Hermine: Open Source compliance with an automated, community-driven tool
A community-driven approach for (mostly FOSS) licences understanding

- Building a **common way of complying with Open Source Licences**. We want every (non FOSS Expert) lawyer to benefit from open source best practices/knowledge and adapt these practices to their specific needs.

- **Our challenge**: to design a digital common (both a software, a database and a community of practice) that allows any lawyer to comply with any Open Source Obligation.

- In line with the dynamics of OS licences from the beginning (general licences, reuse of clauses, use of exceptions, etc.).
Our methodology:

1) Building a **consensus**:
   - which would bring a community-driven shared understanding of licences **WHILE** still allowing flexibility for private reinterpretation within each organisation
   - Concretely, we worked on the methodology of analysis (defining a common "canvas") and the result of the application of this methodology by a community of (non FOSS experts) lawyers (the reference consensus).

2) with a **automated application** in mind:
   - at the level of licence analysis and at the level of application,
   - Applying the FOSS policies (integrating compliance in CI/CD processes).
Vision & history

Our mindset:

- FOSS and Open Data
- End-user-driven project (Lectra, Orange, OVH Cloud, Enedis, RTE, inno³)
- Don't reinvent the wheel / on the shoulders of giants
- A scope as small as possible / as large as necessary
Where we are on the map

Tool orchestrator

- Package Crawler
- License, Copyright & Authors Scanner
- Snippet Scanner (forensics)
- Package Source Archive
- Dependency Analyzer (Source)
- Package Data Repository
  - COTS Management
- Compliance Artefact Generator
- Dependency Analyzer (Binary)
- Case Data (Situation, Inputs, Status)
- Approval Flow (WFE)
- Dependency Analyzer (Container)
- Policies & Rules
  - Legal Solver (determine obligations)
- License Repository (license facts, rights obligations)

- Audit Log
- Reporting and Analytics
- User & Role Management

Derived from the Capability Map v1.6.0 by the OpenChain Automation Work Group under CC-BY-SA-4.0
Standards we rely on

- For licence IDs: SPDX + NexB's/DejaCode LicenseRef-s
- SBOMs: ORT's evaluated model (scopes !)+ SPDX (+ CyclonDX ASAWHT)
- For Components IDs: PURL
- For curations: ORT's format (still WIP)
- Licence analysis data: Goal is to converge AMAP with OSADL (not that simple) (and others FOSSlight, etc.)
Licence analysis

- Main characteristics:
  - Copyleft, Patent grant, Choice of law
  - Foss policy status

- Obligations
  - Passive / Active
  - Trigger:
    - Modification status
    - Exploitation (distribution source / non-source)
  - Related Generic obligation
Licence analysis: policy status

- Always/ Never / Depends on the context:
- « Authorized contexts »:
  - Technical criteria: exploitation, modification, linking type
  - « Business » criteria: the category of your product
Generic obligations

• Grouping obligations from licences considered functionnally equivalent and that could be handled by the same process.
  - « The name of the Copyright Holder may not be used to endorse or promote products derived from this software without specific prior written permission. » (Artistic-1.0)
  - « Neither the name of the copyright holder nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission. » (BSD-3-Clause)
Using reference data

[Diagram showing data flow from Hermine to Gitlab CI to shared.json]

https://gitlab.com/hermine-project/hermine-data

{JSON} Generic obligations

{JSON} Licenses

shared.json

{JSON}
Using reference data

### Shared reference data

<table>
<thead>
<tr>
<th>Licenses</th>
<th>Total</th>
<th>462</th>
</tr>
</thead>
<tbody>
<tr>
<td>Different from shared data</td>
<td>219</td>
<td></td>
</tr>
<tr>
<td>Absent from shared data</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

7 licenses present in the shared database are missing from local data

#### Generic obligations

<table>
<thead>
<tr>
<th>Total</th>
<th>29</th>
</tr>
</thead>
<tbody>
<tr>
<td>Different from shared data</td>
<td>5</td>
</tr>
</tbody>
</table>

2 generics present in the shared database are missing from local data

#### Apache License 2.0

<table>
<thead>
<tr>
<th>Field</th>
<th>Reference</th>
<th>Local value</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patent grant</td>
<td>True</td>
<td>False</td>
<td>Update local field</td>
</tr>
</tbody>
</table>

### License analysis

<table>
<thead>
<tr>
<th>Name</th>
<th>Field</th>
<th>Reference</th>
<th>Local value</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>License and notices in</td>
<td>Generic</td>
<td>Preserve IP mentions in source code</td>
<td>Preserve legal mentions in source code</td>
<td>Update local field</td>
</tr>
<tr>
<td>Source code</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patent peace</td>
<td>Only in local database</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Contributing to reference data (WIP)
Thanks!

Questions?

Front page photo Statyn "Poseidon med brunnskar" på Götaplatsen i Göteborg. Utförd av Carl Milles.
(c) https://commons.wikimedia.org/wiki/User:Historiker under CC-BY-SA-3.0
Ingesting, validating data

- Importing raw SBOMS
- Adding information
- Validating data (curations / choices)