

SPDX TECH REPORT

Accessing the SPDX License List

Abstract

The Software Products Data Exchange (SPDX) maintains the SPDX License List: a list of commonly found open source software licenses for the purposes of being able to easily and efficiently identify such licenses in an SPDX document. The SPDX License List includes a standardized short identifier, full name for each license, vetted license text, other basic information, and a canonical permanent URL. This document describes various methods of accessing the current and previous versions of the list. Several methods of accessing the license list are provided ranging from simply reading the license list on the website to programmatically accessing the license list information online.

This document is organized from the easiest method of access targeted at more casual users to the most sophisticated access targeted at software programmers.

This document does not go into any details on the criteria for what licenses are included on the SPDX License List, how to suggest new licenses, or how to determine if a license you find matches a license on the SPDX License List. Information on the license list process and additional information on SPDX licenses can be found at <http://spdx.org/licenses/> and links provided there

Version 1.0

Author: Gary O'Neill

gary@sourceauditor.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.

Contents

Abstract.....	1
Tech Report License	2
License Summary.....	2
You are free to:	2
Under the following terms:	2
Notices:	2
Reading Licenses on the SPDX License List Web Pages	4
URL.....	4
Table of Contents page	4
License List Version	4
List of Licenses	4
Accessing the Master Files for the SPDX License List on the SPDX Git Repository.....	4
Location.....	4
Tags for Versions.....	4
File Formats.....	4
Programmatically Accessing the Online License List using RDFa.....	5
What is RDFa?	5
RDFa Available Parsers.....	5
Forming the URLs.....	5
Listing the License ID's Available.....	5
RDFa Terms Used	6
Programmatically Accessing the Current SPDX License List using the SPDX Tools Library.....	6
Downloading the Library.....	6
Library APIs	6
Updating the SPDX License List in the Offline Cache	7
Examples	7
Accessing the Text for a Standard SPDX License.....	7
Determining if License Text is Equivalent to an SPDX Standard License Text	8
Accessing Older Versions of the License List	8

Reading Licenses on the SPDX License List Web Pages

If you would like to quickly view the SPDX License List, simply going to the SPDX License List page is the fastest approach. The SPDX License List page always contains the most recent published version of the license list.

URL

The web pages are located at <http://www.spdx.org/licenses>

Table of Contents page

The main index page contains a description of the SPDX License List, pointers to other relevant information, the license list version, and a list of licenses.

License List Version

The version of the license list can be found immediately prior to the list of licenses.

List of Licenses

The license table contains information on the licenses. The definition of the table columns can be found in the Explanation of SPDX License List Fields section on the [License List Overview webpage](#). Not all the fields are listed in this table. Clicking on the full name of each license (or the License Text link) takes you to an individual web page for each license on the SPDX License List which contains the information for all the fields.

Accessing the Master Files for the SPDX License List on the SPDX Git Repository

The web pages found at <http://www.spdx.org/licenses> are generated from the master files for the SPDX License List. The master files include both a spreadsheet listing all the licenses and fields; and the text for each license in a .txt file. All versions of the SPDX License List since version 1.19 are maintained in the [SPDX Git Repository](#).

Location

The summary information can be accessed at <http://git.spdx.org/?p=license-list.git;a=summary>. An archive file for any of the SPDX License List versions can be downloaded by following the "snapshot" link on the left hand side.

Tags for Versions

A tag is created for each SPDX License List once it has been released and published at <http://spdx.org/licenses>. The tag consists of "v" followed by the version number.

File Formats

There are two types of files maintained for each release of the license list:

- Spreadsheet: A spreadsheet file named `spdx_licenselist_vX.XX` where X.XX is the license list version. There are two formats stored, a Microsoft Excel (.xls) file and a Open Office Calc (.ods)

file. The current spreadsheet column definition can be found in the Explanation of License List Fields on the [License List Overview webpage](#). The spreadsheet may include additional columns (that are not official fields) for the purposes of tracking other work on or changes to the SPDX License List.

