# **Audit Policy for OSTC**

## Table of contents

- 1. Introduction
- 2. Scope
- 3. Glossary
- 4. Audit Process
- 5. Archiving of compliance artifacts
- 6. Consolidation of results of audit work
- 7. Remediation of non-compliant Intellectual Property related findings

## 1 Introduction

In OSTC, we believe in building true open source products. "True open source" does not mean identifying the problem, building the solution and donating that entire solution to the world. Rather, "true open source" is about **collaboration**, it is about **sharing** and **discussing ideas, plans and roadmaps** with others.

"True open source" is about **active and responsible members** of a community who share some **common fundamental values**: the freedom to use, study, share and improve software programs; the freedom of choosing technologies based only on their features and quality, and not because of vendor lock-in strategies; the value of interoperability, as a means to achieve such freedom; the values of shared learning, peer review and meritocracy, as a means to enhance developers' skills – and get better software, too; the value of reusing others' code while respecting their rights, in order to build a true software commons; the value of transparency, to share control on technology and protect everyone's digital sovereignty.

This policy is about three of the most important values that we discussed above; i.e. to be compliant while reusing others' code, protect everyone's digital sovereignty, and make sure that our downstream users get a truly open source product.

## 2 Scope

## Objective Scope

This policy covers **Eclipse Oniro Project** (herein after 'Project') and any other open source projects to which OSTC contributes. As of now, the policy is framed in the context of the Project, as it is the only open source project to which OSTC contributes.

### Subjective Scope

This policy is applicable to the source code audits done by auditors, collaborators and any other individuals or entities connected to source code audit of software incorporated in to the Project.

## 3 Glossary

Some of the definitions in this section are taken from Section 2 of the OpenChain Specification 2.0, Copyright 2016-2019 Linux Foundation, licensed under CC-BY-4.0

**Audit Team** The team that does the source code audit of own code and third party code using Open Source auditing tools.

**Legal Team** The team of lawyers who, drafted the Project's IP policy, reviews the decisions made by the Audit Team as per need, provide right advice, and make critical legal decisions.

**Upstream, downstream** Direction from and to which software and other information or technology flows in a chain of interaction from the originating artifacts and their modifications until they reach the final distribution outlet. If entity A provides technology to entity B, that receives and transforms it, A is upstream to B and B is downstream to A.

**Compliance artifacts** [from Openchain 2.0 defintion] a collection of artifacts that represent the output of the OpenChain-compliant process for the published software. The collection may include (but is not limited to) one or more of the following: source code, attribution notices, copyright notices, copy of licenses, modification notifications, written offers, Open Source component bill of materials, and SPDX documents.

**Open Source License** In descending order of preference (a) an OSI-approved license, or (b) a license qualifying as free software according to the list published by the Free Software Foundation (FSF), or (c) a license falling within the Open Source Definition and/or within the FSF definition of Free Software, **but** which is neither OSI-approved nor expressly listed by FSF (note that this latter option requires a previous assessment by the internal Legal Team to be considered "open source").

**Proprietary blobs** Binary executable code owned by proprietary third party vendors.

**license text reference** Reference to a license. It could be a standard english sentence refering to the license of a file, or SPDX standardized short identifier.

documentation files Files that are part of the documentation of an open source component.

**FOSSology** FOSSology is a 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.

**ScanCode** ScanCode is a tool to scan code and detect licenses, copyrights and more.

**Compliance Artifacts** a collection of artifacts that represent the output of the Program for the Supplied Software. The collection may include (but is not limited to) one or more of the following: source code, attribution notices, copyright notices, copy of licenses, modification notifications, written offers, Open Source component bill of materials, and SPDX documents.

**Identified Licenses** a set of Open Source Software licenses identified as a result of following an appropriate method of identifying Open Source components from which the Supplied Software is comprised.

