

Use case 1:

Importing curation data to clear a software package of the same version

Use case 1:

Importing curation data to clear a software package of the same version
Open Source Automation Development Lab (OSADL) eG



OSSelot – The Open Source Curation Database

<https://github.com/Open-Source-Compliance/package-analysis>

- Contains **license and copyright analysis results** for certain packages (created with FOSSology):
 - Readme with metadata of the package, *e.g.* download location, comments
 - SPDX Tag:Value file with concluded licenses, copyright notices and comments on decisions
 - Disclosure document with aggregated license texts, copyright notices and acknowledgments

Use case 1:

Importing curation data to clear a software package of the same version
Open Source Automation Development Lab (OSADL) eG



Reusing curation data (FOSSology)

- Example: **OpenSSL v3.0.3**
- Clearing a software package with FOSSology **reusing existing curation data** involves the following steps:
 1. Converting SPDX Tag:Value file into SPDX RDF format.
 2. Uploading curated source packages into FOSSology.
 3. Importing the SPDX RDF file into the corresponding package in FOSSology.
 4. Clearing any remaining files.

Use case 1:

Importing curation data to clear a software package of the same version
Open Source Automation Development Lab (OSADL) eG



1. Converting SPDX Tag:Value file into SPDX RDF format

- Make sure that the root directories given in the SPDX TV files under "FileName:" is identical to the root directories of the current packages, otherwise the import in step 3 will not work.
- Download the current version of the SPDX tools:
<https://github.com/spdx/tools-java>
- Convert the TV file into an RDF file:

```
$ java -jar tools-java-1.0.4-jar-with-dependencies.jar \  
    Convert [SPDXTV.tag] [SPDXRDF.rdf]
```

Use case 1:

Importing curation data to clear a software package of the same version
Open Source Automation Development Lab (OSADL) eG



2. Upload curated packages into FOSSology



To manage your own group permissions go into **Admin > Groups > Manage Group Users**. To manage permissions for this one upload, go to **Admin > Upload Permissions**.

This option permits uploading a single file (which may be iso, tar, rpm, jar, zip, bz2, msi, cab, etc.) or a directory from a remote web or FTP server to FOSSology. The file or directory to upload must be accessible via a URL and must not require human interaction such as login credentials.

1. Select the folder for storing the uploaded files:

Use Case 1 ▾

2. Enter the URL to the file or directory:

3. (Optional) Enter a viewable name for this file or directory:

Note: If no name is provided, then the uploaded file (directory) name will be used.

4. (Optional) Enter comma-separated lists of file name suffixes or patterns to accept:

5. (Optional) Enter comma-separated lists of file name suffixes or patterns to reject:

6. (Optional) maximum recursion depth (inf or 0 for infinite):

7. (Optional) Enter a description of this file:

8. Apply global decisions for current upload ⓘ

9. Ignore SCM files (Git, SVN, TFS) and files with particular Mime-type ⓘ

cont'd:

10. Visible only for active group ⓘ
 Visible for all groups ⓘ
 Make Public ⓘ

11. Select optional analysis:

- Copyright/Email/URL/Author Analysis
 ECC Analysis, scanning for text fragments potentially relevant for export control
 Keyword Analysis
 MIME-type Analysis (Determine mimetype of every file. Not needed for licenses or buckets)
 Monk License Analysis, scanning for licenses performing a text comparison
 Nomos License Analysis, scanning for licenses using regular expressions
 Ojo License Analysis, scanning for licenses using SPDX-License-Identifier
 Package Analysis (Parse package headers)
 REUSE.Software Analysis (forces *Ojo License Analysis*)
 Software Heritage Analysis

12. Automatic Concluded License Decider ⓘ, based on

- Scanners matches if all Nomos findings are within the Monk findings
 Scanners matches if Ojo or REUSE.Software findings are no contradiction with other findings
 Bulk phrases from reused packages
 New scanner results, i.e., decisions were marked as work in progress if new scanner finds additional licenses

