SPDX Version 2.0 – Requirements and User Stories

SPDX-TR-2014-3

Abstract

The goal of this document is to provide a high-level view of the features targeted for v2.0 of the SPDX specification and to provide a mapping of use cases to features. Certain terminology in the document is taken from the Scrum software development methodology. Scrum is a form of agile software development. The use of the term Epic is one such example. An Agile Epic is a group of related user stories. A User story is an independent, testable requirement.

Keywords

SPDX, 2.0, Specification, Requirements, Technical Working Group

Kirsten Newcomer
Black Duc Software
knewcomer@blackducksoftware.com

Tech Report License

Creative Commons Attribution 3.0 (SPDX License ID CC-BY-3.0)

License Summary

You are free to:

- **Share** — copy and redistribute the material in any medium or format
- **Adapt** — remix, transform, and build upon the material
- for any purpose, even commercially

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

- **Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- **No additional restrictions** — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

- You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.
- No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

See http://creativecommons.org/licenses/by/3.0/legalcode for complete text of the license.

TABLE OF CONTENTS

1. Preface

Document Goal

The goal of this document is to provide a high-level view of the features targeted for v2.0 of the SPDX specification and to provide a mapping of use cases to features. Certain terminology in the document is taken from the Scrum software development methodology. Scrum is a form of agile software development. The use of the term Epic is one such example. An Agile Epic is a group of related user stories. A User story is an independent, testable requirement.

SPDX Mission and Charter

The Software Package Data Exchange (SPDX®) specification is a standard format for communicating the licenses and copyrights associated with a software package. The SPDX specification is created and managed through the Linux Foundation's SPDX workgroup. The grass-roots effort includes representatives from more than 20 organizations—software, systems and tool vendors, foundations and systems integrators—all committed to creating a standard for license and copyright data exchange for software packages.

For further information about the SPDX workgroup mission, charter and structure, as well as the current version of the SPDX specification, visit http://spdx.org.

You can learn more about the SPDX workgroup governance model here: http://spdx.org/about-spdx/governance.

The SPDX workgroup has published three versions of the SPDX standard: v1.0, 1.1 and 1.2. Open source and commercial tooling in support of the standard is available. To date, the standard has been adopted by a small, but significant number of organizations, including Intel, Wind River, and Samsung.

The SPDX workgroup is tracking and measuring adoption in three areas:

- Adoption and use of the SPDX License List
- Generation & consumption of SPDX documents
- Use of SPDX License IDs in software files

To learn more about how adoption is being measured, visit:
http://wiki.spdx.org/view/Business_Team/Adoption

Acknowledgements

A number of participants in the SPDX working groups have contributed ideas and words to this document, including Marshall Clow, Steve Cropper, Daniel German, Mathew German-Pray, Mark Gisi, Michael J. Herzog, Philip Koltun, Scott Lamons, Jilayne Lovejoy, Jack Manbeck, Kirsten Newcomer, Philip Odence, Gary O'Neall, Michel Ruffin, Bill Schineller, Kate Stewart, Ed Warnicke.

2. Change Log

Date	Description	Rationale	Who
November 25, 2013	Very early draft sent to handful of reviewers.	Determine whether document organization and content makes sense at a	Kirsten Newcomer

		high level.	
November 26, 2013	Incorporate changes from very early reviews. Key change: Refer to Package Relationships instead of Package Hierarchy.	The supply chain is more of a time ordered relationship, not necessarily a hierarchy.	Editor: Kirsten Newcomer Reviewers: Phil Odence, Gary O'Neall, Bill Schineller
December 11, 2013	Incorporate changes from additional early reviewers.	Up-level description of key features.	Jack Manbeck
January 15, 2014	Incorporate changes from Business Team review on 12/12/2013	Add more descriptive problem statement to Section 3. Clarify target users in section 4.1. Rework content on key features, section 4.2.	Editor: Kirsten Newcomer Reviewers: Business Team attendees.
January 21-22, 2014	Incorporate changes from Business Team review on 1/16/2014 Electronic comments included in this version for discussion.	Remove tables with links to individual use cases	Editor: Kirsten Newcomer Reviewers: Business team attendees.
February 18, 2014	Incorporate changes from Business Team review on 2/13/2014 and Tech Team input on 2/18/2014	Remove discussion comments. Debated whether to add back table of use cases and decided the link to the wiki is sufficient.	Editor: Kirsten Newcomer Reviewers: Business team & Tech team attendees.

3. Version 2.0 Goals

The primary goal of SPDX version 2.0 is to increase adoption, specifically in the generation and consumption of SPDX documents, through the support of key use cases that arise when managing larger portfolios and the relationships that are created as software packages progress through the supply chain.

Barriers to adoption that 2.0 will address are:

- The current version of the SPDX specification presents information about software packages in a single SPDX document. This fairly simple representation can become cumbersome when you need to express more complex scenarios that arise as software moves through the supply chain. For example, instead of being able to reference a previously created SPDX documents for code that is included in a larger distribution, users must create a new SPDX document that describes the collective code.

