

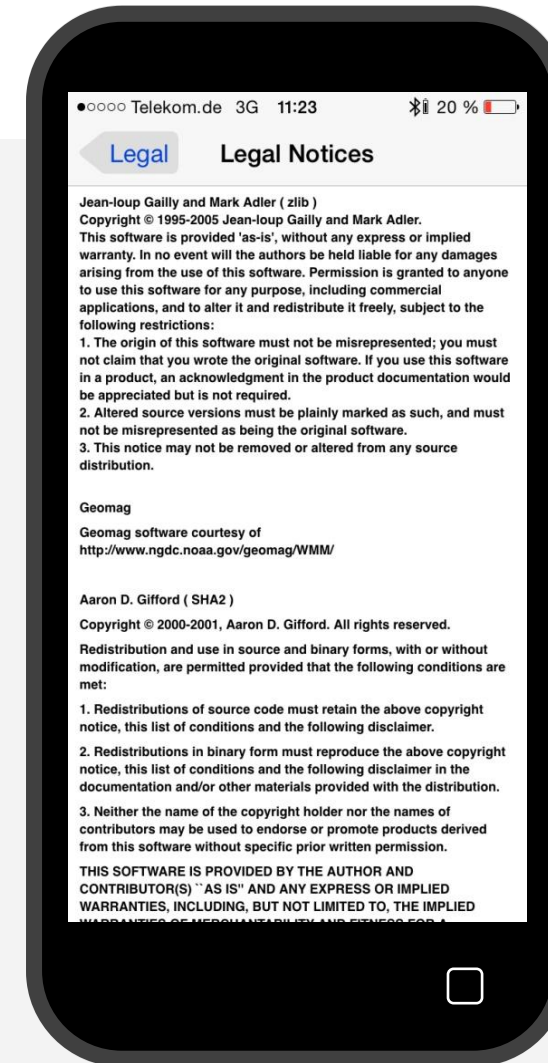
FOSSology: General Introduction

The Problem Actually

You know these examples

Distributing open source software requires to

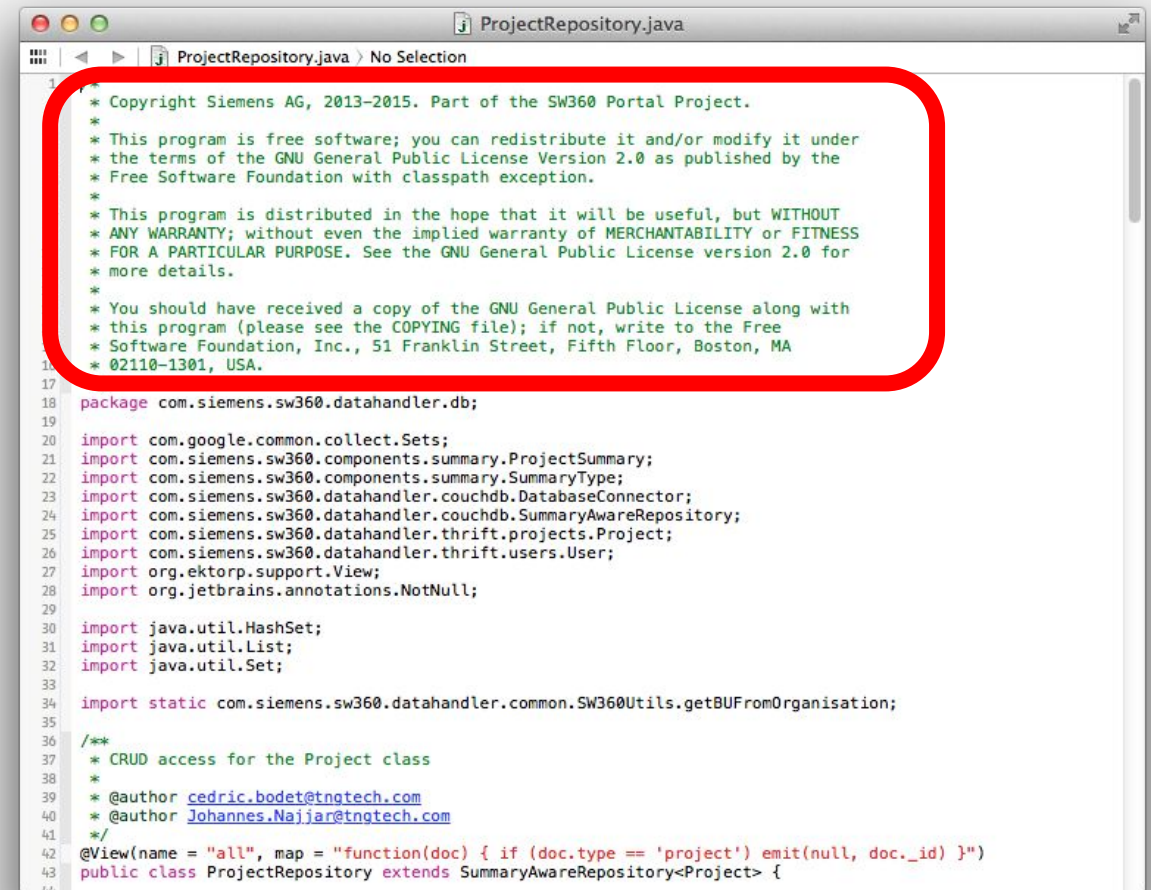
- Provide licenses of involved software
- Provide copyright statements of involved authors
- Provide disclaimers
- ... and much more