13. (Optional) Reuse ⓘ

- Select an already uploaded package for reuse in specific folder
- Enhanced reuse (slower) ⓘ
 Reuse main license/s ⓘ
 Reuse report configuration settings ⓘ
 Reuse deactivated copyrights ⓘ

Upload to reuse:

Upload

Before clearing

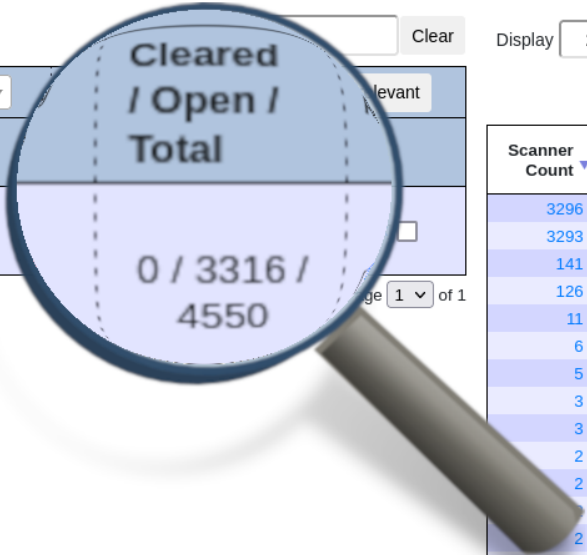
Folder: [Software Repository/ Use Case 1/](#)
[OpenSSL v3.0.3/OpenSSL v3.0.3](#)

[Software Heritage](#) | [License Browser](#) | [File Browser](#) | [Spasht](#) | [Copyright](#) | [ECC](#) | [Email/URL/Author](#) | [Keyword](#) | [Browse](#) | [Export List](#) | [Search](#) | [Bucket](#) | [View](#) | [Conf](#) | [Info](#) | [Refresh](#)

Display files (tree view or flat)

Files	Scanner Results (N: nomos, M: monk, Nk: ninka, I: reportImport, O: ojo, Sp: spasht, Rs: reso)	Edited Results
openssl-openssl-3.0.3	Apache-2.0, Artistic-1.0, Artistic-1.0-Perl, BSD-2-Clause, BSD-3-Clause, BSD-Source-Code, CC0-1.0, Cryptogams, Dual-license, GPL, GPL-1.0, GPL-1.0+, GPL-2.0, GPL-2.0+, LGPL-2.1+, MPL-1.1, No_license_found, OpenSSL, Perl-possibility, Public-domain, RSA-possibility, See-doc.OTHER, See-file	

Showing 1 to 1 of 1 files



Display licenses

Scanner Count	Concluded License Count	License Name
3296	0	Apache-2.0
3293	0	OpenSSL
141	0	Dual-license
126	0	Cryptogams
11	0	BSD-2-Clause
6	0	Artistic-1.0-Perl
5	0	GPL-1.0+
3	0	GPL
3	0	BSD-Source-Code
2	0	See-file
2	0	Perl-possibility
2	0	MPL-1.1
2	0	LGPL-2.1+
2	0	GPL-2.0+
2	0	BSD-3-Clause
1	0	See-doc.OTHER
1	0	RSA-possibility
1	0	Public-domain
1	0	GPL-2.0
1	0	GPL-1.0
1	0	CC0-1.0
1	0	Artistic-1.0

Showing 1 to 22 of 22 licenses

Page of 1

- 3316 files contain license information and must be cleared.

3. Import the SPDX RDF file into the corresponding package in FOSSology



[Home](#) [Search](#) [Browse](#) [Upload](#) [Jobs](#) [Organize](#) [Admin](#) [Help](#)

Report Import

Version: [4.0.0.0], Branch: [HEAD], Commit: [#4b7556] 2022/01/21 08:43 UTC built @ 2022/01/24 08:42 UTC

1. Select the folder that contains the upload:

2. Select the upload you wish to edit:

3. Select report to upload: openssl-3.0.3-SPDX2RDF.rdf

4. Select how the information should be imported:

- Create new licenses as
 - license candidate
 - new license
- Add the License Info as findings from
 - SPDX tag of type licenseInfoInFile
 - SPDX tag of type licenseConcluded
- Add concluded licenses as decisions
 - also overwrite existing decisions
 - import as "to be discussed"
- Add the copyright information as textfindings

Issues with Import Report

Importing the RDF file into FOSSology still has some issues:

- Matching of files happens via full file path and not checksum
→ requires identical root directory
- Import loses the tag "LicenseComment" which holds the explanation for certain decisions of "LicenseConcluded".
- Files with "LicenseInfoInFile: [some value]" and "LicenseConcluded: NOASSERTION" must be cleared manually.
- Curated copyright notices are only imported as "text findings".

