



# Free, Libre, Open Source Expertise Center (FLOSEC) EMEA



## The FOSSology project



Bruno Cornec

Open Source and Linux  
Technology Consultant





# The FOSSology Project

## Introducing myself



- Software engineering and Unices since 1988
  - Mostly Configuration Management Systems (CMS), Build systems, quality tools, on multiple commercial Unix systems
  - Discover Open Source & Linux (OSL) & first contributions in 1993
  - Full time on OSL since 1995, first as HP reseller then @HP
- Currently:
  - Technology **Architect** on OSL for the HP/Intel Solution Center, Grenoble
  - Leading joint HP+Intel+Red Hat initiative
  - OSL HP **Ambassador**
  - EMEA OSL HP **Profession Lead**
  - Solutions Linux Conference board member
  - MondoRescue, Dploy.org, Project-Builder.org **project lead**
  - LinuxCOE, mrepo, tellico, rinse, fossology **contributor**
  - Mandriva, Fedora **packager**





*FOSS-ology*

The study of FOSS  
(Free Open Source Software)



# The FOSSology Project

## Goal



The goal of the FOSSology project is create tools and a framework to reduce fear, uncertainty, and doubt in the use, development, and distribution of open source software.



# The FOSSology Project

## How?

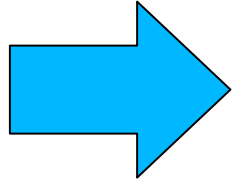


Our goal is to provide tools to solve major FOSS governance problems. Such as:

- License Management
- Code Plagiarism
- Vulnerability tracking



Our goal is to provide tools to solve major FOSS governance problems. Such as:



**License Management**  
Code Plagiarism  
Vulnerability tracking



# The FOSSology Project

## License Discovery



Look at every single file in a package (or distro, or ...)  
Fuzzy match against a library of > 400 known licenses.  
Examine the non-matching portions looking for text  
that could be an unknown license.



# The FOSSology Project

## A Simple FOSSology Process Flow





# The FOSSology Project

Submit software to the FOSSology batch queue





# The FOSSology Project

## File upload screenshot



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

## Upload from URL

User: [fossy](#)  
[logout](#)

This option permits uploading a file from a remote web for FTP server to FOSSology. The file 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 file:

2. Enter the URL to the file:

**NOTE:** If the URL requires authentication or navigation to access, then the upload will fail. Only provide a URL that goes directly to the file. The URL can begin with HTTP://, HTTPS://, or FTP://.

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

4. (Optional) Enter a viewable name for this file:

**NOTE:** If no name is provided, then the uploaded file name will be used.

5. Select optional analysis

- Schedule License Analysis
- Schedule MIME-type Analysis
- Schedule Metadata Analysis
- Schedule Spec File Analysis



# The FOSSology Project

Let's Use FOSSology





## Show Job Queue

User: fossy  
[logout](#)

[Summary](#) | [Detail](#) | [All](#) • [Active](#) | [History](#) • [Refresh](#)

Queued	Scheduled	Running	Finished	Blocked	Failed
--------	-----------	---------	----------	---------	--------

busybox-1.10.2.tar.bz2			
Job/Dependency	Job Name: wget		<a href="#">Reset</a>   <a href="#">Delete</a>   Priority: < 0 >
<a href="#">37584</a>	wget	1 item Elapsed scheduled: 00:00:00 Elapsed running: 00:00:00	2008-05-22 05:15
Job/Dependency	Job Name: unpack		<a href="#">Reset</a>   <a href="#">Delete</a>   Priority: < 0 >
<a href="#">37585 / 37584</a>	unpack	1,328 item Elapsed scheduled: 00:00:15 Elapsed running: 00:00:15	2008-05-22 05:15
Job/Dependency	Job Name: license		<a href="#">Reset</a>   <a href="#">Delete</a>   Priority: < 0 >
<a href="#">37586 / 37585</a>	sqlagent	1 item Elapsed scheduled: 00:00:46 Elapsed running: 00:00:46	2008-05-22 05:16
<a href="#">37587 / 37586</a>	filter_license	0 items Elapsed scheduled: 00:00:00	2008-05-22 05:16



