate cruft – files not relevant for package licensing obligations

How to model hierarchical / associated packages

License of SPDX files: currently PDDL; CC0 with disclaimer under discussion

Develop guidelines for templatizing of licenses

Adoption of SPDX: SPDX as required standard in supply chain

Adoption of SPDX in upstream projects