



Describing License Information in SPDX

It's Easier than you think

Gary O'Neall, Source Auditor Inc.

- What is SPDX, who is using it, and why?
- The complex – examples of SPDX documents
- A simple start – using SPDX license ID's
 - Overview of the SPDX license list
 - Including license identifiers in your source code
 - Tools to support using license IDs
- A bit more work – SPDX documents for original source
 - Example of an SPDX document
 - Tools to support creating SPDX document
- Advanced SPDX usage – automating your supply chain
- Questions



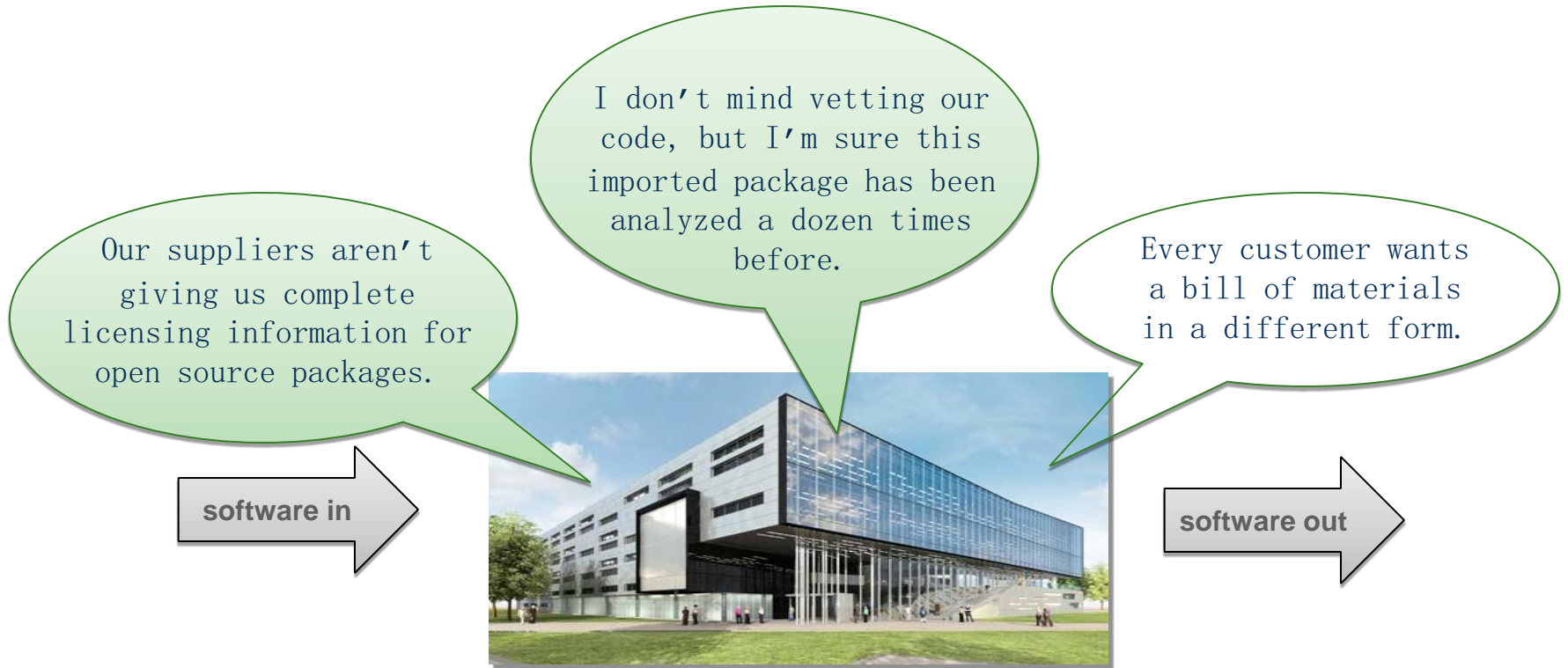
What is SPDX?

- Standard:
 - A standard format for communicating the components, licenses and copyrights associated with a software package
- Guiding principles:
 - Human and machine readable
 - Focus on capturing facts; avoid interpretations
- Vision:
 - To help reduce redundant work in determining software license information and facilitate compliance



Who is using SPDX?