**SPDX** the format standard created by the Linux Foundation's SPDX (Software Package Data Exchange) WorkingGroup for exchanging license and copyright information for a given software package. A description of the SPDX specification can be found at www.spdx.org.

## 4 The audit process

It is critical for the Project to make sure that the licenses of the source code and binaries that are distributed are those that comply with the Open Source Definition — in brief, they allow software to be freely used, modified, and shared. There could be exceptions though, that needs to be considered according to situations. The audit process has an important role in ensuring this.

The source code audit process in OSTC ensures that information about third party software is available to the downstream users, and all software distributables are accompanied by corresponding SBOMS, that comply with the existing accepted industry standards. In the Project, we reuse the existing tools available in yocto/bitbake, modified to suit our requirements, to collect the source code of all third party components consumed, for license audits. The tool that we developed based on yocto/bitbake is called Aliens4Friends, and the next section briefs about the tool.

The source code of third party components are uploaded to Fossology with the help of Aliens4Friends and automated using gitlab runners. The Fossology agents for unpacking and license identification are run on the source code, triggered by Alines4Friends. Usually, The Audit Team does an initial analysis of the uploaded components, and identify, report and initiate discussions on very obvious license issues, without going in to the in-depth analysis of the source code. The next step is to analyse the source code of each component in detail by navigating through the files and validating the scanner identified licenses. This step is needed because Fossology scanners identify licenses based on keywords and text matches that can introduce false positives as well.

Obligations, restrictions and rights granted by most prevalent open source Licenses (as those listed in 'Choose a License' website) are reviewed and documented by the Audit Team using functionalities provided by our license scanning tool, and are included in each software component's internal report generated through such tool. For uncommon open source licenses (and for uncommon variants of common licenses), an assessment by the Legal Team is required, which will be managed through the OSS issue tracker described in the Appendix (the "OSS IssueTracker"); issues in this respect should be opened by any OSTC team member who encounters an uncommon Identified License that appears not to have been reviewed yet in the OSS Issue Tracker or in the OSS wiki described in the Appendix (the "OSS Wiki"). The final outcome of the assessment will be included in the database of the license scanning tool by the Audit Team.

Please see the 'Getting started with Fossology' document to know how a source code audit is done in Fossology with detailed step by step explanation.

#### Aliens4Friends

It is a tool for Software Composition Analysis (SCA), expressly designed to analyze Yocto/bitbake builds — but it could be usefully adopted in any software composition context where a package manager is missing, and where source code provenance and license/copyright metadata are often missing, messy, uncertain and/or imprecise. The tool uploads source code packages to Fossology, integrate ScanCode as a scanner along with Fossology, find matching Debian packages and import their license and copyright metadata to Fossology, schedule scan agents like ojo for REUSE-compliant packages, collects SBOM in SPDX format, and elaborate the statistics on progress and results of the audit work and show them in a dashboard. More information about the tool could be found in the git repository.

The Aliens4Friends tool applies the license information of packages for which a matching Debian package is found, to files of the corresponding source package in Fossology. This reduces the effort and time required for audit work. The Audit Team may have to audit and apply decision only on files that do not have a license decision applied.

The following section lists different kinds of license texts or files that the Audit Team comes across when they audit source code, and the recommended way to identify each one of them. The types of license texts and files are not exhaustive. There would be updates to this document as and when a new type of license or file is audited.

## Audit of most common types of files found in source code packages

### Source code files

Some of the most common source code files have the extensions .c, .cpp, .java, .py etc. Those files needs to be audited and decisions related to applicability of license texts or references found on such files should be made, as they would be used by downstream users like device makers when they distribute thier applications or products.

Some files that contain source code may be a part of Example folders or Test folders. From the perspective of auditing of Project, such files are treated like other source files, as bundling of such files for distribution could be mandatory in many use cases of downstream users.

The source code files should be audited in the same way as detailed in the Fossology Basic Workflow.

