

SCANOSS / OSSKB 소개

OpenChain KWG Tooling SG

2022. 02. 15.

LG전자 박원재 선임연구원

SW센터 SW공학연구소 Open Source Task

SCANOSS

- Spain에 본사를 둔 기업으로, 'SCANOSS' 플랫폼을 제공

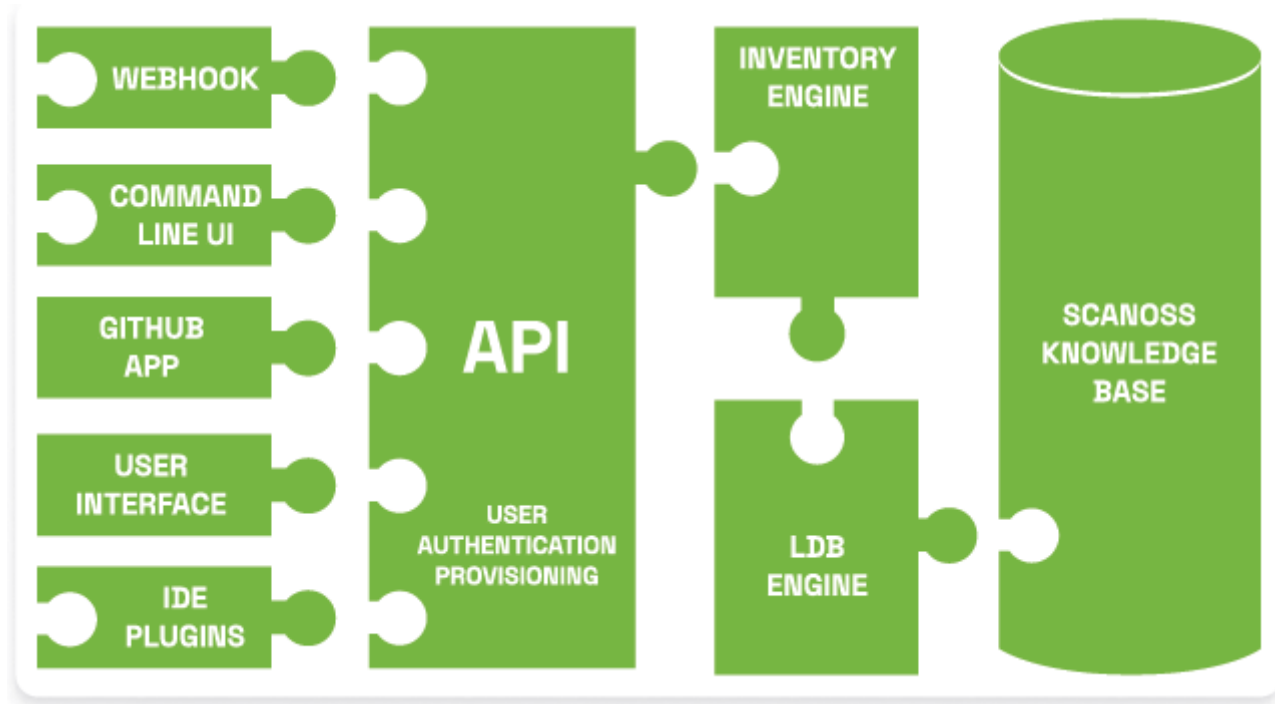
SCANOSS Solutions Product About us Get in touch

Minimize the risk of Open Source, while coding, carefree.

It's time to democratize Open Source Risk Management. Discover how you can benefit from it in true 360 degrees.

Get in touch

SCANOSS Platform



SCANOSS Platform

❑ Database Engine(LDB Engine)

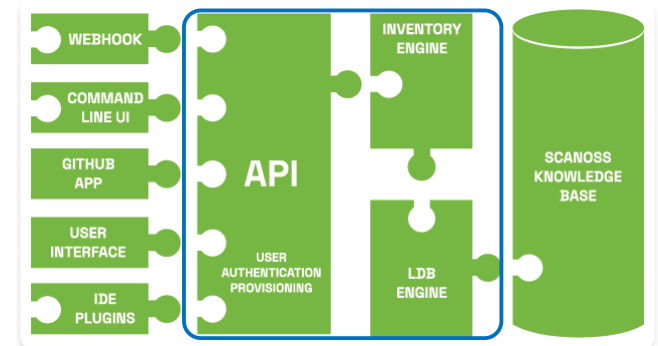
- GitHub : <https://github.com/scanoss/ldb>
- Scanning 시 쿼리를 최소화 하기 위해 설계된 Database Engine