It is about finding licenses

Finding Licenses

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



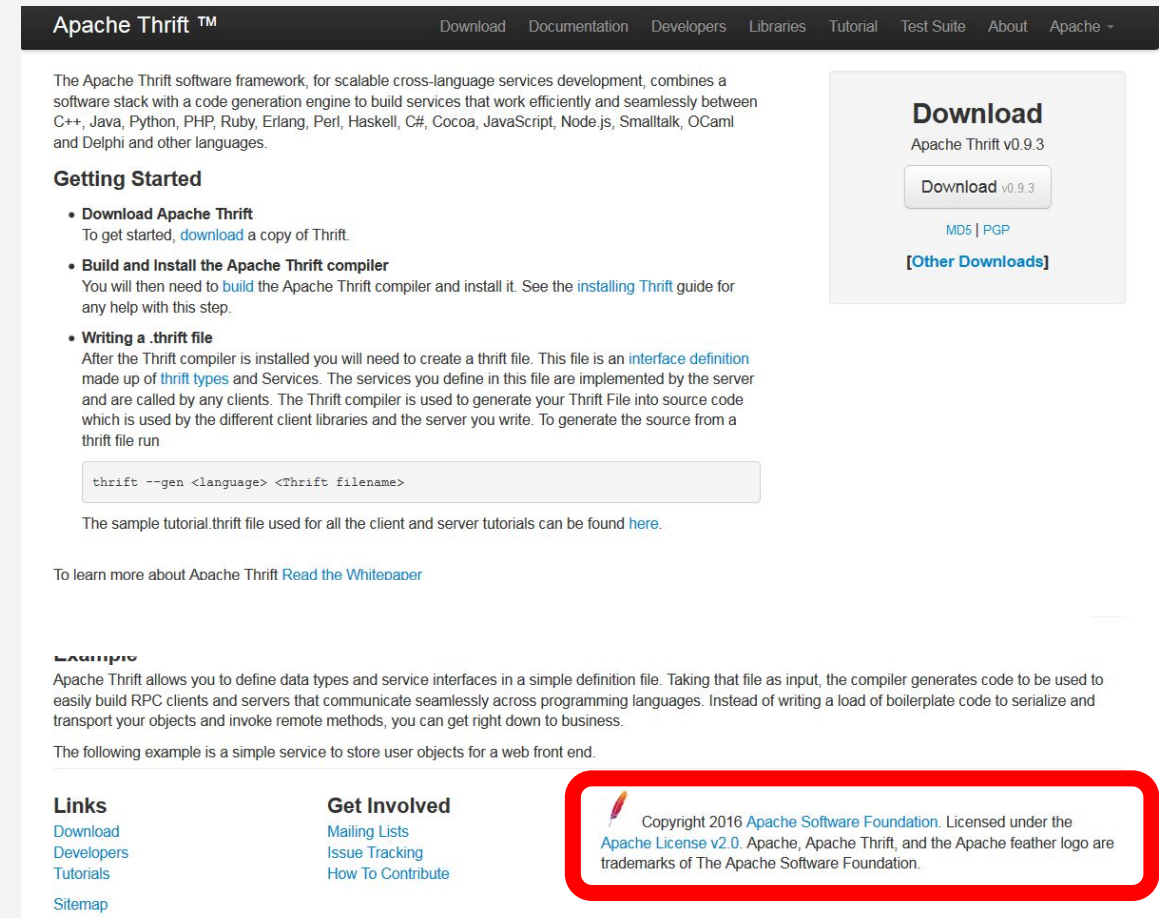
The screenshot shows a code editor window titled 'ProjectRepository.java'. A red rounded rectangle highlights a multi-line comment block at the top of the file, which contains the license text for the SW360 Portal Project. Below the highlighted text, the code continues with package declarations, imports, and class definitions.

```
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.getBUFromOrganisation;
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
```