- SPDX has over 20 active participants
- Active participation by organizations in the middle of the software supply chain (e.g. TI, HP, Wind River, Samsung)
- Tools providers are incorporating SPDX for import and export of software “Bill of Materials”
- Consultants using the format as a standard output
- Open source projects using the SPDX license identifiers (e.g. Das U-Boot, Ruby projects, GitHub)





Example SPDX document Formats

Tag/Value

```
##-----
## Package Information
##-----

PackageName: time-1.7.tar.gz
PackageFileName: time-1.7.tar.gz
PackageDownloadLocation: NOASSERTION
PackageVerificationCode: dd5cf0b17bfef4284c6c22471b277de7beac407c
PackageChecksum: SHA1: dde0c28c7426960736933f3e763320680356cc6a
PackageLicenseConcluded: GPL-2.0+
PackageLicenseDeclared: GPL-2.0+
PackageLicenseInfoFromFiles: GPL-2.0
PackageLicenseInfoFromFiles: GPL-2.0+
PackageLicenseInfoFromFiles: MIT
PackageLicenseInfoFromFiles: LicenseRef-1
PackageLicenseInfoFromFiles: LicenseRef-2
PackageLicenseInfoFromFiles: LicenseRef-3
PackageCopyrightText: NOASSERTION

##-----
```

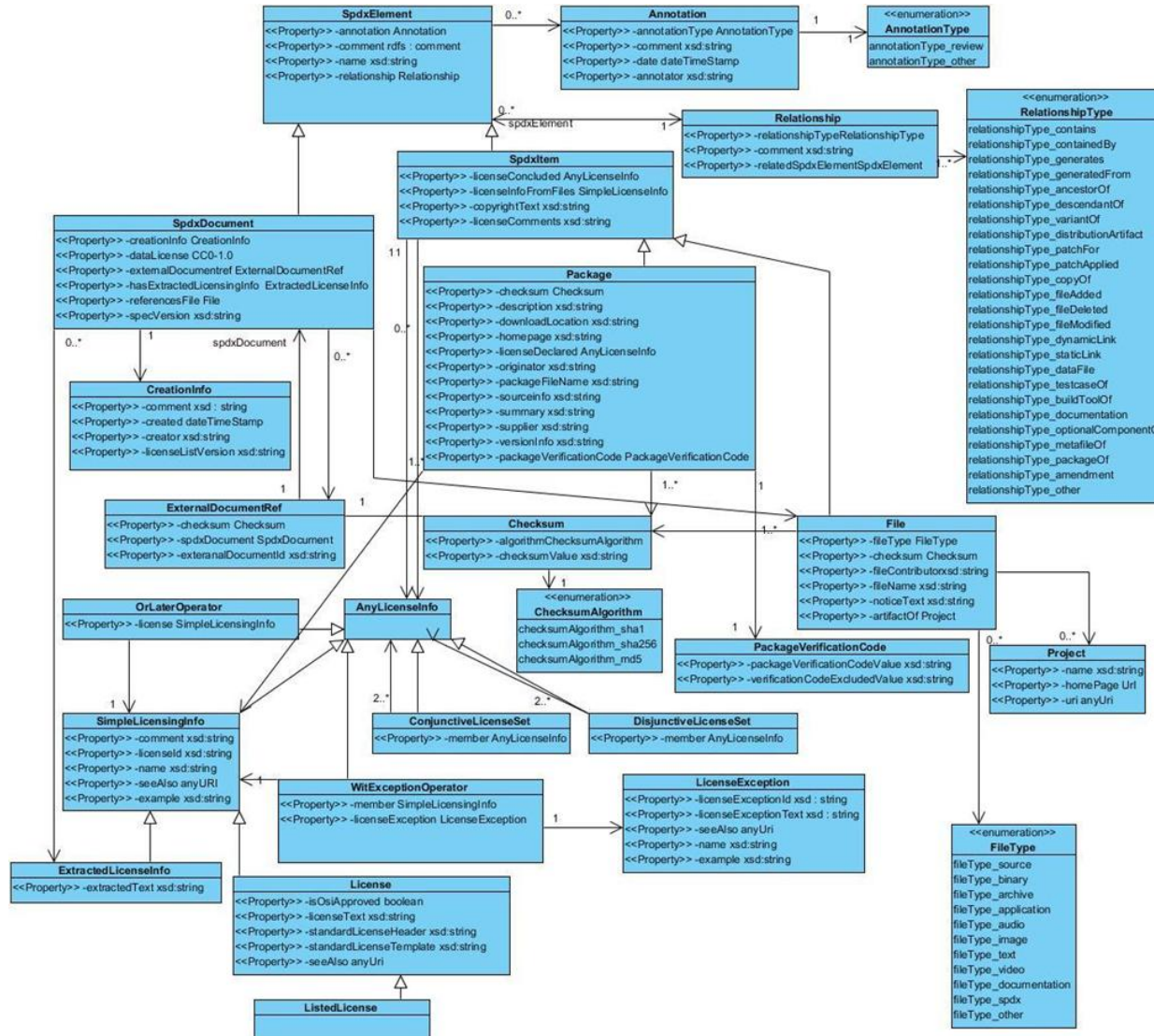
```
- <rdf:Description rdf:nodeID="A0">
  <rdfs:comment/>
  - <copyrightText>
    * <I>Copyright</I> (C) 1999 by Randolph Chung &
  </copyrightText>
  <licenseComments/>
  <licenseInfoInFile rdf:resource="http://spdx.org/lice
  <licenseConcluded rdf:resource="http://spdx.org/rdf
  <fileType rdf:resource="http://spdx.org/rdf/terms#file
  <checksum rdf:nodeID="A1">
  <fileName>networking/hostname.c</fileName>
  <rdf:type rdf:resource="http://spdx.org/rdf/terms#Fi
  </rdf:Description>
- <rdf:Description rdf:nodeID="A2">
  <checksumValue>dc90a437e03f31ab04e7059d8da5f88b28cde77d</checksumValue>
  <algorithm rdf:resource="http://spdx.org/rdf/terms#checksumAlgorithm_sha1"/>
  <rdf:type rdf:resource="http://spdx.org/rdf/terms#Checksum"/>
  </rdf:Description>
- <rdf:Description rdf:nodeID="A3">
  <rdfs:comment/>
  <copyrightText rdf:resource="http://spdx.org/rdf/terms#none"/>
  <licenseComments/>
```