Further, the only way SPDX users have to track changes to licensing and copyright data for software elements as they move through the supply chain is through the use of comment fields.

Additional enhancement requests

- The current version of the SPDX specification includes data that can be used to verify that the files in the associated software package match the files analyzed when the SPDX document was created. It also includes fields to provide information about the creator and reviewers of the SPDX document. However, there is no verification mechanism provided for the creator and reviewer data.

- In versions 1.0, 1.1 and 1.2, the smallest unit of a copyrightable thing is a file. However, there are numerous examples of files with multiple copyrights and where individual copyrights apply to a subset of the file. While the current version of the specification allows me to document multiple copyrights for a single file, it does not allow me to associate an individual copyright with a specific subset of that file.

4. Features & Users

The use cases targeted for the 2.0 version of the SPDX specification have been collected in a series of open meetings, some held during Linux Collaboration Summits and some held as part of the regular SPDX workgroup meetings.

Full details of the use cases are documented here: http://wiki.spdx.org/view/Technical_Team/Use_Cases/2.0. A designation of OK indicates that the technical team has a sufficient understanding of the use case to move forward, and that the use case has a solid advocate.

Target Users

A supply chain can include the following categories of users:

- Upstream provider: Upstream represents the point of origin for a software package, such as an open source project.
- Midstream providers: Also, aggregators; intermediate packager. For example, those who combine code from one ore more open source packages, and may also include code of their own, for consumption by those further downstream. Midstream users are both recipients of SPDX documents and providers of SPDX documents.
- Distributors: Those who distribute a product that includes code from more than one upstream and/or midstream providers. The distribution may include their own code as well. Distributors are both recipients of SPDX documents and providers of SPDX documents.

SPDX v2.0 is focused on better meeting the needs of the Midstream and Distributor members of the supply chain.

Key Features

SPDX v2.0 includes two major features and two possible features, described below. One or more epics are included for each major feature.

Document, Package, and File Relationships:

- As a software package progresses through the supply chain, downstream recipients of the package may 1) use all or 2) only part of the package; recipients of the SPDX documents may then 3) add content to the package, 4) delete content, or 5) combine the content of multiple packages into one. In addition, recipients, who in turn become creators of SPDX docu-

ments, may 6) modify individual files within a package or 7) make other changes that activate license obligations.
    - **EPIC**: As a midstream provider or distributor, I want the SPDX specification to give me a way to represent relationships between SPDX documents so that I can provide an SPDX document for my new package without having to recreate all the data from the upstream provider in my new SPDX document.
    - **EPIC**: As a distributor I am building a product that passes through several entities in my supply chain. I need to know what software has been added, deleted and what individual files have been changed so that I can assess impact on licenses and license obligations.
- Midstream or upstream suppliers may bundle separate and distinct artifacts together to form an application or product.
    - **EPIC**: As a upstream or midstream provider I want to be able to say here is an application and it consists of these pieces and each piece has its own SPDX document but they are related in that they are provided as part of that application.
- To address these epics, SPDX v2.0 will provide a way to represent relationships between
    - SPDX documents
    - Software packages
    - Files within a software package
    - Files in different software packages

Implementation note: documenting relationships between elements requires unique ids for those elements.

Audit / Revision History

- **EPIC**: As a recipient of SPDX documents, I would like the specification to provide fields to capture information about how the contents of the SPDX document have changed over time so that I have a documented audit trail that I can share with all members of the supply chain.

Features that may be included

Signing

The current version of the SPDX specification includes data that can be used to verify that the files in the associated software package match the files analyzed when the SPDX document was created. It also includes fields to provide information about the creator and reviewers of the SPDX document. However, there is no verification mechanism provided for the creator and reviewer data.

- **EPIC**: When receiving an SPDX document, I would like to have a way to verify the creators and reviewers of the SPDX document. I would also like to provide my downstream recipients with a document they can verify.

Snippets

- **EPIC**: As a member of a supply chain working with such files, I want the SPDX specification to make it easier for me to clearly represent which lines in an individual file are associated with which copyright, so that more accurate and complete information about the copyrights for the file are available to me and other members of my supply chain.

Migration

A note about migration: Every data element in the SPDX 1.2 specification will have a home in the SPDX 2.0 specification. It is possible that some data elements will be deprecated, if it is determined that a different home is appropriate. Further, creators and recipients of SPDX documents that conform to previous versions of the specification may find that new homes are available for content that was previously presented in Comment fields.

What 2.0 is NOT

SPDX 2.0 does not attempt to address use cases directly related to license obligations or compliance.

5. High-level Model

The working model can be found here:

http://wiki.spdx.org/view/Technical_Team/Proposals/2012-02-01/Merged_Model_Proposal