Use case 1:
Importing curation data to clear a software package of the same version

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

Reusing curation data (FOSSology)

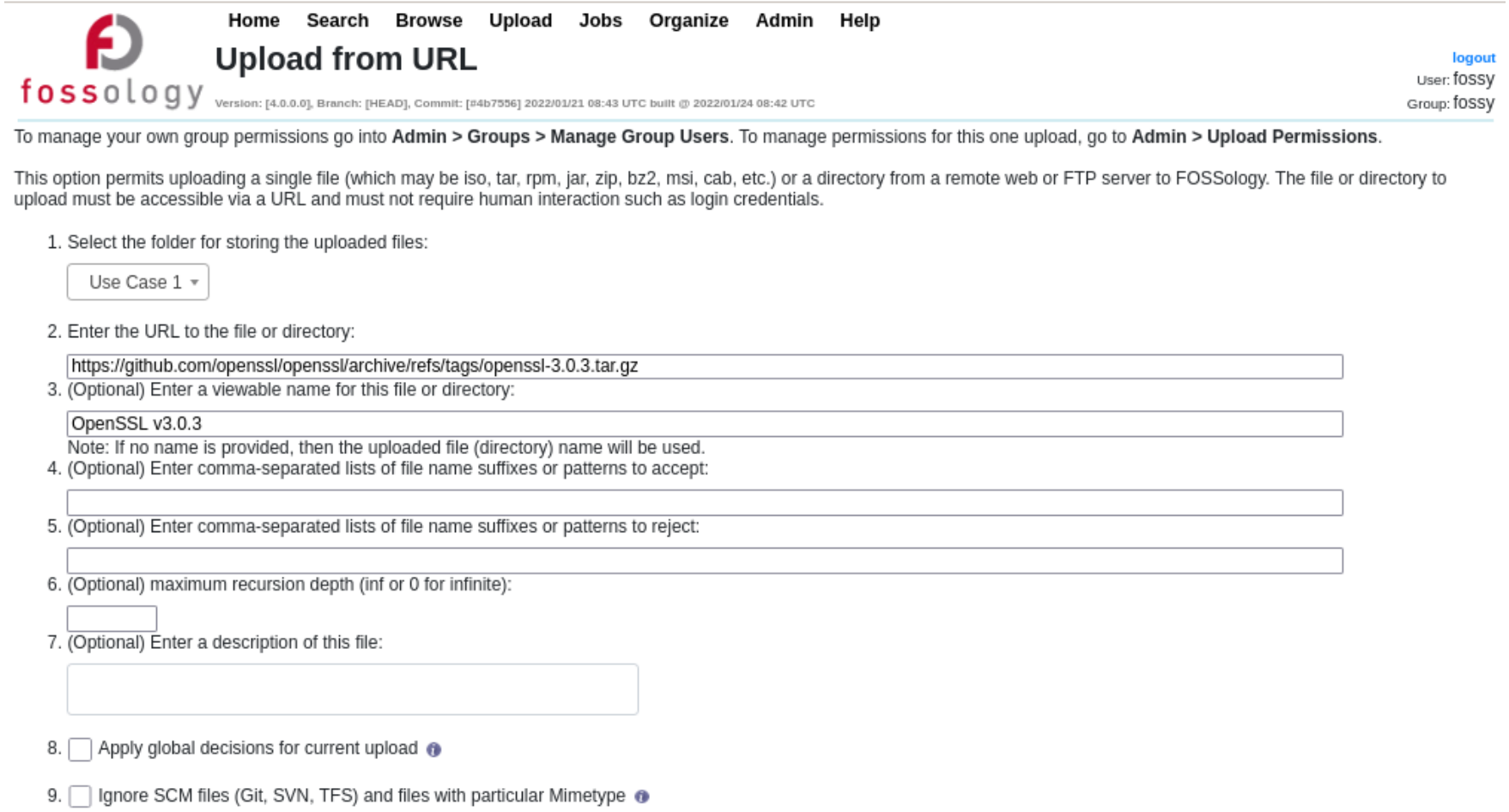
- Example: **OpenSSL v3.0.3**
- Clearing a software package with FOSSology **reusing existing curation data** involves the following steps:
  1. Converting SPDX Tag:Value file into SPDX RDF format.
  2. Uploading curated source packages into FOSSology.
  3. Importing the SPDX RDF file into the corresponding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

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]
```

2. Upload curated packages into FOSSology



cont'd:

10. ○ Visible only for active group
    ◉ Visible for all groups
    ○ Make Public

11. Select optional analysis:
    - [x] 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)
    - [x] Monk License Analysis, scanning for licenses performing a text comparison
    - [x] Nomos License Analysis, scanning for licenses using regular expressions
    - [x] 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 Software Repository (fossy:8)
    - [ ] Enhanced reuse (slower)
    - [ ] Reuse main license/s
    - [ ] Reuse report configuration settings
    - [ ] Reuse deactivated copyrights

    Upload to reuse:

    Select upload to reuse

Upload

Before clearing

fossology

Home Search Browse Upload Jobs Organize Admin Help

License Browser

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

logout
User: fossy
Group: fossy

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

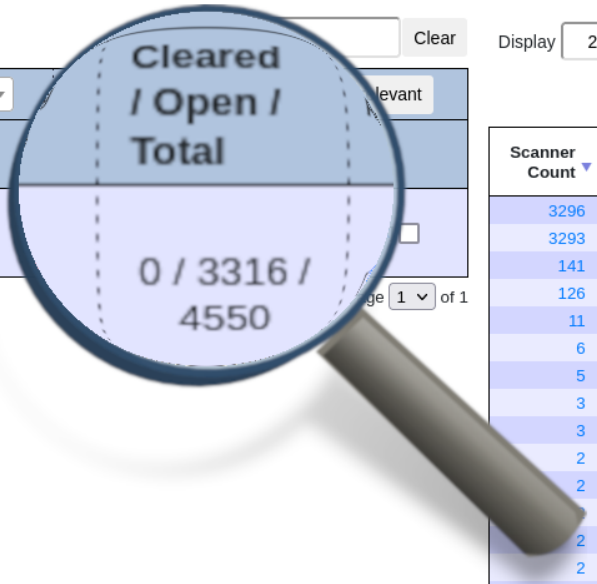
Display 50 files (**tree view** or flat)

Clear

	-- filter for scan results --	-- filter for edited results --	levant
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

ge 1 of 1



- 3316 files contain license information and must be cleared.

Display 25 licenses

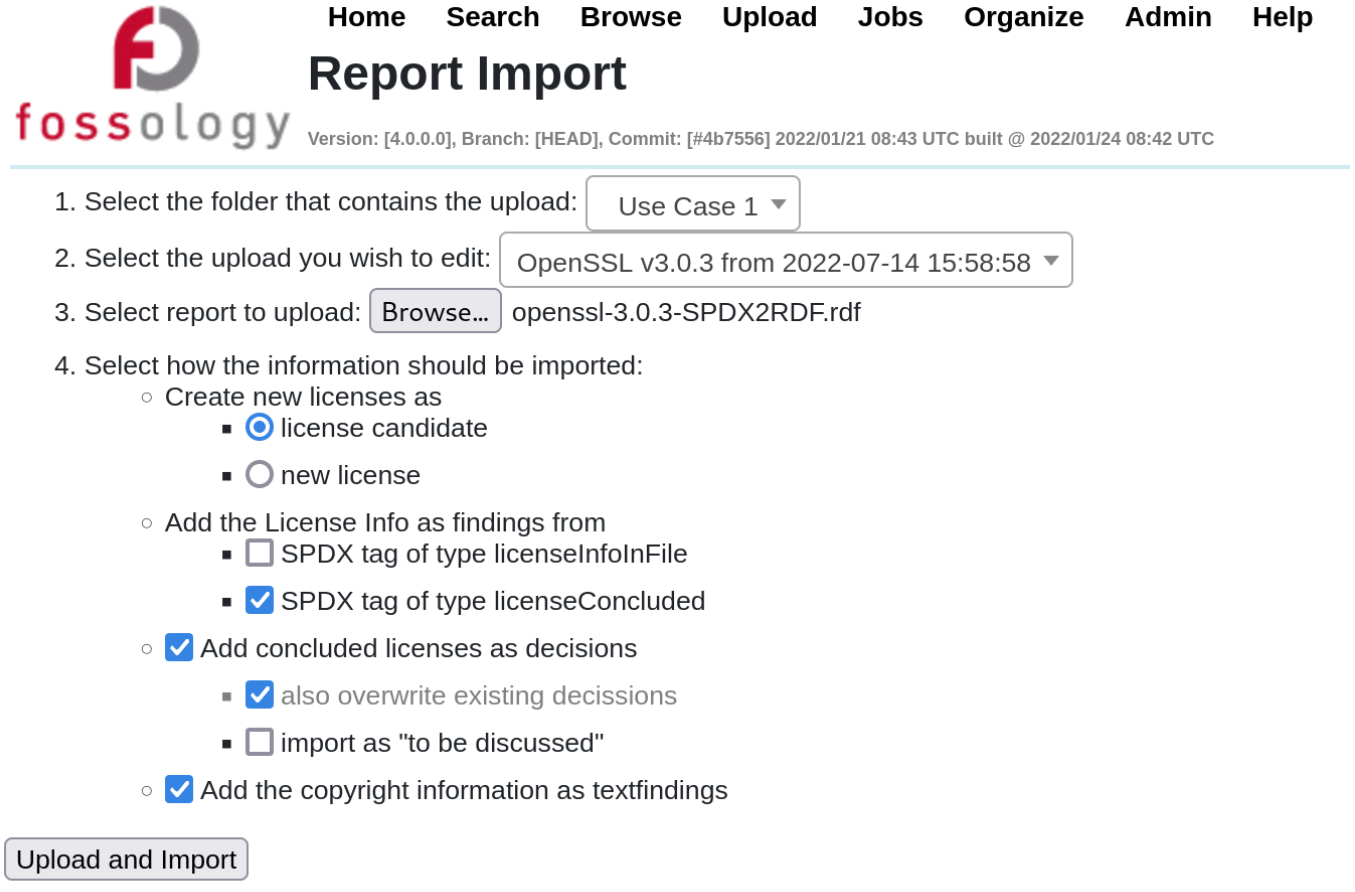
Clear

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
	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 1 of 1

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

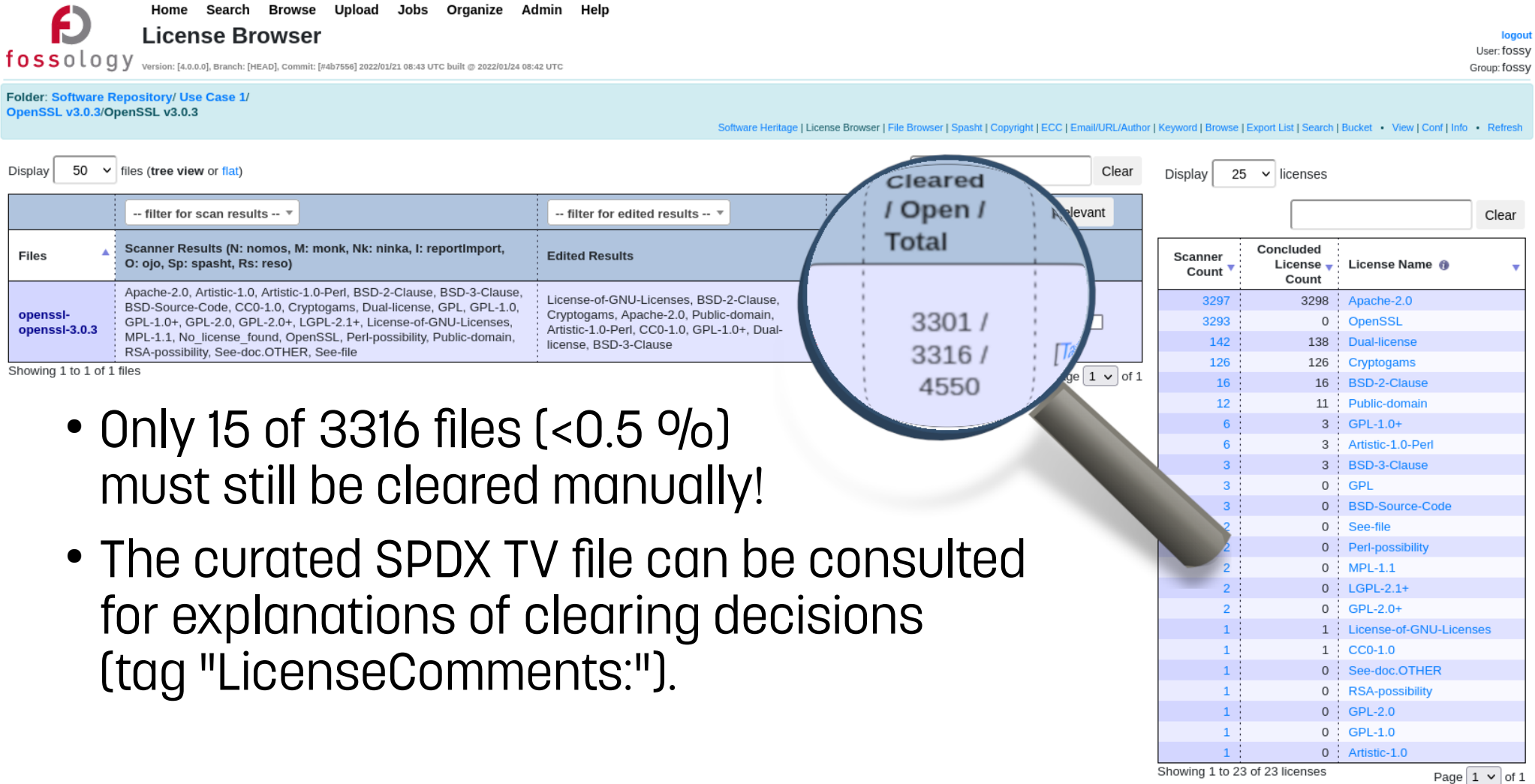


Issues with Import Report

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 looses the tag "LicenseComment" which holds the explanation for certain decisions of "LicenseConcluded".
- Files with "LicenseInfoInFile: [some value]" and "LicenseConluded: NOASSERTION" must be cleared manually.
- Curated copyright notices are only imported as "text findings".

4. Clearing any remaining files



- 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:").