RDF

Spreadsheet

6.1 File Name	6.2 File Type	6.3 File Checksum	6.4 License Concluded	6.5 License Info in File	6.6 L
time-1.7.AUTHORS	OTHER	7961F4CEFFD030CE1FED DF7BBA22523CC1A67F	NOASSERTION	NONE	
time-1.7.ChangeLog	OTHER	4A872E2C972E38B502B228 37C26136CC2647339	NOASSERTION	NONE	
time-1.7.configure	OTHER	A54A6E0A73D19F322E7E71 A5F0E23B59F2B8B15	LicenseRef-3	LicenseRef-3	
time-1.7.configure.in	OTHER	63F77F8E8E9083E8FD0E4 BE3B856872A06358BF	NOASSERTION	NONE	
time-1.7.COPYING	OTHER	0750269958584BBD0E4B526F 5C4D3B817E0D3558	GPL-2.0	GPL-2.0	
time-1.7.error.c	SOURCE	37BFDB644AE05071B7FA89 BD48B2110CB8D3E12D	GPL-2.0+	GPL-2.0+	
time-1.7.getopt.c	SOURCE	4EEC2F371CDEA3FA5F96A3 7B44820244898C75	GPL-2.0+	GPL-2.0+	
time-1.7.getopt.h	SOURCE	512168AACCC0C10F2E2D E76A4AF347AACC05	GPL-2.0+	GPL-2.0+	
time-1.7.getopt1.c	SOURCE	177C2F08A07203F875A6E 3C0C02F7B93C9F602	GPL-2.0+	GPL-2.0+	
time-1.7.getpagesize.h	SOURCE	1EF16700B723876F8322695 B81AFC2A5E18C8D0	NOASSERTION	NONE	
time-1.7.INSTALL	OTHER	BD0CE867F86293AFC2069 E8B8A842ARC3E238C	NOASSERTION	NONE	