# The FOSSology Project

Let's Use FOSSology





Folder: [Software Repository/ busybox-1.10.2.tar.bz2/](#)  
[busybox-1.10.2.tar.bz2/](#) busybox-1.10.2

[Clear](#) • [Browse](#) | [License](#) • [Detail](#) | [Simple](#) • [Refresh](#)

Count	Files	License
110	<a href="#">Show</a>	Phrase
10	<a href="#">Show</a>	GPLv2 reference 3
10	<a href="#">Show</a>	<b>INRIA-OSL</b>
6	<a href="#">Show</a>	LGPL GNU C Library variant
2	<a href="#">Show</a>	GPL from FSF reference 2
1	<a href="#">Show</a>	GPLv2

```

.indent.pro
AUTHORS
Config.in
INSTALL
LICENSE [2 licenses]
Makefile
Makefile.custom
Makefile.flags
Makefile.help
README
TODO
TODO_config_nommu
applets/
arch/
archival/
console-tools/
coreutils/
debianutils/
docs/ [5 licenses]
e2fsprogs/ [94 licenses]
editors/
examples/
findutils/
include/ [3 licenses]
init/
libbb/ [1 license]
libpwdgrp/ [1 license]
loginutils/

```



## License Browser

User: fussy  
[logout](#)

**Folder:** [Software Repository/ Debian/ etch-4.0r1/ source/ debian-40r1-source-DVD-3.iso/ debian-40r1-source-DVD-3.iso/ pool/ main/ x/ xen-3.0/ xen-3.0\\_3.0.3-0.orig.tar.gz/ xen-3.0-3.0.3-0](#)

[Clear](#) • [Browse](#) | [License](#) • [Detail](#) | [Simple](#) • [Refresh](#)

Count	Files	License
97	<a href="#">Show</a>	LGPL 2.0 reference
74	<a href="#">Show</a>	MIT (modern) with sublicense
66	<a href="#">Show</a>	LGPL GNU C Library variant
53	<a href="#">Show</a>	Phrase
46	<a href="#">Show</a>	LGPL 2.1 reference
39	<a href="#">Show</a>	Intel-OSL
33	<a href="#">Show</a>	Python CWI
32	<a href="#">Show</a>	GPLv3 reference 2
24	<a href="#">Show</a>	GPLv2 reference 6
21	<a href="#">Show</a>	BSD As-Is clause
17	<a href="#">Show</a>	GPL from FSF reference 2
13	<a href="#">Show</a>	GPLv2 reference 5
10	<a href="#">Show</a>	X.Net License
8	<a href="#">Show</a>	GPLv2 reference 3
5	<a href="#">Show</a>	LGPL 2.1
5	<a href="#">Show</a>	GPLv2
4	<a href="#">Show</a>	INRIA-OSL
3	<a href="#">Show</a>	BSD UCREgents 2
2	<a href="#">Show</a>	Python Software Foundation 2.1.1
2	<a href="#">Show</a>	FSF variant 4

[COPYING](#) [2 licenses]  
[ChangeLog](#)  
[Config.mk](#)  
[Makefile](#)  
[README](#)  
**[buildconfigs/](#)**  
**[config/](#)**  
**[docs/](#)**  
**[extras/](#)** [5 licenses]  
[install.sh](#)  
**[linux-2.6-xen-sparse/](#)** [7 licenses]  
**[patches/](#)**  
**[tools/](#)** [443 licenses]  
**[unmodified\\_drivers/](#)**  
**[xen/](#)** [105 licenses]



# The FOSSology Project



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



## License Browser

User: fossy  
[logout](#)

Folder: [Software Repository/ Fedora-8-source-DVD.iso/ Fedora-8-source-DVD.iso](#)

[Clear](#) • [Browse](#) | [License](#) • [Detail](#) | [Simple](#) • [Refresh](#)