Use case 1:

Importing curation data to clear a software package of the same version
Open Source Automation Development Lab (OSADL) eG



4. Clearing any remaining files

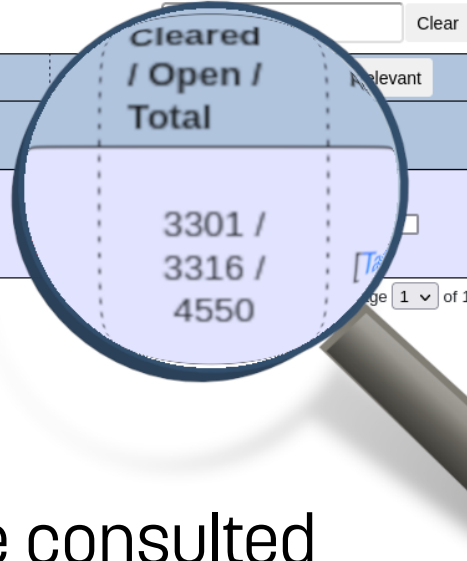
Folder: Software Repository/ Use Case 1/
OpenSSL v3.0.3/OpenSSL v3.0.3

Software Heritage | License Browser | File Browser | Spasht | Copyright | ECC | Email/URL/Author | Keyword | Browse | Export List | Search | Bucket | View | Conf | Info | Refresh

Display 50 files (tree view or flat)

Files	-- filter for scan results --	-- filter for edited results --	relevant
openssl-openssl-3.0.3	Scanner Results (N: nomos, M: monk, Nk: ninka, I: reportImport, O: ojo, Sp: spasht, Rs: reso)	Edited Results	
	Apache-2.0, Artistic-1.0, Artistic-1.0-Perl, BSD-2-Clause, BSD-3-Clause, BSD-Source-Code, CC0-1.0, Cryptogams, Dual-license, GPL, GPL-1.0, GPL-1.0+, GPL-2.0, GPL-2.0+, LGPL-2.1+, License-of-GNU-Licenses, MPL-1.1, No_license_found, OpenSSL, Perl-possibility, Public-domain, RSA-possibility, See-doc.OTHER, See-file	License-of-GNU-Licenses, BSD-2-Clause, Cryptogams, Apache-2.0, Public-domain, Artistic-1.0-Perl, CC0-1.0, GPL-1.0+, Dual-license, BSD-3-Clause	

Showing 1 to 1 of 1 files



Display 25 licenses

Scanner Count	Concluded License Count	License Name
3297	3298	Apache-2.0
3293	0	OpenSSL
142	138	Dual-license
126	126	Cryptogams
16	16	BSD-2-Clause
12	11	Public-domain
6	3	GPL-1.0+
6	3	Artistic-1.0-Perl
3	3	BSD-3-Clause
3	0	GPL
3	0	BSD-Source-Code
2	0	See-file
2	0	Perl-possibility
2	0	MPL-1.1
2	0	LGPL-2.1+
2	0	GPL-2.0+
1	1	License-of-GNU-Licenses
1	1	CC0-1.0
1	0	See-doc.OTHER
1	0	RSA-possibility
1	0	GPL-2.0
1	0	GPL-1.0
1	0	Artistic-1.0

Showing 1 to 23 of 23 licenses

Page 1 of 1

- Only 15 of 3316 files (<0.5 %) must still be cleared manually!
- The curated SPDX TV file can be consulted for explanations of clearing decisions (tag "LicenseComments:").