# The State of Open Source Licensing and Ways to Improve It

Martin Michlmayr Hewlett-Packard

tbm@hp.com



# Agenda

- Why licensing matters
- Open source licensing
  - Contributor agreements
  - Copyright assignment
- Tools for the detection of licenses: FOSSology
- Standard for exchange of license information: Software Package Data Exchange (SPDX)
- Not covered: licenses; legal advice



# Target audience

- Companies using open source, especially those that also distribute it
  - Must understand the importance of honouring licenses
  - Identify licenses and follow them
  - Work with projects to ensure their intentions are followed
- Open source projects
  - Ensure that licensing is done right
  - Work with companies that use and distribute their software
- Researchers
  - Can shed light on best practices
  - Can help improve state of licensing



# Why is licensing an important topic?

- Increasing adoption and penetration of open source
- Companies are getting sued, leading to more awareness:
  - SCO: question of code ownership
  - BusyBox, gpl-violations.org: complying with FOSS licenses



# Problems with FOSS licensing

- Misunderstanding of FOSS licenses: you have obligations
- FOSS licenses and licensing can be complex and complicated
- Keeping track of what FOSS is being used
- Keeping track of FOSS licenses used by an application and how they interact



## Your obligations: copyleft

- GPL: requires source code to be offered to those who receive binaries
- AGPL: additionally requires that the complete source code be made available to any network user



#### Your obligations: permissive

- MIT: The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.
- BSD (3 clause): Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.



#### Who gets sued?

- Whoever distributes the software without compliance
- No excuses: 'software from ODM in Taiwan'
- Indemnification may help
- But reputation is destroyed quickly



#### Contributor Agreements

- Make legal questions around contributions explicit
- Often requires copyright assignment or grants



# Fedora Project Contributor Agreement (FPCA)

- Defines default licenses that are used unless explicit license is given
- Current defaults:
  - Code: MIT
  - Content: Creative Commons Attribution ShareAlike 3.0 Unported
- Does not assign copyright to Fedora or Red Hat



#### Debian

- Every Debian developer has to agree to the DFSG and Social Contract
- DFSG: Debian Free Software Guidelines
- Social Contract: Debian will remain 100% free (according to DFSG)
- debian/copyright: describes upstream copyright/license and that of packaging



#### Linux kernel

#### Developer's Certificate of Origin

- The contribution was created by me and I the have right to submit under indicated open source license
- The contribution is based on previous work that is also under indicated license
- The contribution was provided directly to me by someone who certified it and I didn't modify it
- I understand that the contribution and project are public and recorded

Signed-off-by: Martin Michlmayr <tbm@cyrius.com>



# Copyright assignments

#### Why?

- Preserves the ability to relicense code
- Ensures sufficient rights to enforce licences in court
- Avoids and prevents later competing copyright claims

#### Why not?

- Gives copyright holder a lot of power
- Makes it harder to contribute



# Tools for compliance work

- Binary Analysis Tool
- FOSSology
- Open Source License Checker
- Proprietary tools from Black Duck, Palamida, etc



# **FOSSology**

- FOSSology is a framework to study the source code of FOSS applications in a number of ways
- Main functionality: detection of licenses in open source applications
- Originally developed by HP, it is an open project with an open source license



# **FOSSology**

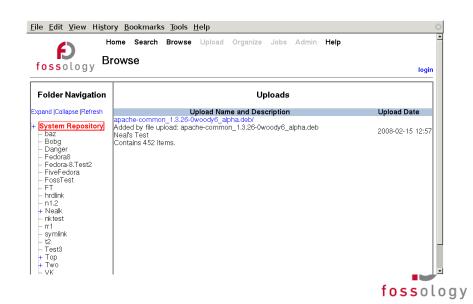
- You load code into the repository
- You analyse it and put the results into a database
- You view the results

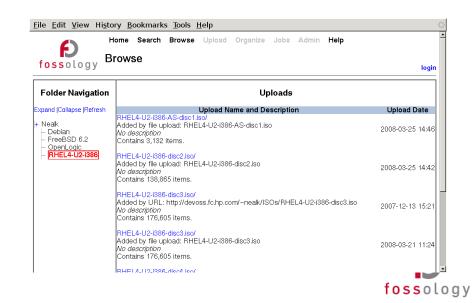




The menu at the top contains all the primary capabilities of FOSSology.