Count	Files	License	
51915	<a href="#">Show</a>	Dual MPL GPL	<a href="#">BackupPC-3.0.0-3.fc8.src.rpm</a> [64 licenses]
51666	<a href="#">Show</a>	MPL 1.1 reference	<a href="#">ConsoleKit-0.2.3-1.fc8.src.rpm</a> [81 licenses]
50098	<a href="#">Show</a>	Phrase	<a href="#">GConf2-2.20.1-1.fc8.src.rpm</a> [163 licenses]
49427	<a href="#">Show</a>	LGPL 2.1 reference	<a href="#">GeoIP-1.4.3-1.fc8.src.rpm</a> [47 licenses]
42723	<a href="#">Show</a>	LGPL GNU C Library variant	<a href="#">ImageMagick-6.3.5.9-1.fc8.src.rpm</a> [1,571 licenses]
34578	<a href="#">Show</a>	GPLv2 reference 5	<a href="#">MAKEDEV-3.23-1.2.src.rpm</a> [12 licenses]
25788	<a href="#">Show</a>	GPL from FSF reference 2	<a href="#">MySQL-python-1.2.2-4.fc8.src.rpm</a> [4 licenses]
25670	<a href="#">Show</a>	Eclipse Public License 1.0 reference	<a href="#">NetworkManager-0.7.0-0.5.svn3030.fc8.src.rpm</a> [355 licenses]
25339	<a href="#">Show</a>	eCos	<a href="#">NetworkManager-openvpn-0.7.0-2.svn3047.fc8.src.rpm</a> [141 licenses]
20574	<a href="#">Show</a>	Apache Software License 2.0 reference	<a href="#">NetworkManager-vpnc-0.7.0-0.4.svn3030.fc8.src.rpm</a> [80 licenses]
20345	<a href="#">Show</a>	FSF	<a href="#">ORBit2-2.14.10-2.fc8.src.rpm</a> [135 licenses]
19322	<a href="#">Show</a>	LGPL 2.0 reference	<a href="#">OpenEXR-1.6.0-5.fc8.src.rpm</a> [262 licenses]
14147	<a href="#">Show</a>	GPLv3 reference 2	<a href="#">PolicyKit-0.6-1.fc8.src.rpm</a> [94 licenses]
13234	<a href="#">Show</a>	FSF variant 4	<a href="#">PolicyKit-gnome-0.6-1.fc8.src.rpm</a> [54 licenses]
11299	<a href="#">Show</a>	GPLv2 reference 6	<a href="#">Pound-2.4-0.1.d.fc8.src.rpm</a> [26 licenses]
10973	<a href="#">Show</a>	X.Net License	<a href="#">PyOpenGL-3.0.0-0.4.a6.fc8.src.rpm</a> [5 licenses]
10095	<a href="#">Show</a>	GPLv2 reference 3	<a href="#">PyQt-3.17.1-1.fc7.src.rpm</a> [350 licenses]
5384	<a href="#">Show</a>	MIT (modern) with sublicense	<a href="#">SDL-1.2.12-2.fc8.src.rpm</a> [656 licenses]
3021	<a href="#">Show</a>	INRIA-OSL	TRANS.TBL
2940	<a href="#">Show</a>	FSF variant 2	<a href="#">Xaw3d-1.5E-10.1.src.rpm</a> [377 licenses]
2840	<a href="#">Show</a>	Apache Software License 1.1	<a href="#">a2ps-4.13b-69.fc8.src.rpm</a> [540 licenses]
2730	<a href="#">Show</a>	MIT HP-DEC variant	<a href="#">acl-2.2.39-10.fc8.src.rpm</a> [108 licenses]
2676	<a href="#">Show</a>	MIT (oldstyle) with disclaimer 1	<a href="#">acpid-1.0.6-3.fc8.src.rpm</a> [8 licenses]
2634	<a href="#">Show</a>	BSD As-Is clause	<a href="#">adns-1.2-6.fc8.src.rpm</a> [67 licenses]
2333	<a href="#">Show</a>	X11	<a href="#">agg-2.5-4.fc8.src.rpm</a> [198 licenses]
2137	<a href="#">Show</a>	Open Group	<a href="#">akode-2.0.1-9.fc8.src.rpm</a> [134 licenses]
2021	<a href="#">Show</a>	GPLv2 reference 8	<a href="#">alacarte-0.11.3-4.fc8.src.rpm</a> [106 licenses]
1981	<a href="#">Show</a>	Internet Software Consortium	<a href="#">alchemist-1.0.37-2.fc8.src.rpm</a> [55 licenses]
1950	<a href="#">Show</a>	GPLv2 reference 2	<a href="#">alsa-lib-1.0.15-1.fc8.src.rpm</a> [234 licenses]
			<a href="#">alsa-plugins-1.0.14-5.fc8.src.rpm</a> [71 licenses]
			<a href="#">alsa-utils-1.0.15-1.fc8.src.rpm</a> [110 licenses]