#### **Build files**

From the perspective of source code audits in Project, such files are treated like other source files, as bundling of such files for distribution could be mandatory in many use cases of downstream users.

#### Documentation files

From the perspective of source code audits in Project, such files are treated like other source files, as bundling of such files for distribution could be mandatory in many use cases of downstream users.

## Audit of special non-code files found in source code packages

#### Image files

The Audit Team checks the folders in which the image files are present, or the heirarchical parent folders, for information on the license of the images. e.g. A README file could be present, or information about licenses of the files may be included in the README file of the project. If not found, the team does research in the internet to find out information about licenses or details of copyright holders of the images. The Audit Team, if needed, consults the legal team before making a decision. If no relevant information could be found, that should be documented with details of the research done.

#### Sound files

The Audit Team checks the folders in which the sound files are present, or the heirarchical parent folders, for information on the license of the sound files. e.g. A README file could be present, or information about licenses of the files may be included in the README file of the project. If not found, the team does research in the internet to find out information about licenses or details of copyright holders of the sound files. The Audit Team, if needed, consults the legal team before making a decision. If no relevant information could be found, that should be documented with details of the research done.

#### Font files

From the perspective of source code audits in Project, such files are treated like other source files when license information is present in the font files. If no license information is present, the Audit Team checks the folders in which the font files are present, or the heirarchical parent folders, for information on the license of the font files. e.g. A README file could be present, or information about licenses of the files may be included in the README file of the project. If not found, the team does research in the internet to find out information about licenses or details of copyright holders of the font files. The Audit Team, if needed, consults the legal team before making a decision. If no relevant information could be found, that should be documented with details of the research done.

#### Proprietary blobs

The Audit Team checks the heirarchical parent folders for any license information of the proprietary blobs. If not found, the team does research in the internet to find out information about licenses or details of copyright holders of the propietary blobs. The Audit Team, if needed, consults the legal team before making a decision.

Project projects are generally intended to be released as source only distributions. Distribution of application and/or library binaries should be generally avoided within Project projects; however, some binary blobs could have to be distributed along with Project, to enable compatibility with certain hardware devices and components. License obligations, restrictions and rights related to such binary blobs shall be always reviewed and assessed by the Legal Team. Since the Project is intended to be implemented downstream by device makers, who will typically perform a binary distribution of modified parts of such software, Project working group commits to provide them with some basic assessment and information on license obligations, restrictions in the context of a typical binary/firmware distribution, including information about the existence and the license conditions of possible third party binary blobs.

**Note**: Fossology scanners identify licenses by text matches and keyword matches. Some of the image files, sound files, font files, proprietary blobs etc. may not contain license texts or keywords. Fossology scanners won't consider such files as 'to be audited'. So, the Audit Team may not come across such files during audit. This may cause such files to go without a license decision. To avoid this, and to bring such files for auditing, we have included an automation in our compliance pipeline, to stamp a custom license identifier to such files. During the audit, when the Audit Team finds files with those identifiers, they should follow the respective audit process for the file type, as mentioned in this section.

#### Audit of files with reference to non open source licenses

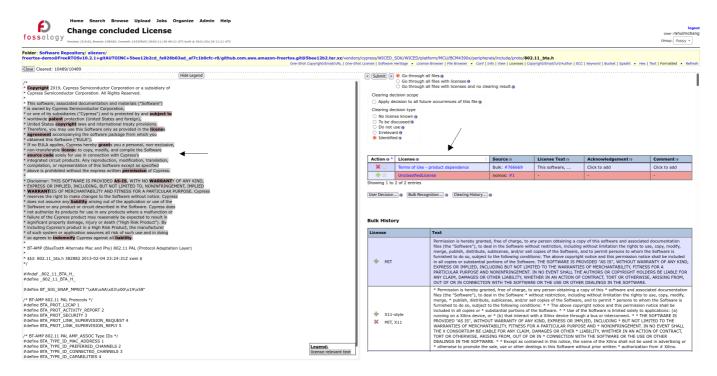
One of the main focus areas for audit of third party code in Project is to identify, document, and decide on the distribution of source code files with licenses that do not provide all the freedoms that are supposed to be guaranteed by an open source license. The importance of identifying such licenses and deciding on their usage is critical for a project that is fully open source.