The Big (and complex) Picture





Starting Simple: SPDX License List Identifiers

[projects](#) / [u-boot.git](#) / [blob](#)

[summary](#) | [shortlog](#) | [log](#) | [commit](#) | [commitdiff](#) | [tree](#)
[history](#) | [raw](#) | [HEAD](#)

Merge branch 'master' of [git://git.denx.de/u-boot-mpc85xx](#)

[\[u-boot.git\]](#) / [post](#) / [post.c](#)

```
1 /*
2  * (C) Copyright 2002
3  * Wolfgang Denk, DENX Software Engineering, wd@denx.de.
4  *
5  * SPDX-License-Identifier:      GPL-2.0+
6  */
7
```




Advantages to License ID's

- Reduces effort to determine the actual license used in a source file
- List of (common) open source* licenses
 - Over 300 licenses and over 20 common exceptions
- spdx.org/licenses contains the text, reference URL's, templates for matching, whether OSI approved and headers
- Backed by an active organization which maintains the license list
- Matching guidelines to help determine if the license text matches the text



More Info on the License List

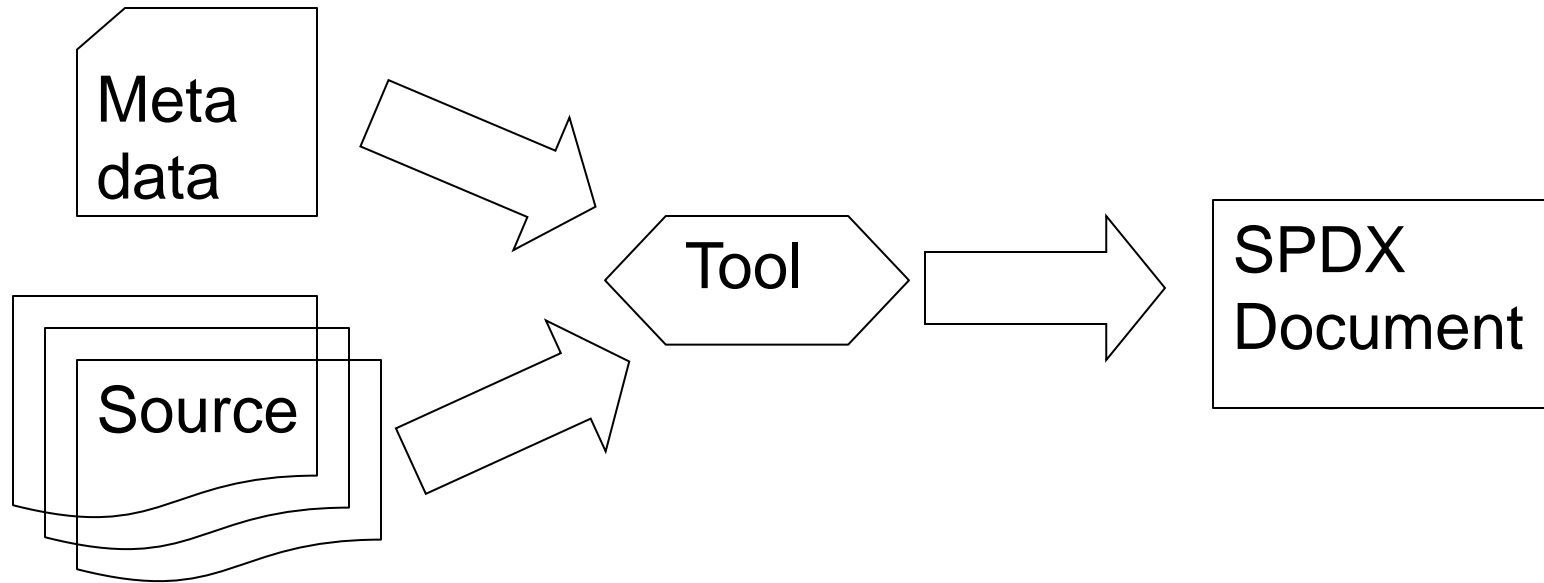
- Human and machine readable at <http://spdx.org/licenses>
- RDFa machine readable access
- JSON file at <http://spdx.org/licenses/licenses.json>
- For tools support, see the tech report “Accessing SPDX Licenses”: <http://spdx.org/publications/tool-documentation/accessing-spdx-licenses>
- License ID's in Source:
http://wiki.spdx.org/view/Technical_Team/SPDX_Meta_Tags