## View License

User: fossy  
[logout](#)

Folder: [Software Repository/ busybox-1.10.2.tar.bz2/](#)  
[busybox-1.10.2.tar.bz2/ busybox-1.10.2/ libbb/ restricted\\_shell.c](#)

[View](#) | [Info](#) | [License](#) | [Meta](#) • [Hex](#) | [Text](#) | [Formatted](#) • [Refresh](#)

Match		Item
91%	<a href="#">view</a>	<a href="#">ref</a> INRIA-OSL

## License: BSD/BSD.old/INRIA-OSL

Copyright (c) 2000-2005 INRIA, France Telecom  
All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. 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.
3. Neither the name of the copyright holders nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING BUT NOT LIMITED TO

```
/* vi: set sw=4 ts=4: */
```

```
/*
```

```
* Copyright 1989 - 1991, Julianne Frances Haugh <joc  
* All rights reserved.
```

```
*
```

```
* Redistribution and use in source and binary forms, with or without  
* modification, are permitted provided that the following conditions  
* are met:
```

- \* 1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- \* 2. 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.
- \* 3. Neither the name of Julianne F. Haugh nor the name of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

```
* THIS SOFTWARE IS PROVIDED BY JULIE HAUGH AND CONTRIBUTORS "AS IS"  
* ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING BUT NOT LIMITED TO  
* IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE  
* ARE DISCLAIMED. IN NO EVENT SHALL JULIE HAUGH OR CONTRIBUTORS BE LIABLE
```



# The FOSSology Project

## License Templates



## fossology License Groups

Folder: [System Repository/ Bobg/ gnome-commander-1.2.5.tar.gz/](#)  
[gnome-commander-1.2.5.tar.gz](#)

[Clear](#) • [Browse](#) | [License](#) | [License Group](#)

Count	Files	License Groups	<a href="#">gnome-commander-1.2.5/</a> [432 licer
363	<a href="#">Show</a>	+ FSF License Groups	
9	<a href="#">Show</a>	...FSF Free Documentation Licenses	
354	<a href="#">Show</a>	...FSF GPL-Compatible Free Software Licenses	
10	<a href="#">Show</a>	+ Fedora License Groups	
13	<a href="#">Show</a>	Phrase	
419	<a href="#">Show</a>	+ Similar Text	

Elapsed time: 0 seconds



# The FOSSology Project

## Disturbing Image





# The FOSSology Project

## Know Thy Licenses



### Skype's GPL Follies

Posted by Serdar Yegulalp, May 9, 2008 12:29 PM

**From:** David Turner <novalis@fsf.org>  
**To:** linux-kernel@vger.kernel.org  
**Subject:** Linksys/Cisco GPL Violations

### xtreme Networks Latest Target Of GPL Enforcement

This is the sixth lawsuit the Software Freedom Law Center has filed against xtreme Networks seeking to uphold open source software principles.

By [Charles Babcock](#)  
**InformationWeek**  
July 22, 2008 01:30 PM

### Verizon Settles Open Source Software Law: Alleged GPL violation spurs accusations, lawsuit

By [Bruce Byfield](#) on September 14, 2006 (8:00:00 AM)

The issue centered on claims that a subcontractor used an open source software package called BusyBox in Verizon's wireless routers.

By [Paul McDougall](#)  
**InformationWeek**  
March 17, 2008 12:50 PM

### Asus resolves Eee GPL violation, releases asus\_acpi code changes

By [Ryan Paul](#) | Published: November 27, 2007 - 05:45PM CT

Defendant is ordered to disclose to Plaintiff how many items of the data storage unit "[...] Wireless G Network media Storage DSM-G600" Defendant has acquired, the number of items of this data storage unit Defendant has sold, the identity of the commercial purchasers of the device, and from whom



# The FOSSology Project

## Know Thy Licenses



### Skype's GPL Follies

Posted by Serdar Yegulalp, May 9, 2008 12:29 PM

**From:** David Turner <novalis@fsf.org>  
**To:** linux-kernel@vger.kernel.org  
**Subject:** Linksys/Cisco GPL Violations

### xtreme Networks Latest Target Of GPL Enforcement

This is the sixth lawsuit the Software Freedom Law Center has filed against xtreme Networks seeking to uphold open source software principles.

