

FOSSology - License Review and Analysis

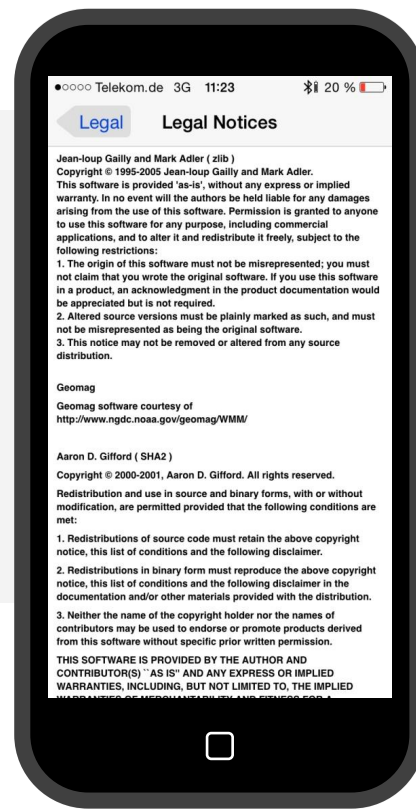
FOSDEM 2018 - 3rd of February, Brussels - The FOSSology Project

The Problem Actually

You know these examples

Distributing open source software requires to

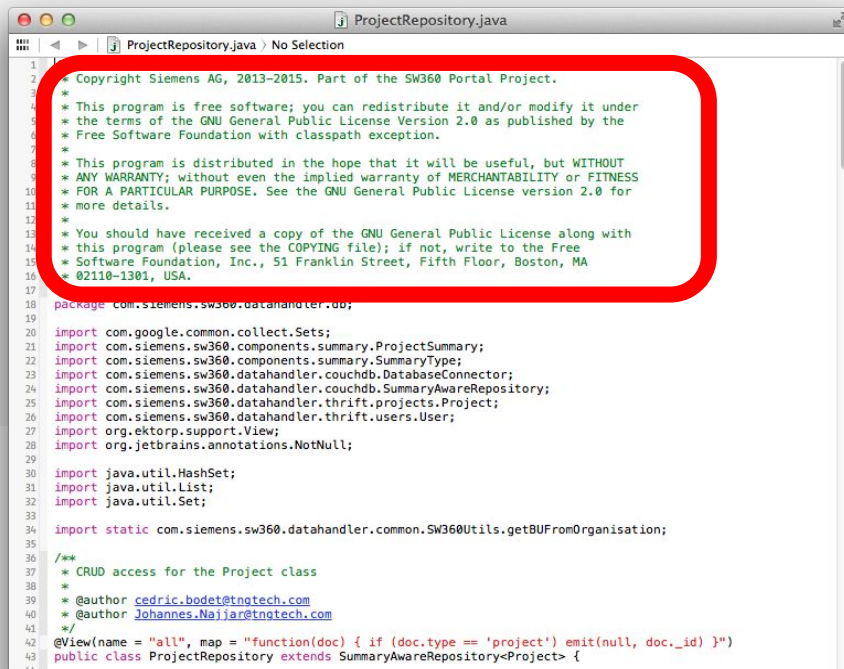
- Provide licenses of involved software
- Provide copyright statements of involved authors
- Provide disclaimers
- ... and much more