A bit more work: Using License Expressions

- Expresses a combination of licenses (e.g. “Dual Licensing”)
- AND – Conjunctive licenses
- OR – Disjunctive licenses
- With – Exceptions
- + - Or later versions

- Example: GPL-2.0+ OR MIT – Early versions
jQuery
- Example: GPL-2.0-with-autoconf-exception AND
GPL-3.0+ AND LGPL-2.1+ AND GPL-2.0+
- Example: (GPL-2.0+ OR MIT) AND BSD-3-Clause



- Requires some tooling since the verification code is generated by checksums



Tools to Support SPDX Creation

- Wind River website (spdx.windriver.org)
- FOSSology for analyzing licenses (<https://fossologyspdx.ist.unomaha.edu/>)
- Maven plug-in (<https://github.com/goneall/spdx-maven-plugin>)
- Yocto – UNO built SPDX extension
- Eclipse Plugin's under development (https://github.com/TCF-30/SPDX_Eclipse_Plugin and <https://github.com/goneall/SpdxEclipsePlugin>)
- FOSSology (SPDX support in beta) (<http://www.fossology.org/projects/fossology/wiki/WhoUsesFOSSology>)
- Others... - See spdx.org/tools

- Leverages existing meta data
- Additional configuration parameters for SPDX specifics
- Builds relationships based on dependencies
- Can be included in a “parent POM” for a larger open source project

```
<plugins>
  <plugin>
    <groupId>org.spdx</groupId>
    <artifactId>spdx-maven-plugin</artifactId>
    <executions>
      <execution>
        <id>build-spdx</id>
        <goals>
          <goal>createSPDX</goal>
        </goals>
      </execution>
    </executions>
    <configuration>
      <spdxDocumentNamespace>
        http://spdx.org/documents/spdx-toolsv2.0-rc1
      </spdxDocumentNamespace>
      <defaultFileComment>test file comment
      </defaultFileComment>
      <documentComment>Initial submission for the SPDX
document bake-off</documentComment>
      <documentAnnotations>
        <param>
          <annotator>Person: Gary O'Neill</annotat
          <annotationDate> 2015-07-23T18:30:22Z
          </annotationDate>
          <annotationType>OTHER</annotationType>
          <annotationComment>Initial submission fo
the SPDX 2.0 document bake-off
          </annotationComment>
        </param>
      </documentAnnotations>
    </configuration>
  </plugin>
</plugins>
```



Advantages to SPDX Documents

- Can be very descriptive
 - In addition to licenses, contains copyrights, relationships, creation and origin information
- Can be easily translated to human readable form (e.g. a spreadsheet)
- Some verification supported (package verification code, checksums)
- Integrates with various license compliance tools
- REALLY helps the downstream consumers – changes from identification to (maybe) verification



Tips for Keeping it Simple: Not all fields are required

Document Creation Information

- 2.1 SPDX Version.
- 2.2 Data License
- 2.3 SPDX Identifier
- 2.4 Document Name
- 2.5 SPDX Document Namespace
- 2.8 Creator
- 2.9 Created

1 per document

Package Information

- 3.1 Package Name
- 3.2 Package SPDX Identifier
- 3.7 Package Download Location
- 3.8 Package Verification Code
- 3.12 Concluded License
- 3.13 All Licenses Information from Files.
- 3.14 Declared License.
- 3.16 Copyright Text

1 per package described in document

1 per file in each package

File Information

- 4.1 File Name
- 4.2 File SPDX Identifier
- 4.4 File Checksum
- 4.5 Concluded License
- 4.6 License Information in File
- 4.8 Copyright Text

- Beware of over-complicating the relationships
- Refer to best practices:
http://wiki.spdx.org/view/Technical_Team/Best_Practices
- Refer to other example SPDX documents, such as the results of the SPDX “bake-off” results
- Monitor the spdx tech mailing lists



Example – SPDX Tools:

<rdf:RDF ...