An Example – What do we find?

The Example

- (Apache projects are known for homogenous licensing)
- Such a project declares its licensing on its Website
- However, Open Source implies using other open source
- As such, a projects can contain also parts from other open source projects



The screenshot shows the Apache Thrift website. The header includes the Apache Thrift logo and navigation links: Download, Documentation, Developers, Libraries, Tutorial, Test Suite, About, and Apache. The main content area has a 'Download' section on the right with a 'Download v0.9.3' button and links for 'MD5 | PGP' and '[Other Downloads]'. The 'Getting Started' section on the left includes instructions for downloading, building, and writing a Thrift file. A code block shows the command 'thrift --gen <language> <Thrift filename>'. The footer contains 'Links' (Download, Developers, Tutorials, Sitemap) and 'Get Involved' (Mailing Lists, Issue Tracking, How To Contribute). A red box highlights the copyright notice: 'Copyright 2016 Apache Software Foundation. Licensed under the Apache License v2.0. Apache, Apache Thrift, and the Apache feather logo are trademarks of The Apache Software Foundation.'

Using FOSSology with this Example

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

logout
User: fossy
Group: fossy

2.1.0-ng, commits: [#0d99362] 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: 25 files

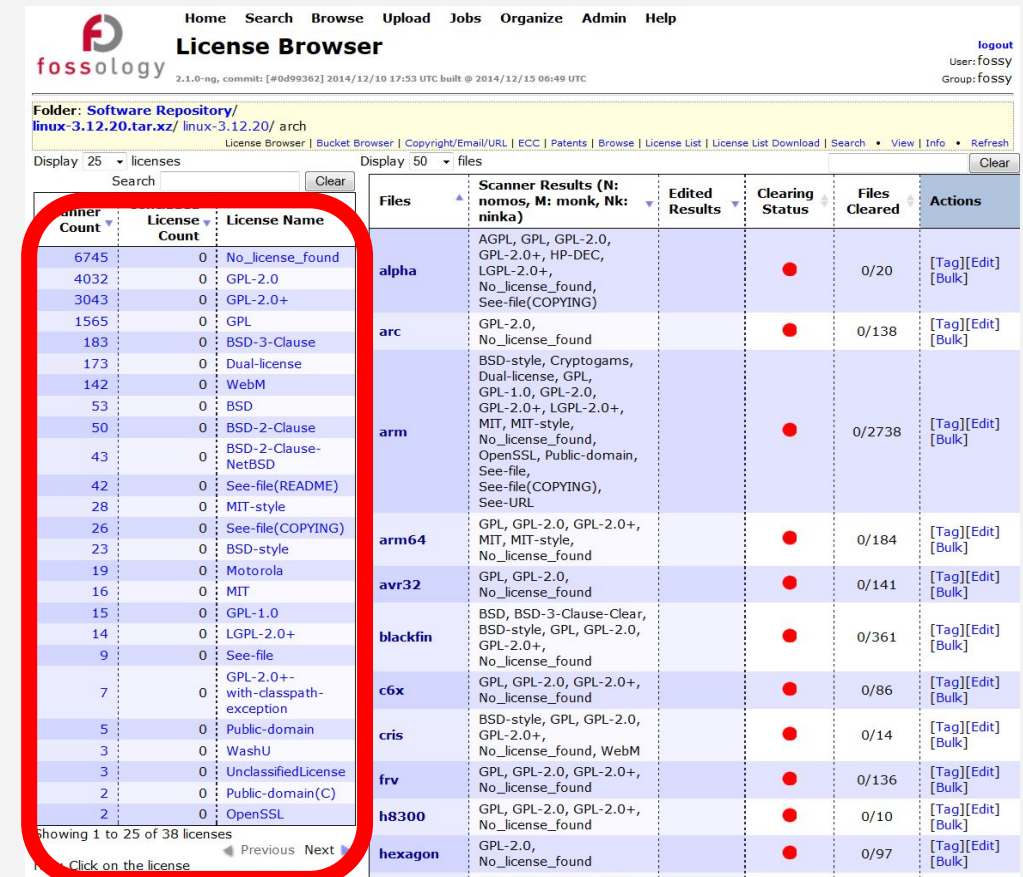