It is about finding licenses

Finding Licenses

- License texts
- References to licenses
- Written texts explaining licensing
- License relevant statements



```
1  * Copyright Siemens AG, 2013-2015. Part of the SW360 Portal Project.
2  *
3  * This program is free software; you can redistribute it and/or modify it under
4  * the terms of the GNU General Public License Version 2.0 as published by the
5  * Free Software Foundation with classpath exception.
6  *
7  * This program is distributed in the hope that it will be useful, but WITHOUT
8  * ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS
9  * FOR A PARTICULAR PURPOSE. See the GNU General Public License version 2.0 for
10 * more details.
11 *
12 * You should have received a copy of the GNU General Public License along with
13 * this program (please see the COPYING file); if not, write to the Free
14 * Software Foundation, Inc., 51 Franklin Street, Fifth Floor, Boston, MA
15 * 02110-1301, USA.
16
17 package com.siemens.sw360.datahandler.db;
18
19
20 import com.google.common.collect.Sets;
21 import com.siemens.sw360.components.summary.ProjectSummary;
22 import com.siemens.sw360.components.summary.SummaryType;
23 import com.siemens.sw360.datahandler.couchdb.DatabaseConnector;
24 import com.siemens.sw360.datahandler.couchdb.SummaryAwareRepository;
25 import com.siemens.sw360.datahandler.thrift.projects.Project;
26 import com.siemens.sw360.datahandler.thrift.users.User;
27 import org.ektorp.support.View;
28 import org.jetbrains.annotations.NotNull;
29
30 import java.util.HashSet;
31 import java.util.List;
32 import java.util.Set;
33
34 import static com.siemens.sw360.datahandler.common.SW360Utils.getBUFFromOrganisation;
35
36 /**
37  * CRUD access for the Project class
38  *
39  * @author cedric.bodet@tngtech.com
40  * @author Johannes.Najjar@tngtech.com
41  */
42 @View(name = "all", map = "function(doc) { if (doc.type == 'project') emit(null, doc._id) }")
43 public class ProjectRepository extends SummaryAwareRepository<Project> {
44
45 }
```

Problem of many Licenses (“Proliferation”)

Open Source and Reuse

- It is natural that an OSS project reuses available <https://github.com/fossology/fossology>
- Likely OSS from other projects is found
- For example, FOSSology will find 25 other licensing relevant text occurrences in Apache thrift



Home Search Browse Upload Jobs Organize Admin Help

License Browser

2.1.0-rc1, commit: [40d99302] 2014/12/10 17:53 UTC built @ 2014/12/15 06:49 UTC

Folder: Software Repository/
thrift-0.9.1.tar.gz/
thrift-0.9.1.tar/ thrift-0.9.1

License Browser | Bucket Browser | Copyright/Email/URL | ECC | Patents | Browse | License List | License List Download | Search | View | Info | Refresh

Display 50 files

Scanner Count	Concluded License Count	License Name
2421	0	Apache-2.0
819	0	No_license_found
132	0	FSF
94	0	UnclassifiedLicense
13	0	Freeware
8	0	GPLv2+
6	0	GPL-exception
6	0	autoConfException
4	0	Zlib
4	0	MIT
4	0	LGPL-2.1
3	0	SeeFile
3	0	MIT-style
2	0	Trademark-ref
2	0	GPLv3+
2	0	GPL-3.0+-with-bison-exception
2	0	GPL-2.0-with-autoconf-exception
2	0	GPL-2.0+
2	0	BisonException
2	0	Apache-possibility
1	0	X11
1	0	WebM
1	0	See-file
1	0	See-doc(OTHER)

Files	Scanner Results (N: nomos, M: monk, Nk: ninka)
adocal	Freeware, FSF, GPL-2.0-with-autoconf-exception, GPLv2+,
compiler	Apache-2.0, BisonException, FSF, GPL-3.0+-with-bison-exc
contrib	Apache-2.0, Freeware, No_license_found, See-file, SeeFile,
debian	Apache, LGPL-2.1, MIT, MIT-style, No_license_found, Undas
doc	Apache-2.0, LesserGPLv2.1+, LGPL-2.1, MIT, MIT-style, MI
lib	Apache-2.0, Apache-possibility, BSD-3-Clause, FSF, No_lic
test	See-doc(OTHER), SeeFile, UnclassifiedLicense, WebM
tutorial	Apache-2.0, FSF, No_license_found, UnclassifiedLicense
.travis.yml	Apache-2.0 [Nk: 100%][N]
adocal.m4	FSF [M: 94%][N], autoConfException [Nk: 100%], GPLv2+
CHANGES	UnclassifiedLicense [Nk], Apache-possibility [N]
config.guess	autoConfException [Nk: 100%], GPLv2+ [Nk: 100%], GPL
config.h	No_license_found [Nk][N]
config.hin	No_license_found [Nk][N]