- Text files: The actual text for each license is contained in a text file (.txt) with a filename corresponding to the License Identifier. Some text files may include template information for the license as specified in the [License Matching Guidelines](#) and Appendix II License Matching Guidelines and Templates in the [SPDX specification version 1.2](#) or later.

Programmatically Accessing the Online License List using RDFa

To allow easy, well-defined programmatic access to the current SPDX Licenses List, the web pages are encoded using a [W3C](#) standard RDFa. Using RDFa parsers, a program can access SPDX License List information in a more reliable fashion than manually parsing the HTML pages themselves. The RDF terms used on the pages are also well-defined and documented in the [SPDX RDF terms page](#).

What is RDFa?

RDFa (Resource Descriptor Framework in Attributes) extends HTML (and XHTML) with attributes allowing RDF data to be easily extracted from the same web pages which can also be rendered using web browsers.

The following resources can provide you further information on RDFa:

- [Wikipedia article](#)
- [W3C RDFa primer](#)
- [Community website for RDFa](#) contains a number of resources for developers and users

RDFa Available Parsers

There are a number of RDFa parsers available for various languages. [The RDFa community page for developers](#) contains a list of libraries for JavaScript, PHP, Python, Ruby, C, and Java.

In addition to the libraries mentioned above, [Java-RDFa](#) used in conjunction with [Apache Jena](#) provides rich RDF functionality for Java applications. These are the libraries used by the [SPDX Workgroup Tools](#).

Forming the URLs

To retrieve details for a license on the SPDX License List, access the HTML file at [http://spdx.org/licenses/\[licenseID\]](http://spdx.org/licenses/[licenseID]) where [licenseID] is the SPDX license identifier. Note that there is no file type appended.

Listing the License ID's Available

A list of license ID's can be accessed by querying the RDFa located at <http://spdx.org/licenses/index.html> for resource license ID (<http://spdx.org/rdf/terms#licenseID>).

RDFa Terms Used

The following RDF terms are used in both the index.html and individual license pages:

- <http://spdx.org/rdf/terms#licenseId> - A human readable short form license. The license Id must be of the form "LicenseRef-"[idString] where [idString] is a unique string containing letters, numbers, ".", "-", or "+".

The following RDF terms are used on the individual license pages:

- <http://www.w3.org/2000/01/rdf-schema#label> or <http://spdx.org/rdf/terms#name> - License name (rdfs:label will be replaced with spdx:name on the next version of the license list)
- <http://spdx.org/rdf/terms#licenseText> - The full text of the license
- <http://spdx.org/rdf/terms#standardLicenseHeader> - Text for a license notice as specifically delineated by the license or license appendix
- <http://spdx.org/rdf/terms#licenseTemplate> - License template (if applicable) which describes sections of the license which can be varied. See [License Matching Guidelines](#) and Appendix II: License Matching Guidelines and Templates section of the specification for format information.
- <http://spdx.org/rdf/terms#isOsiApproved> - Indicates if the OSI has approved the license
- <http://www.w3.org/2002/07/owl#sameAs> or <http://www.w3.org/2000/01/rdf-schema#seeAlso> - Cross reference URL pointing to additional copies of the license (owl:sameAs will be replaced by rdfs:seeAlso on the next version of the license list)

Programmatically Accessing the Current SPDX License List using the SPDX Tools Library

The SPDX Tools are Java utilities and libraries made available by the SPDX workgroup under the Apache 2.0 license. The SPDX Tools library contains a class `SPDXLicenseInfoFactory` which will parse SPDX license names, access the online SPDX License List using RDFa, and provide detailed information on the licenses. If the web pages are not accessible (or if a Java property is set to use only offline licenses), a local cache of license information will be used.

Downloading the Library

Source code for the SPDX libraries can be accessed in the Linux Git repository (git.spdx.org) in the `spdx-tools.git` project. Implementation versions and a changelog is maintained in the root project directory.

The binary libraries are also included in each of the [SPDX workgroup tools](#).

The binary jar files can also be downloaded from the [Maven Central Repository](#) using the group ID `org.spdx` and the artifact ID `spdx-tools`.