Scanner Count	Concluded License Count	License Name
2421	0	Apache-2.0
819	0	No_license_found
132	0	FSF
94	0	UnclassifiedLicense
13	0	Freeware
8	0	GPLv2+
6	0	GPL-exception
6	0	autoConfException
4	0	Zlib
4	0	MIT
4	0	LGPL-2.1
3	0	SeeFile
3	0	MIT-style
2	0	Trademark-ref
2	0	GPLv3+
2	0	GPL-3.0+-with-bison-exception
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 No_license_found, UnclassifiedLicense, Zlib
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, MIT UnclassifiedLicense
lib	Apache-2.0, Apache-possibility, BSD-3-Clause, FSF, No_licen See-doc(OTHER), SeeFile, UnclassifiedLicense, WebM
test	Apache-2.0, FSF, No_license_found, UnclassifiedLicense
tutorial	Apache-2.0, FSF, No_license_found, Trademark-ref, Unclass
.travis.yml	Apache-2.0 [Nk: 100%][N]
adocal.m4	FSF [M: 94%][N], autoConfException [Nk: 100%], GPLv2+ with-autoconf-exception [N]
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]

Using FOSSology with this Example

Another Example: Linux Kernel Project

- Looking at the Linux kernel will result in thousands of files containing license relevant information
- Analyzing them yields about 30 distinctive licensing statements



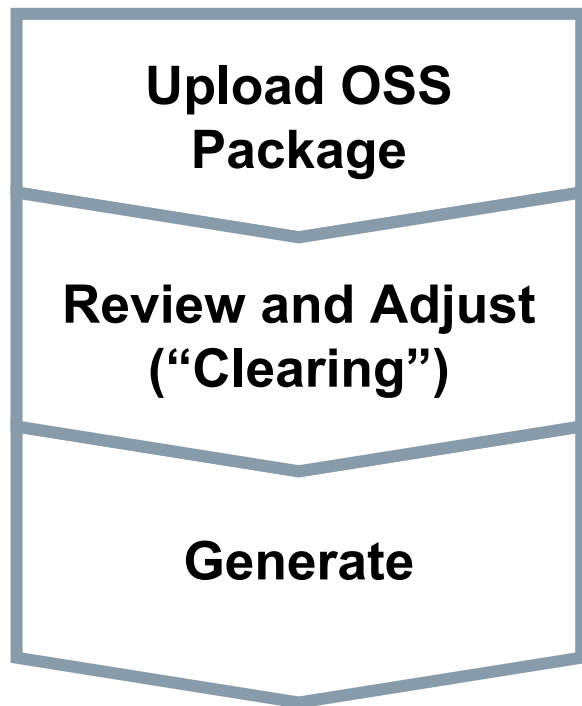
The screenshot shows the FOSSology License Browser interface. The left pane displays a table of licenses found in the project, with a red circle highlighting it. The right pane shows a list of files and their associated license information.