There are different kinds of such license references. e.g.

### Product dependent licenses

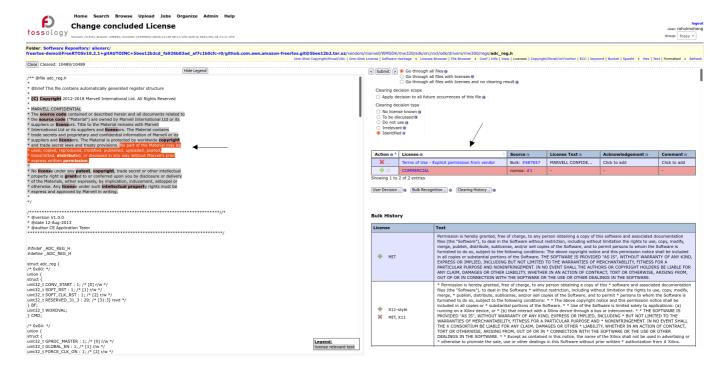
Licenses that grant rights for usage and further distribution based on usage on a particular type of product do not satisfy the criterion 8 of the Open Source Definition. The Audit team identifies such licenses found in the audited source code and documents it for further discussions with the development team. A custom

license identifier called 'Terms of Use - product dependence' is used to identify such licenses. Please see below, a screenshot of such an identification in Fossology:



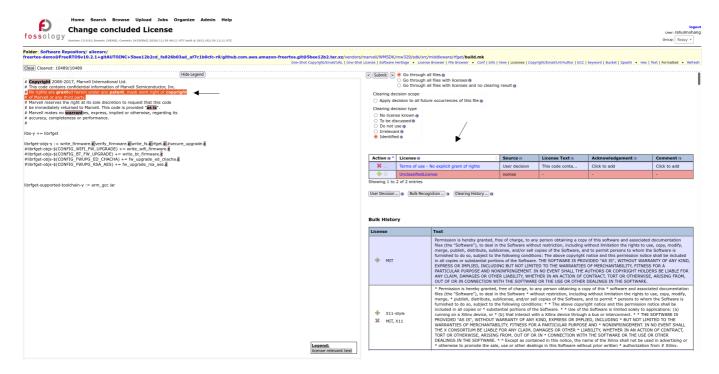
#### Licenses that require explicit permission from vendor

Licenses that restrict usage and distribution without a written consent by the owner of the copyright to the software do not satisfy the criterion 7 of Open Source Definition. The Audit team identifies such licenses found in the audited source code and documents it for further discussions with the development team. A custom license identifier called 'Terms of Use - Explicit permission from vendor' is used to identify such licenses. Please see below, a screenshot of such an identification in Fossology:



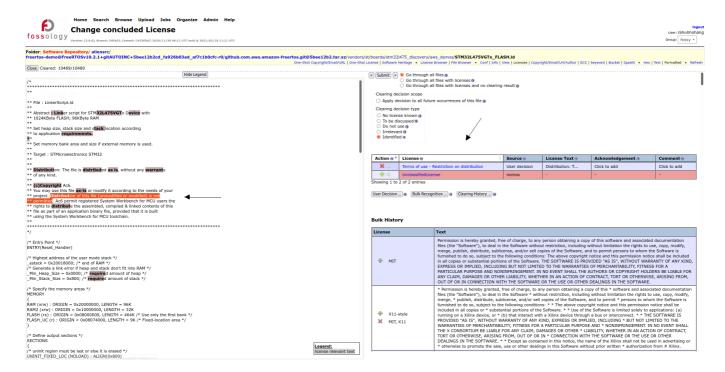
### Licenses that do not have explicit grant of rights of use