Library APIs

The primary interfaces to the SPDX License List are the following methods in the `org.spdx.rdfparser.SPDXLicenseInfoFactory` class:

- `String getLicenseListVersion()` - Return the license list version used by the SPDX tools library
- `SPDXStandardLicense getStandardLicenseById(String licenseId)` - Return a SPDX license associated with the license ID parameter
- `String[] getStandardLicenseIds()` - Return an array of all SPDX standard license ID's
- `isStandardLicenseId(String licenseId)` - Returns true if the license ID is associated with an SPDX standard license ID
- `SPDXLicenseInfo parseSPDXLicenseString(String licenseString)` - Parses a license string compliant with the specification and returns an object representing the licenses. The parser will handle standard license ID's, conjunctive license sets, disjunctive license sets, and non-standard license IDs.

The class `SPDXStandardLicense` contains fields for all the properties of a license on the SPDX License List (e.g. text, id, name, sourceUrls, template, standardLicenseHeader, comments).

The package `org.spdx.licenseTemplate` contains classes to manage the license templates used in SPDX licenses. See the JavaDocs for this package (located in the doc directory of the source code or in the javadoc jar file in Maven) for a description of the interface to this package. The primary interfaces are in the `org.spdx.licenseTemplate.SpdxLicenseTemplateHelper` class.

The class `org.spdx.compare.LicenseCompareHelper` contains methods to compare license text to determine if they are equivalent according to the license matching [guidelines](#) maintained by the SPDX workgroup. The method `boolean LicenseCompareHelper.isTextStandardLicense(SPDXStandardLicense license, String licenseText)` will return true if the `licenseText` is equivalent to a license on the SPDX License List considering both the license template replaceable/optional text as well as the license matching guidelines.

Updating the SPDX License List in the Offline Cache

By default, the SPDX Tools Library will access the current SPDX License List web pages using RDFa to acquire the license information. If there is no Internet access to <http://spdx.org/licenses> or if the configuration property `OnlyUseLocalLicenses` is set to true, a local offline cache of the licenses is used.

The license cache can be found in the `resources/stdlicenses` directory in the same directory where the library is installed. The files in this directory are exact copies of the HTML files found on the <http://spdx.org/licenses> web pages. If you wish to update the licenses, these files can be replaced.

Also within the `resources/stdlicenses` directory is a file `licenses.properties`. This file contains the configuration property `OnlyUseLocalLicenses`. Editing this text file and setting `OnlyUseLocalLicenses=true` will always use the offline cache.

Examples

Accessing the Text for a Standard SPDX License

To print out the text for Apache 1.1 to the console:

```

public static void main(String[] args) {
    try {
        SPDXStandardLicense apache11 = SPDXLicenseInfoFactory
            .getStandardLicenseById("Apache-1.1");
        System.out.println(apache11.getText());
    } catch (InvalidSPDXAnalysisException e) {
        System.out.println("Error getting Standard License: " +
            e.getMessage());
    }
}

```

Determining if License Text is Equivalent to an SPDX Standard License Text

To determine if the license text in the string licenseTxt is equivalent to the license text for the BSD 3 clause license:

```

public static void main(String[] args) {
    String licenseText = "...";
    try {
        SPDXStandardLicense bsd3clause = SPDXLicenseInfoFactory
            .getStandardLicenseById("BSD-3-Clause");
        boolean matches = LicenseCompareHelper
            .isTextStandardLicense(bsd3clause, licenseText);
        if (matches) {
            System.out.println("Matches");
        } else {
            System.out.println("Does not match");
        }
    } catch (InvalidSPDXAnalysisException e) {
        System.out.println("Error getting lic: "+e.getMessage());
    } catch (SpdxCompareException e) {
        System.out.println("Error comparing: "+e.getMessage());
    }
}

```

Accessing Older Versions of the License List

An archive of the older versions of the HTML pages for the SPDX License List, starting with version 1.17, can be found at http://spdx.org/licenses/archive/archived_ll_vX.XX where vX.XX is the version of the license list.