Scanner Count	License Count	License Name
6745	0	No_license_found
4032	0	GPL-2.0
3043	0	GPL-2.0+
1565	0	GPL
183	0	BSD-3-Clause
173	0	Dual-license
142	0	WebM
53	0	BSD
50	0	BSD-2-Clause
43	0	BSD-2-Clause-NetBSD
42	0	See-file(README)
28	0	MIT-style
26	0	See-file(COPYING)
23	0	BSD-style
19	0	Motorola
16	0	MIT
15	0	GPL-1.0
14	0	LGPL-2.0+
9	0	See-file
7	0	GPL-2.0+-with-classpath-exception
5	0	Public-domain
3	0	WashU
3	0	UnclassifiedLicense
2	0	Public-domain(C)
2	0	OpenSSL

Files	Scanner Results (N: nomos, M: monk, Nk: ninka)	Edited Results	Clearing Status	Files Cleared	Actions
alpha	AGPL, GPL, GPL-2.0, GPL-2.0+, HP-DEC, LGPL-2.0+, No_license_found, See-file(COPYING)		●	0/20	[Tag][Edit][Bulk]
arc	GPL-2.0, No_license_found		●	0/138	[Tag][Edit][Bulk]
arm	BSD-style, Cryptogams, Dual-license, GPL, GPL-1.0, GPL-2.0, GPL-2.0+, LGPL-2.0+, MIT, MIT-style, No_license_found, OpenSSL, Public-domain, See-file, See-file(COPYING), See-URL		●	0/2738	[Tag][Edit][Bulk]
arm64	GPL, GPL-2.0, GPL-2.0+, MIT, MIT-style, No_license_found		●	0/184	[Tag][Edit][Bulk]
avr32	GPL, GPL-2.0, No_license_found		●	0/141	[Tag][Edit][Bulk]
blackfin	BSD, BSD-3-Clause-Clear, BSD-style, GPL, GPL-2.0, GPL-2.0+, No_license_found		●	0/361	[Tag][Edit][Bulk]
c6x	GPL, GPL-2.0, GPL-2.0+, No_license_found		●	0/86	[Tag][Edit][Bulk]
cris	BSD-style, GPL, GPL-2.0, GPL-2.0+, No_license_found, WebM		●	0/14	[Tag][Edit][Bulk]
frv	GPL, GPL-2.0, GPL-2.0+, No_license_found		●	0/136	[Tag][Edit][Bulk]
h8300	GPL, GPL-2.0, GPL-2.0+, No_license_found		●	0/10	[Tag][Edit][Bulk]
hexagon	GPL-2.0, No_license_found		●	0/97	[Tag][Edit][Bulk]

How does FOSSology work? – Overview 1 of 2

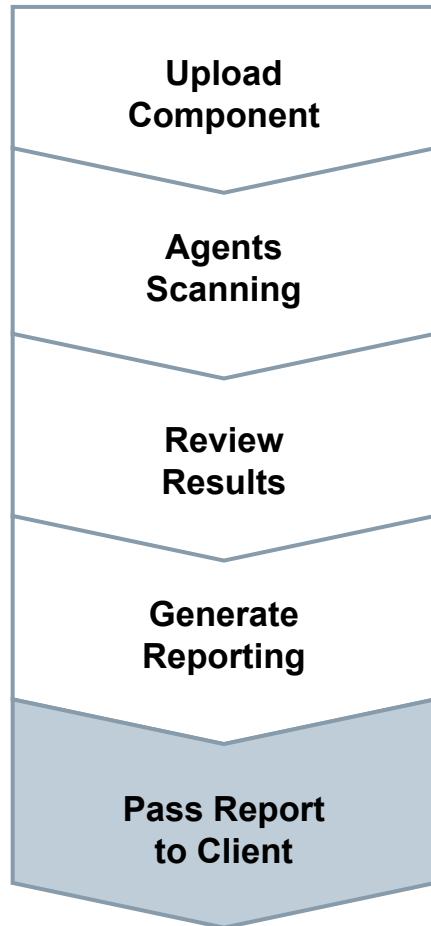
See more details the Basic Workflow Description: <https://www.fossology.org/get-started/basic-workflow>