❑ Inventory Engine

- GitHub : <https://github.com/scanoss/engine>
- Source Code File 혹은 미리 계산된 Winninging 핑거프린트(WFP)와 KB Data를 비교
- Scan 결과를 JSON으로 출력

❑ API

- GitHub : <https://github.com/scanoss/api>
- Inventory Engine과 통신하며 Audit 관리, User 관리, 권한 관리, Project 관리, Authentication 등의 기능을 제공



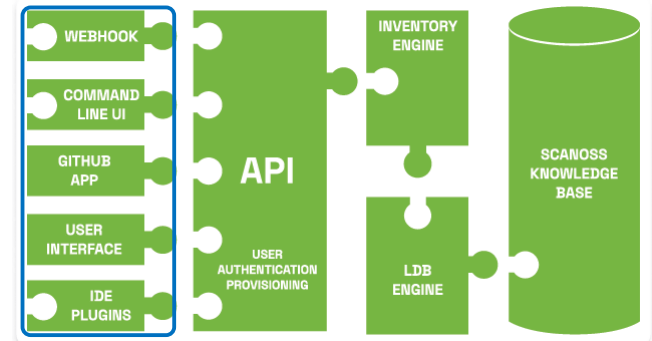
SCANOSS Platform

□ Webhook

- GitHub : <https://github.com/scanoss/webhook>
- Github, GitLab, BitBucket API와 Integration 하여 Source Code Scanning을 제공

□ Command-line Interface

- GitHub : <https://github.com/scanoss/scanoss.py>
- REST API를 이용해 CLI를 구성하는 예제를 언어별로 제공했으나, Deprecate 시키는 중
 - Python : <https://github.com/scanoss/scanner.js>
 - js : <https://github.com/scanoss/scanner.js>
 - (deprecated) Java : <https://github.com/scanoss/scanner.java>
 - scanner.php : <https://github.com/scanoss/scanner.php>
 - (deprecated) C : <https://github.com/scanoss/scanner.c>
- 최근 Python을 통한 CLI로 집중
 - Pypi : <https://pypi.org/project/scanoss/>



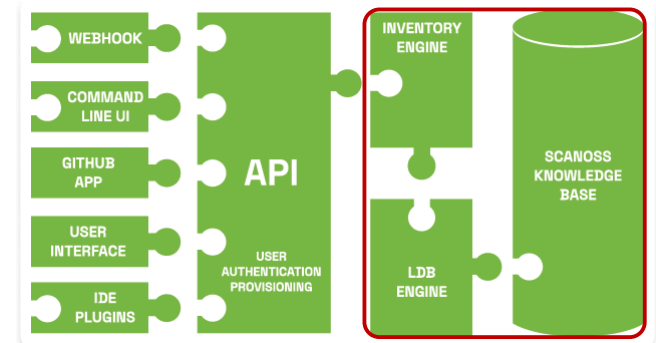
SCANOSS Platform

❑ Mining Tool

- GitHub : <https://github.com/scanoss/minr>
- Data mining을 하여 KB를 생성하는 command-line 툴
- Component의 Source Code를 Download, Extract Indexing
- Minr는 복수개의 Machine / Instance에서 실행되어 결과(data)를 쉽게 연결 시켜 하나의 KB로 구성할 수 있음
- 기업에서는 Proprietary Code에 대한 KB를 생성하여 별도로 관리 가능

❑ Open Indexing Algorithm(wfp)

- GitHub : <https://github.com/scanoss/wfp>
- Full File 해싱 알고리즘 대비 Snippet에 강점을 두는 Winnowing 알고리즘을 Implement
 - 학계에서 표절 검사에 널리 사용되는 알고리즘