- . Browse: If you don't know where to start, try browsing the currently uploaded projects.
- Search: Look through the uploaded projects for specific files.
- Login: Depending on your account's access rights, you may be able to upload files, schedule analysis tasks, or even add new users.



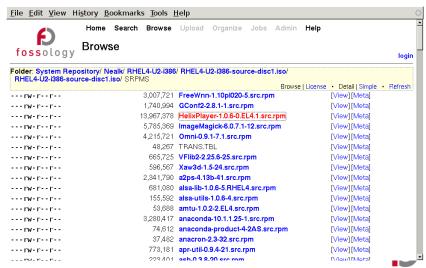




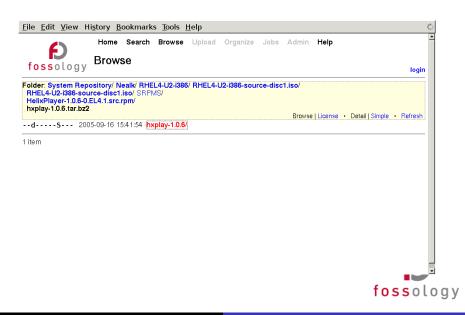
fossology

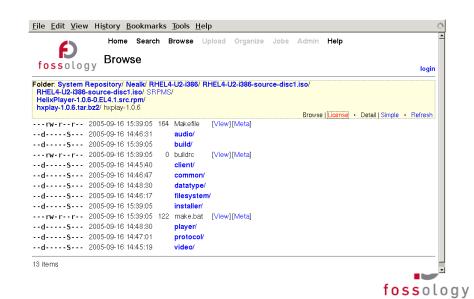




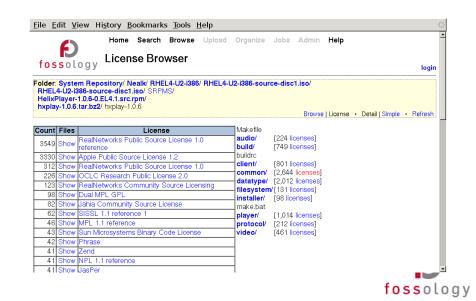






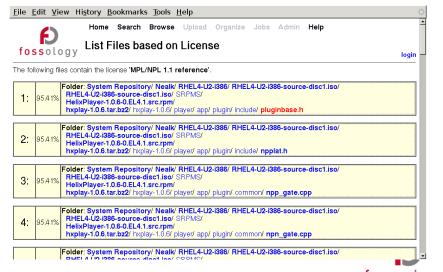


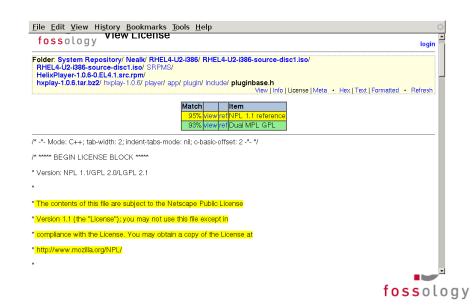
















# FOSSology: the new release

- Buckets
- New license algorithm
- Copyright agent





Home Search Browse Upload Organize Jobs Admin Obsolete Help

#### **Bucket Browser**

Folder: Software Repository/ bobg/ RedHat.tar/ RedHat

Copyright/Email/URL | Nomos License | Browse | Bucket Browser | Nomos License List | Nomos License List Dov

Count	Files	Bucket	amtu-1.0.2-2.EL4.ia64.rpm
52	Show	Approved BAD Licenses (Override all versions)	anaconda-9.1-8.RHELia64.rp anaconda-help-9.1-3.RHEL.nc anaconda-images-9.1-2.RHEL anaconda-product-3-1AS.noa anaconda-runtime-9.1-8.RHE aspell-nl-0.50-3.i386.rpm aspell-nl-0.50-3.src.rpm
4	Show	GPL/LGPL version 3 licenses	
3	Show	Known commercial licenses and/or packages	
3	Show	Package LIKELY to be commercial	
2	Show	SHIP-HOLD licenses	
1	Show	Suspected bad licenses	