Examples for Licensing – Clarification Needed 3

(TrueCrypt 7.1a Source.zip/ Common/ Cache.c)

Legal Notice: Some portions of the source code contained in this file were derived from the source code of Encryption for the Masses 2.02a, which is Copyright (c) 1998-2000 Paul Le Roux and which is governed by the 'License Agreement for Encryption for the Masses'. Modifications and additions to the original source code (contained in this file) and all other portions of this file are Copyright (c) 2003-2008 TrueCrypt Developers Association and are governed by the TrueCrypt License 3.0 the full text of which is contained in the file License.txt included in TrueCrypt binary and source code distribution packages. */

...

Another real world example:

- The text is actually occurs with this formatting in file
- Very special occurrence in fact that requires review

Examples for Licensing – Clarification Needed 2

*(from zlib-1.2.8.tar/ zlib-1.2.8/ contrib/ amd64/
amd64-match.S)*

```
/*  
 * match.S -- optimized version of longest_match()  
 * based on the similar work by Gilles Vollant,  
 * and Brian Raiter, written 1998  
 *  
 * This is free software; you can redistribute it and/or  
 * modify it  
 * under the terms of the BSD License.  
 * Use by owners of Che Guevarra  
 * parafernalia is prohibited, where possible,  
 * and highly discouraged  
 * elsewhere.  
 */
```

...

Another real world example:

- What was meant to be fun (or a political statement), is difficult for license analysis
- *Question: Can this be ignored or shall the origination check for ownership of referred parafernalia?*

Examples for Licensing – Clarification Needed 4

(TrueCrypt 7.1a Source.zip/ Crypto/ AesSmall.h)

The free distribution and use of this software in both source and binary form is allowed (with or without changes) provided that:

1. distributions of this source code include the above copyright notice, this list of conditions and the following disclaimer;
2. distributions in binary form include the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other associated materials;
3. the copyright holder's name is not used to endorse products built using this software without specific written permission.

Another real world example:

- How does the organization decide which license to choose
- There may be an external reason for choosing either one or the another

ALTERNATIVELY. provided that this notice is retained in full. this product may be distributed under the terms of the GNU General Public License (GPL). in which case the provisions of the GPL apply INSTEAD OF those given above.

DISCLAIMER

This software is provided 'as is' with no explicit or implied warranties in respect of its properties, including, but not limited to, correctness and/or fitness for purpose.

Examples for Licensing – Attention Needed 5

```
/******
```

```
* Copyright (C) 2008 - 2015 ***, Inc. All rights reserved.
```

```
*
```

```
* Permission is hereby granted, free of charge, to any person obtaining a copy  
* of this Software and associated documentation files (the "Software"), to deal  
* in the Software without restriction, including without limitation the rights  
* to use, copy, modify, merge, publish, distribute, sublicense, and/or sell  
* copies of the Software, and to permit persons to whom the Software is  
* furnished to do so, subject to the following conditions:
```

```
*
```

```
* The above copyright notice and this permission notice shall be included in  
* all copies or substantial portions of the Software.
```

```
*
```

```
* Use of the Software is limited solely to applications:
```

```
* (a) running on a *** device. or
```

```
* (b) that interact with a *** device through a bus or interconnect.
```

```
*
```

```
* THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR  
* IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY,  
* FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE  
* *** CONSORTIUM BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY,  
* WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF  
* OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE  
* SOFTWARE.
```

```
*
```

```
*****/
```

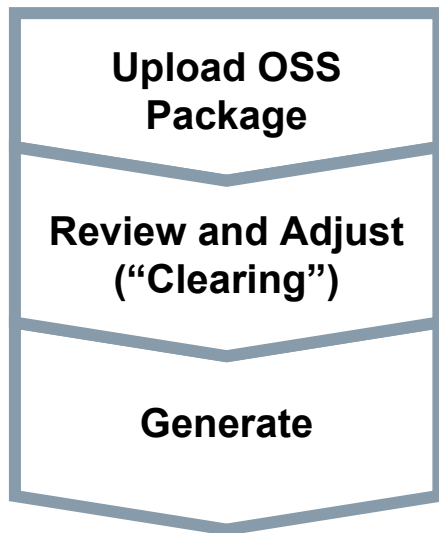
Another real world example:

- It is actually based on an MIT license text
- MIT license: very popular and permissive
- Added two conditions inside the original license text
 - (not so permissive)
- Very hard to identify with regular expression matching

How does FOSSology work?