- 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



How does FOSSology work? – Overview 2 of 2



- Uploading source code archive (*.zip, *.tar.gz, etc)
- Agents scan for license relevant text
- Copyrights, Export Control (ECC), your keywords to look for etc.
- Review scanner results for wrong license classification
- Review other scanner findings (copyrights, ECC)
- Result of the “clearing”
 - SPDX reporting
 - Generated notice or readme file
 - debian-copyright



What is FOSSology?

A Web server application for license and copyright compliance of software components.

FOSSology Project

<https://www.fossology.org/>

- Published first in 2008, GPL-2.0
- 2015: Linux Foundation collaboration project
- Web server based and command line interfaces
- Scanning agents searching for license and copyright relevant hits (and more ...)
- A multi-user / multi-tenant Web UI for review organizing clearing job

FOSSology Development

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

- Standard Web application stack:
 - Linux, Apache 2, PostgreSQL, PHP,
- Web-based UI in PHP, but scanners written in C / C++
- Two ways to interact:
 - Web user interface
 - Command line utilities

FOSSology Feature Overview

A Web server application for license and copyright compliance of software components.

License Scan features

- Regular expression scanner
- Text similarity scanner
- License (text) management
- Aggregation of licenses in hierarchical view
- License histogram
- Supporting concluded vs. found license
- Bulk processing of files with same licensing
- Reusing of license conclusions