<referencesFile>

<File rdf:nodeID="A2">

<fileName>resources/stdlicenses/GFD

<fileType rdf:resource="http://spdx.org/

<rdfs:comment></rdfs:comment>

<licenseComments></licenseComments>

<licenseInfoInFile rdf:resource="

<checksum>

```
SPDXVersion: SPDX-2.0
DataLicense: CC0-1.0
DocumentNamespace: http://spdx.org/documents/spdx-toolsv2.0-rc1
DocumentName: SPDX tools
SPDXID: SPDXRef-DOCUMENT
DocumentComment: <text>Initial submission for the SPDX 2.0 document
                 bake-off</text>
```

	A	B	C	D	
1	Package Name	SPDX Identifier	Package Version	Package FileName	Package Organization
2	SPDX Tools		1.2.3		
3					
	sources/stdlicenses/Aladdin	OTHER		SHA1: a6c1b0fe85fabc1f4091fe92418f8c868b2a99	
	sources/stdlicenses/Apache-1.0	OTHER		SHA1: 02b0b1329d48f6ccd8b8672b64a4e5cbbbd3	
	sources/stdlicenses/Apache-1.1	OTHER		SHA1: 62fd412692a0b480c4c12aaa3e70978c65at	
	sources/stdlicenses/Apache-2.0	OTHER		SHA1: 341c45dc91b2a5a8362c4f52bc236713638c	
	sources/stdlicenses/Artistic-1.0	OTHER		SHA1: 528752aa62a047aca0cdc0f08ddd415812bf4i	
	sources/stdlicenses/Artistic-1.0-Perl	OTHER		SHA1: 3d35d044ca8814765fa36c95f138850cde78f	
	sources/stdlicenses/Artistic-1.0-cl8	OTHER		SHA1: e91beda1dc616fdb9f906739b3dcb084708fa	
	sources/stdlicenses/Artistic-2.0	OTHER		SHA1: 9a57ad1c535cd4331127880825cc3807bdd	
	sources/stdlicenses/BSD-2-Clause	OTHER		SHA1: 422745c62dd4505bd0e083b1801e2aa59f5a	
	sources/stdlicenses/BSD-2-Clause-FreeBSD	OTHER		SHA1: 0aa11fe1c536c726009cb737282e5e5ddcf8!	
	sources/stdlicenses/BSD-2-Clause-NetBSD	OTHER		SHA1: e49cda4d4810c79e2696a3d5f86e534d6adc	
	sources/stdlicenses/BSD-3-Clause	OTHER		SHA1: 0971dc218eb6bd9f2816a9961d7d6c0f64b6i	
	sources/stdlicenses/BSD-3-Clause-Attribution	OTHER		SHA1: 10145ae9fadc88f43d281aea49a878d9eb23!	
	sources/stdlicenses/BSD-3-Clause-Clear	OTHER		SHA1: af286b1cb13070b290ddce244935e1b8854t	
	sources/stdlicenses/BSD-3-Clause-LBNL	OTHER		SHA1: 1a8ffb24315a417c41ebcb084d68cc03ded5;	
	sources/stdlicenses/BSD-4-Clause	OTHER		SHA1: 3dcbddbbef77f1abe42882090b2a99242903	
	sources/stdlicenses/BSD-4-Clause-UC	OTHER		SHA1: 4f34ab6358fbb67749b21c0d0fa44cac5d638	
	sources/stdlicenses/BSD-Protection	OTHER		SHA1: 987925513dad869c7c912ed4a25fc28cb1a2	