See more details the Basic Workflow Description: <https://www.fossology.org/get-started/basic-workflow>

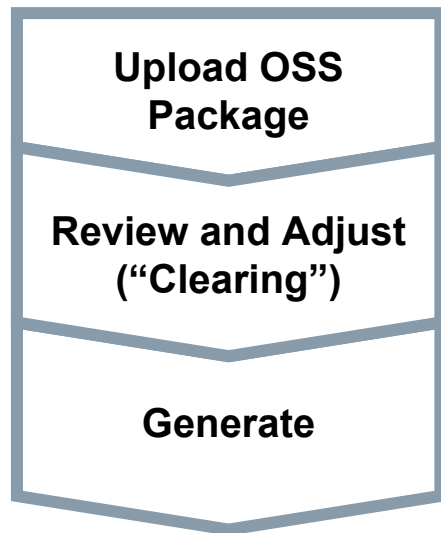


- Upload an open source package to the server
- Select scan agents that analyze the software
- Review what scanners have found
- Review license occurrences and correct findings if necessary
- Generate report output
- For example list of licenses or SPDX

What is the point of FOSSology?



See more details the Basic Workflow Description: <https://www.fossology.org/get-started/basic-workflow>



- Upload an open source package to the server
- ~~Select scan agents that analyze the software~~
- Review what scanners have found
- Review license occurrences and correct findings if necessary
- Generate report output
- For example list of licenses or SPDX

Key facts



- Linux Foundation collaboration project
- GPLv2 licensed
- Linux, and only Linux application
- Mostly C/C++ and PHP
- Frontend runs on Apache httpd
- Provides also a Command Line Interface
- Backend schedules multiple Agents in parallel
- PostgreSQL as database
- Provides scripts for Docker and Vagrant

fossology / **fossology**

<> Code

! Issues 208

🔗 Pull requests 15

📁 Projects 0

📖 Wiki

FOSSology is an open source license compliance software system and toolkit. A export control scans from the command line. As a system, a database and web workflow. License, copyright and export scanners are tools used in the workflow.

fossology

spdx

license-management

license

Manage topics

🕒 7,579 commits

🌿 59 branches

Tag: 3.2.0rc1 ▾

New pull request



mcjaeger Merge pull request #918 from fossology/contrib/word-report ...

📁 debian

feat(report): new word report

📁 examples

chore(docs): updating readme and changelog, co

📁 install

feat(copyright): allow to have multiple copyright d

📁 pbconf

feat(report): new word report

Feature: Analysis Documentation and Sharing



SPDX Export allows for exporting and exchanging analysis results

Use Case

- Exchanging licensing documentation with SPDX
- Can I have documentation of my analysis?
- Can I provide comprehensive reporting what was analysed?
- **How can tell others about the licensing details?**

Solution

Export a SPDX report

- FOSSology generates SPDX output as RDF and tag-value
 - Structured and parseable
 - Many tools already support it
- Contains license and copyright listings
- All Information is equipped with the hashes of the corresponding files

Feature: SPDX Import



SPDX Import allows for applying SPDX license analysis information to uploaded source code packages

Use Case

- Licensing information in SPDX files require also to see the original source code
- If you receive an SPDX file from another (unknown) organisation, review is maybe necessary
- **How can I review SPDX license information?**

Solution

- FOSSology allows for uploading SPDX files
- The imported information can be reviewed in the Web UI
- Select different options for importing
- New and custom licenses are created as license candidates

Thank you very much - ... some links:



- <https://github.com/fossology/fossology>
- <https://www.fossology.org/>

Try it Yourself:

```
$ docker run -p 8081:80 fossology/fossology
```