By Charles Babcock  
InformationWeek  
July 22, 2008 01:30 PM

### Verizon Settles Open Source Software Law

The issue centered on claims that a subcontractor used an open source software package called BusyBox in Verizon's wireless routers.

By Paul McDougall  
InformationWeek  
March 17, 2008 12:50 PM

### Alleged GPL violation spurs accusations, lawsuit

By Bruce Byfield on September 14, 2006 (8:00:00 AM)

### Asus resolves Eee GPL violation, releases asus\_acpi code changes

By Ryan Paul | Published: November 27, 2007 - 05:45PM CT

Defendant is ordered to disclose to Plaintiff how many items of the data storage unit "[...] Wireless G Network media Storage DSM-G600" Defendant has acquired, the number of items of this data storage unit Defendant has sold, the identity of the commercial purchasers of the device, and from whom





Hosted by FOSSBazaar, under the Linux Foundation.

Oregon State University  
- Open Source Lab (OSUOSL)

Public availability June 2008 (available today with authorization)



Hosted by FOSSBazaar, under the Linux Foundation.

Oregon State University  
- Open Source Lab (OSUOSL)

Public availability June 2008 (available today with authorization)



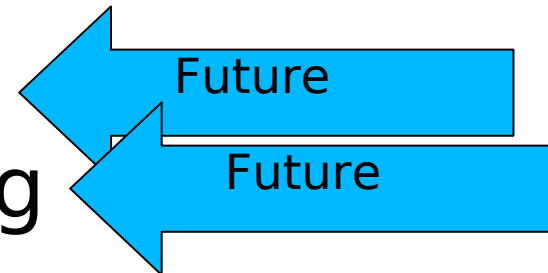


## Major FOSS Governance problems:

**License discovery**

Code plagiarism

Vulnerability tracking





# The FOSSology Project



## View Meta Data

Folder: **Software Repository/ Fedora8-Test2/ d-f/ eclipse/ eclipse/ eclipse/ eclipse.spec**

Item	Meta Data		
1	Unpacked file type	application/x-rpm-spec	
2	Copyright	Eclipse Public License	
3	Epoch	1	
4	Group	Development/Languages	
5	Group	Text Editors/Integrated Development Environm	
6	License	Eclipse Public License	
7	Name	eclipse	73 Requires   lucene-contrib >= 1.9.1
8	Name	eclipse-cvs-client	74 Requires   mx4j >= 2.1
9	Name	eclipse-jdt	75 Requires   regexp
10	Name	eclipse-pde	76 Requires   rpmlib(VersionedDependencies) <= 3.0.3-1
11	Name	eclipse-pde-runtime	77 Requires   tomcat5 >= 5.5.23-9jpp.4
12	Name	eclipse-platform	78 Requires   tomcat5-jasper-eclipse >= 5.5.23-9jpp.4
13	Name	eclipse-rcp	79 Requires   tomcat5-jsp-2.0-api >= 5.5.23-9jpp.4
14	Name	libswt3-gtk2	80 Requires   tomcat5-servlet-2.4-api >= 5.5.23-9jpp.4
15	Obsoletes	ecj	81 Requires   /usr/lib64/eclipse/plugins/org.eclipse.swt.gtk
16	Obsoletes	eclipse-sdk	82 Summary   An open, extensible IDE
17	Provides	ecj	83 Summary   Eclipse Compiler for Java
18	Provides	eclipse	84 Summary   Eclipse CVS Client
19	Provides	eclipse-sdk	85 Summary   Eclipse Java Development Tools
20	Release	19	86 Summary   Eclipse platform common files
21	Requires	ant-antlr	87 Summary   Eclipse Plugin Development Environment
22	Requires	ant-apache-bcel	88 Summary   Eclipse Plugin Development Environment run
23	Requires	ant-apache-log4j	89 Summary   Eclipse Rich Client Platform
24	Requires	ant-apache-oro	90 Summary   SWT Library for GTK+-2.0
25	Requires	ant-apache-regexp	91 URL   <a href="http://www.eclipse.org/">http://www.eclipse.org/</a>
26	Requires	ant-apache-resolver	92 Version   3.3.0
27	Requires	ant-commons-logging	
28	Requires	ant-javamail	

Total meta data records: 92



# The FOSSology Project

Other uses for FOSSology



Copyright geneology

Trademark search

Vulnerability tracking

Dependency graphs

Distro, package, file diffs

Localization reports

?