Licenses that do not have an explicit statement that grants the rights to use and distribution may restrict the Project as well as the downstream users from making use of the software licensed under such licenses, in any way. The Audit team identifies such licenses found in the audited source code and documents it for further discussions with the development team. A custom license identifier called 'Terms of Use - No explicit grant of rights' is used to identify such licenses. Please see below, a screenshot of such an identification in Fossology:



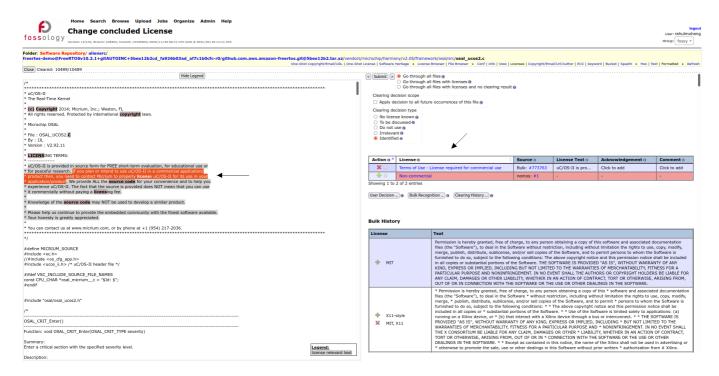
## Licenses that explicitly restricts distribution of software

Licenses that explicitly restricts distribution of software in source code form or binary form do not satisfy the criterion 2 of the Open Source Definition. The Audit team identifies such licenses found in the audited source code and documents it for further discussions with the development team. A custom license identifier called 'Terms of Use - Explicit permission from vendor' is used to identify such licenses. Please see below, a screenshot of such an identification in Fossology:



Licenses that demand additional permission to be taken for commercial distribution

Licenses that, by written statement, demands additional license to be obtained for commercial use, that do not satisfy the criterion 1 of the Open Source Definition. The Audit team identifies such licenses found in the audited source code and documents it for further discussions with the development team. A custom license identifier called 'Terms of Use - License required for commercial use' is used to identify such licenses. Please see below, a screenshot of such an identification in Fossology:



## 5 Archiving of Compliance Artifacts

It is important to manually audit all components used in Project to check the licenses and identify the license decisions in Fossology. It is important that the results of the audit should be stored in a form that is tool independent, readable, and understandable to all stake holders involved. In Fossology, 'bill of materials' could be generated for each source code package uploaded and audited. The bill of materials may be generated in different formats like SPDX RDF, SPDX tag:value, and doc. In the Project, the bill of materials are generated in SPDX RDF format, for each source code package, and stored in our repositories.

## 6 Consolidation of results of audit work

The results from audit of various source packages are consolidated in an excel sheet, that acts as a repository. The respository could be used to retrieve details of audit of a individual source packages using filters. This helps in retrieving audit information without going through individual license clearing reports. The excel sheet would be stored in internal repositories of the Project.

## 7 Remediation of non-compliant Intellectual Property related findings

If there are any non-compliant intellectual property related findings in the audit of the source code, or if license non-compliances are found out by any internal stakeholders, including developers/testers/documentation, issues should be created for each finding in the private gitlab repository, Private SBOM issues, that is maintained by the IP-compliance team.

The IP-compliance team discuss the non-compliant IP finding within the team, and if needed, involve the respective developement team who works with the components in question. If the non-compliant situation

could be rectified by an action by the upstream open source project, the audit team may raise the issues upstream.

If that is time consuming or not feasible, the team may look for options to remove the file where the non-compliant finding was found, or upgrade the component with a newer version where the issue does not exist, or replace the component with another component.

The status of the issue may be updated at Private SBOM issues, and when a decision is made on the resolution and subsequent action is taken, the issue should be closed.