Unique buckets: 6

- 16 source packages
- · 49 binary packages
- · 25 binary packages with no source package

anaconda-images-9.1-2.RHEI anaconda-product-3-1AS.noz anaconda-runtime-9.1-8.RHE aspell-nl-0.50-3.i386.rpm aspell-nl-0.50-3.src.rpm caching-nameserver-7.2-10.1 caching-nameserver-7.2-10.2 comps-4WS-0.20080523.src.l comps-4WS-0.20080523.src.l

fossology



Home Search Browse Upload Organize Jobs Admin Obsole

#### Copyright/Email/URL Browser

Folder: Software Repository/ bobg/ RedHat.tar/ RedHat

Bucket Browser | Nomos License | Browse | Copyright/Email/URL | Nomos License List

Jump to: Emails | URLs

Count	Files	Copyright Statements		
142	Show	Copyright (C) 2001 the		
75	Show	Copyright Notice		
69	Show	Copyright Statement		
54	Show	Copyright (C) The Internet Society ( 2003		
41	Show	Copyright (C) 1995 1997 Eric Young (eay@mincom.oz.au) All rights reserved .		
40	Show	(c) )		
40	Show	written by Eric Young (eay@mincom.oz.au)		
26	Show	Copyright Statement .		
26	Show	P. Overell ,		
20	Show	Copyright (c) 2001 Carnegie Mellon University . All rights reserved .		
20	Show	I		
18	Show	Copyright (C) The Internet Society ( 1997		
18	Show	P. Hedstrom		





Home Search Browse Upload Organize Jobs Admin Obsolete Help

#### **Nomos License Browser**

Folder: Software Repository/ bobg/

RedHat.tar/ RedHat

Bucket Browser | Copyright/Email/URL | Browse | Nomos License | Nomos License List | Nomos License Li

Count	Files	License Name	amtu-1.0.2-2.EL4.ia64.rp
4735	Show	No License Found	anaconda-9.1-8.RHEL.ia6
909	Show	(C)Apache	anaconda-help-9.1-3.RH
858	Show	Apache_v2.0	anaconda-images-9.1-2. anaconda-product-3-1AS
812	Show	ATT	anaconda-runtime-9.1-8
787	Show	Misc-Copyright	aspell-nl-0.50-3.i386.rpn
383	Show	(C)U-Washington	aspell-nl-0.50-3.src.rpm
289	Show	GPL_v2+	caching-nameserver-7.2
247	Show	CMU	caching-nameserver-7.2
1		ene	compe AME A SAMORESS



## Scope of the problem

- Prior to distributing a collection of software, each package needs to be reviewed to ensure compliance with all the licenses.
- Supply chain for products now requires software copyright and licensing information for lawsuit avoidance and risk mitigation.
- A package's declared license may not always match the licenses of individual files inside the package itself.
- A package may consist of thousands of files with different licenses in the files
- Need a standard way of referring to the legal compliance 'bill-of-materials' of a software package and be able exchange information with others.

#### Solution: SPDX

- Define a file format for license information to accompany open source packages
  - Focus: Just the facts no interpretations
- Benefits
  - Provides a unified method for exchanging license information
  - Avoids due diligence redundancy where the same source code package is analyzed multiple times by different receivers



#### Structure of standard

- Identification: meta data to associate analysis results with a specific package
- Overview: Facts that are properties for entire package (e.g. package name, declared license)
- File Specific: Facts that are specific to each file included in a package (e.g. filename, copyright)



#### Resources

- Organizations
  - FSF Free Software Licensing and Compliance Lab
  - FSFE Freedom Task Force (FTF)
  - gpl-violations.org
  - Open Source Initiative (OSI)
  - Software Freedom Law Center
- Communities
  - FOSSBazaar
  - FSFE Legal Network
- News and journals
  - Groklaw
  - International Free and Open Source Software Law Review
- Conferences
  - FSFE ELN (European Legal Network)
  - EOLE European Open Source Law Event
- Software