Your input here

An open source project. Contributions are encouraged.





# The FOSSology Project

What's new in fossology 1.1.0



Email notification of job status on per user basis.

Roll-up of many bug fixes.

Several new license templates added.

Code cleanup for improved efficiency and maintainability.

Tutorial section, with examples, added to fossology.org.

Lots of new tests added to the automated test suite.

RPM packages for RHEL4/5, CentOS4/5 in addition to Debian ones

Many improvements to scheduler to improve robustness (watchdog).





# The FOSSology Project

What's planned in fossology 1.2.0



New heuristic based license analyzer (based on small phrases and phrases relative to other phrases) aka fo\_nomos.

New machine learning license analyzer (based on sentence clustering). Currently we are calling this F1. If there are results from both this analysis and fo\_nomos, the results will be combined for reporting.

Creation of categories in license reports. ( e.g. “good licenses”, “bad licenses”, “commercial licenses”, “files with no license”, ...)

Report files with no license. Distro reports

Add capability for reanalysis without breaking data persistence ie. do new analysis without removing previous analysis results. This can be used, to compare new and old analysis results, and to insure that report url's are persistent.

Display license changes by package version, on a per file basis between versions of any archive (rpm, tar, etc)

Report Copyrights, URL's, and email addresses.





# The FOSSology Project

## Open Source Governance Workshop



Workshop designed to guide through the top issues around management of Open Source in the enterprise.

Targeted at a cross-organizational audience, including auditing, legal, procurement, operational risk management, technology strategy, and line-of-business departments

### Open Source Baseline

Business Drivers

Various open source touch points in your company

Awareness, responsibilities, risks, processes

### Legal Aspects of Open Source Governance

Assessment of Free and Open source software phenomenon

Detailed discussion of Open Source Licenses

Bridging the legal and technical communities

Other considerations: WEB-based services, mergers and acquisitions, other

### Automating Open Source Compliance

Open Source discovery

License detection and analysis

### Open Source Policy Best Practices

Use of open source - when appropriate, when not appropriate for your business

Review of licenses, product distribution considerations

Considerations for employee contribution to open source community

Company relationship with community

### Open Source Governance Processes

Best practices for open source tracking, review and management

Open Source Compliance Lifecycle, workflow

Building Internal Open Source Communities





# The FOSSology Project

## Contact - Thanks



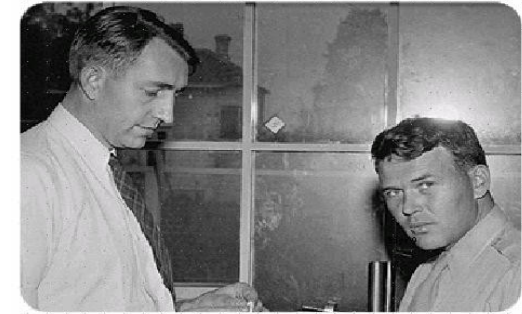
[Bruno.Cornec@hp.com](mailto:Bruno.Cornec@hp.com)

(Open Source and Linux Technology Architect  
at the HP/Intel Solution Center)

<http://www.hp.com/linux>

<http://opensource.hp.com>

<http://fossology.org>



Thanks goes to:

Linus Torvalds, Richard Stallman, Eric Raymond, Nat Makarevitch, René Cougnenc, Eric Dumas, Rémy Card, Bdale Garbee, Bryan Gartner, Craig Lamparter, Lee Mayes, Gallig Renaud, Andree Leidenfrost, Phil Robb, Bob Gobeille, Martin Michlmayr among others, for their work and devotion to the Open Source Software cause... and my family for their patience :-)

Changes are never easy to make. There is comfort and safety in tradition, but change must come, no matter how painful or expensive it may be."

*Bill Hewlett*