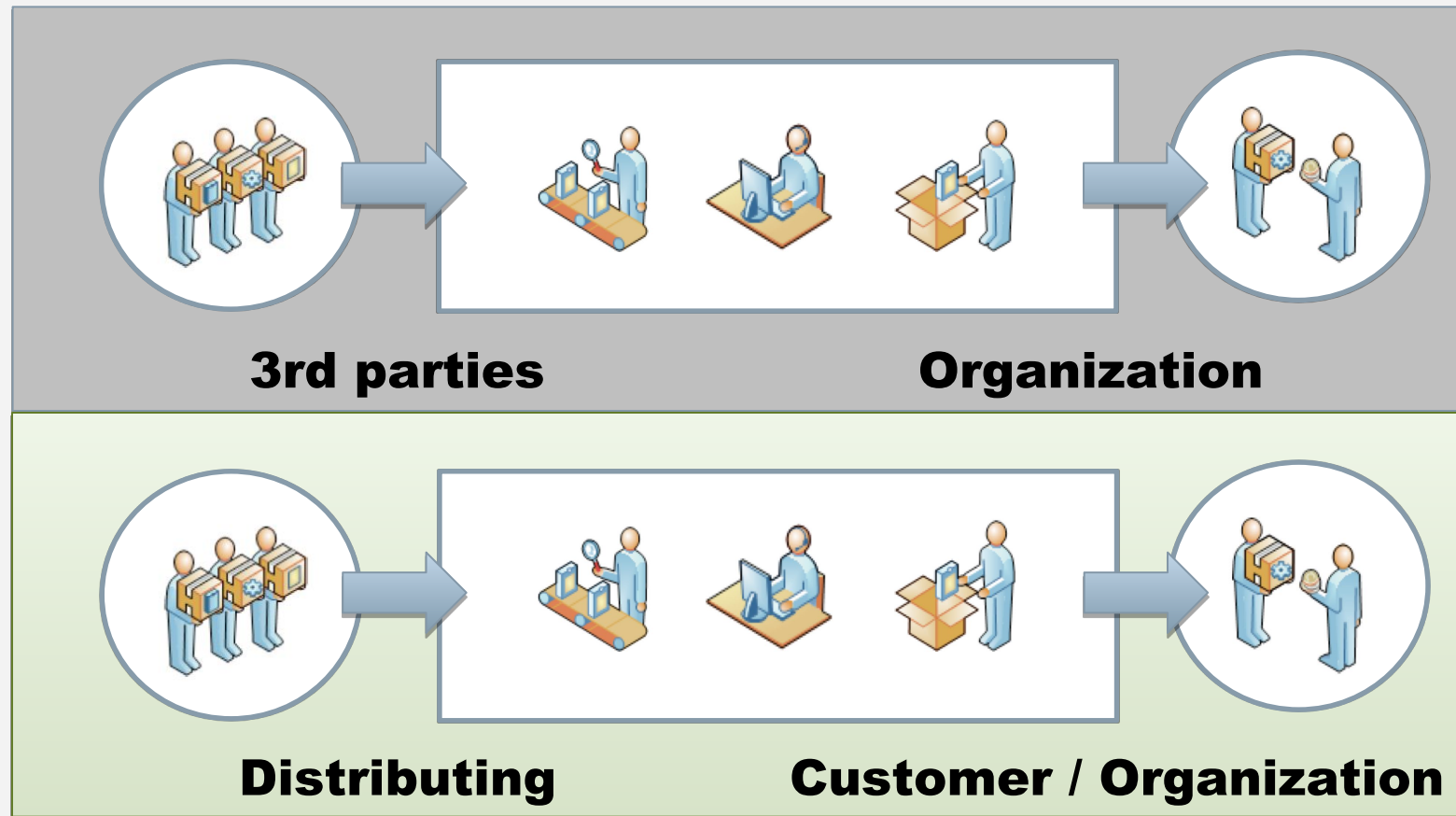
Other features

- Copyright, authorship statements scanner
- Export control and customs scanner
- Command line interfaces
- Reporting
 - SPDX RDF and tag-value
 - Debian-copyright
 - Plain text output
- Files sorting in buckets
- User, group and upload management

FOSSology is Open Source

What happens actually?


- As an organization distributing software there is responsibility for license compliance
- This work must be actually provided by 3rd parties as well!
- *With FOSSology, a tool is freely available supporting all kind of organizations with their license compliance efforts*
 - *GPL-2.0 licensed*



FOSSology is Popular

List of Public Supporters

- See the FOSSology.org website for the most recent list of public supporters.
- *Not on this list?* – Consider supporting the project with putting your logo there. To be included, send email to fossology-steering@fossology.org.




ABOUT FOSSOLOGY

Learn more about FOSSology, the project governance, and license details.



GET STARTED WITH FOSSOLOGY

Instructions on how to get a local instance of FOSSology up and running from source.













IN THE NEWS






The latest announcements, news articles and events related to FOSSology.

FOSSology is an open source license compliance software system and toolkit. As a toolkit you can run license, copyright and export control scans from the command line. As a system, a database and web ui are provided to give you a compliance workflow. License, copyright and export scanners are tools available to help with your compliance activities.

Companies and Organizations Using FOSSology







© 2016 FOSSology Workgroup a Linux Foundation Collaborative Project. All Rights Reserved.
Linux Foundation is a registered trademark of The Linux Foundation. Linux is a registered trademark of Linux Torvalds.
Please see our [privacy policy](#) and [terms of use](#).

FOSSology Conclusions

Major Takeaways

- **It is open source software**
 - No vendor lock-in, it can be shared among partners
- **It is a server-based Web application**
 - Allowing for multi user operations, allowing for reuse
- **It is design for efficiency**
 - Reducing the actual efforts for analyzing OSS components

You like it and would like to contribute?

- Put your logo on the FOSSology homepage if you use it.
Send email to fossology-steering@fossology.org
- Report issues on Github: <https://github.com/fossology/fossology>
- Consider contributing to the code base with your extensions



© Siemens AG 2016, Linux Foundation

CC-BY-SA 4.0

<https://creativecommons.org/licenses/by-sa/4.0/>

Internet

<https://www.fossology.org>

Github

<https://github.com/fossology/fossology>

Further Links

<https://github.com/sw360/sw360portal>