Using SPDX in the Software Supply Chain

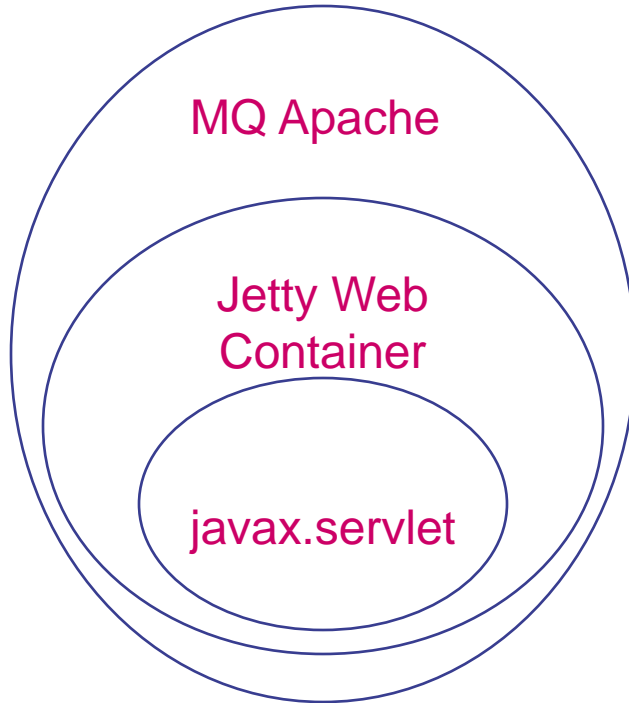
- Importing upstream SPDX files
 - Validating upstream SPDX files
 - Scanning and generating SPDX for packages without SPDX files
 - Creating output SPDX files
-
- Will typically involve multiple packages, multiple documents and relationships



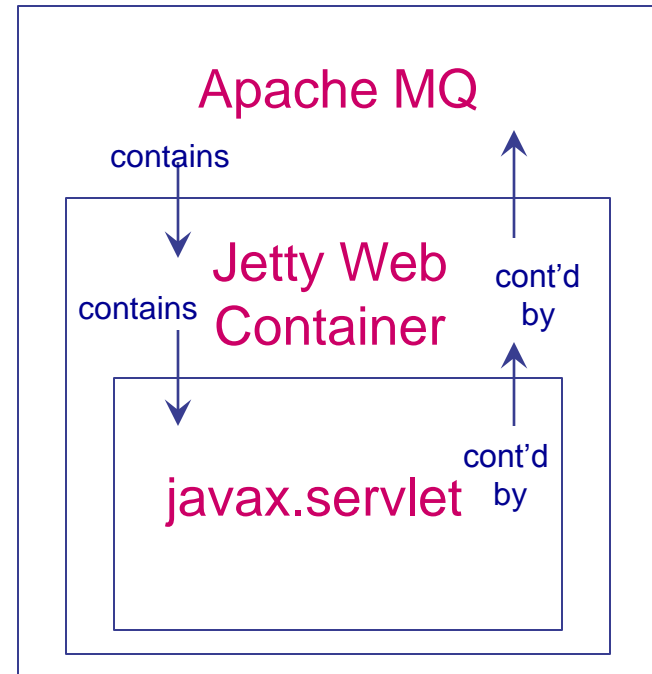
Advantages to using SPDX in the Software Supply Chain

- Maintain fidelity of the licensing data
- Reduce the effort for downstream consumers
- Reduce the effort for suppliers who are using SPDX

Package



SPDX Doc





Tips on keeping it simple in the supply chain

- Reference existing SPDX documents using external SPDX document relationships
- Choose a common “namespace” for your SPDX documents
- Use relationships and annotations to provide any corrections



More Information

- spdx.org
 - General information
- SPDX License list: spdx.org/licenses
- SPDX Tools: spdx.org/tools
- wiki.spdx.org
 - Workgroup wiki's for Technical, Business, and legal teams
 - Contains information on joining the mailing lists and calls



BACKUPS

- Open source (Apache 2.0 licensed) Java code providing access the SPDX Standard Libraries
- Binary .jar files available:
 - <http://spdx.org/spdx-tools/tools-from-the-spdx-workgroup>
- Source code available:
 - git.spdx.org – spdx-tools project
 - <https://github.com/spdx/tools>
- Bugs / enhancement ideas can be reported:
 - bugs.linuxfoundation.org – project SPDX/tools
 - Github issues tracking
- Contributions welcome!



More Information

- See the tech report “Accessing SPDX Licenses”:
<http://spdx.org/publications/tool-documentation/accessing-spdx-licenses>
- Join the spdx-tech mailing list
- Email me at gary@sourceauditor.com