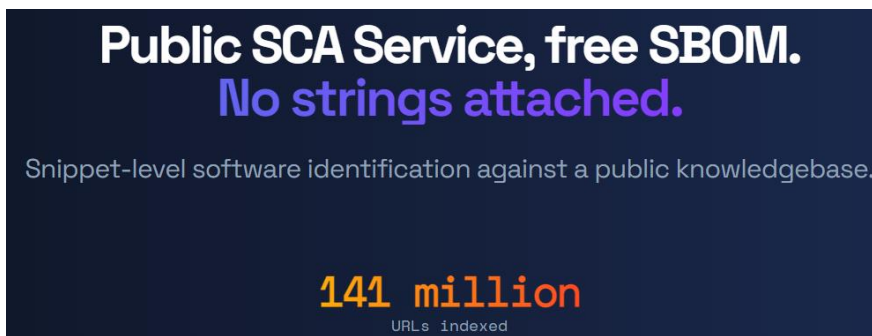


Open Data

□ OSSKB

- SCANOSS를 위한 Open Data KB
- Minr을 이용해 KB가 생성되고, LDB Engine을 통해 조회

□ 무상으로 제공되고 있으며, Minr를 이용해 KB를 지속적으로 업데이트중

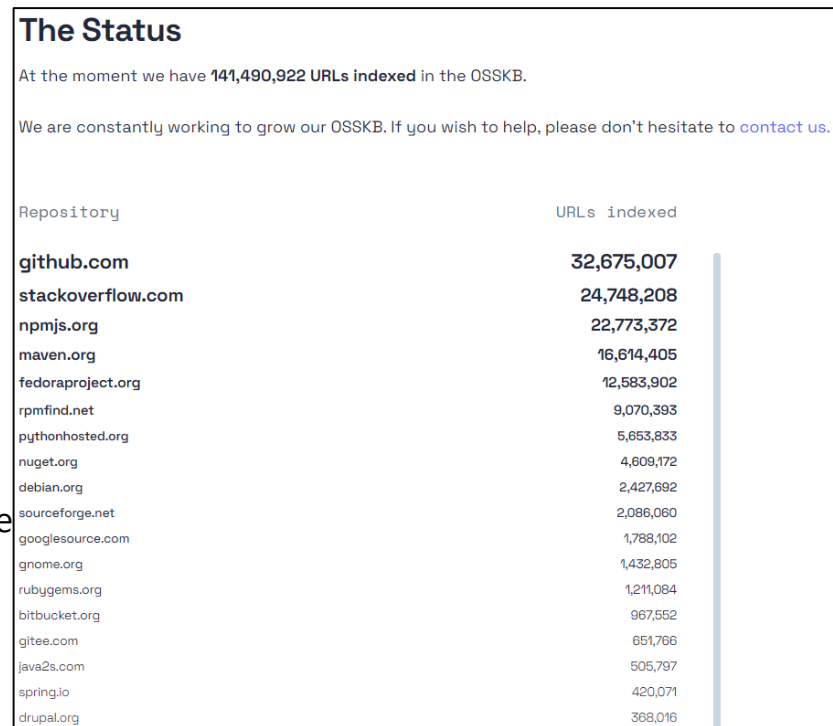


Public SCA Service, free SBOM.
No strings attached.
Snippet-level software identification against a public knowledgebase.

141 million
URLs indexed

□ 저장되는 정보

- Components : ID, name, author, url, license, dependencies
- Files : ID, component ID, file path, file size, license
- Snippets : ID, file ID line number where snippet starts



The Status

At the moment we have **141,490,922** URLs indexed in the OSSKB.

We are constantly working to grow our OSSKB. If you wish to help, please don't hesitate to [contact us](#).

| Repository | URLs indexed |
|--------------------------|-------------------|
| github.com | 32,675,007 |
| stackoverflow.com | 24,748,208 |
| npmjs.org | 22,773,372 |
| maven.org | 16,614,405 |
| fedoraproject.org | 12,583,902 |
| rpmfind.net | 9,070,393 |
| pythonhosted.org | 5,653,833 |
| nuget.org | 4,609,172 |
| debian.org | 2,427,692 |
| sourceforge.net | 2,086,060 |
| googlesource.com | 1,788,102 |
| gnome.org | 1,432,805 |
| rubygems.org | 1,211,084 |
| bitbucket.org | 967,552 |
| gitee.com | 651,766 |
| java2s.com | 505,797 |
| spring.io | 420,071 |
| drupal.org | 368,016 |

Software Transparency Foundation

□ Software 공급망의 투명성을 보장하자는 취지의 재단

The screenshot displays the website for the Software Transparency Foundation. At the top left is the logo, which consists of the letters 'STF' in a stylized font. To the right of the logo is the text 'Software Transparency Foundation'. Further right are navigation links for 'Mission', 'Sponsors', 'Projects', and 'Board'. A 'Contact Us' button is located in the top right corner. The main heading is 'OUR MISSION' followed by 'Solving Software Supply Chain Transparency'. Below this, there are four distinct sections, each with an icon, a title, and a description of a goal.

Software Transparency Foundation Mission Sponsors Projects Board [Contact Us](#)

OUR MISSION

Solving Software Supply Chain Transparency

- **Driving adoption of SBOM**


By providing Open Source tools for generating, notarizing, relating and validating Software Bills of Materials (SBOM), we will facilitate broad adoption. This will ease Software Supply Chain traceability, drastically improving cybersecurity and license compliance.

Our goal is to provide a complete set of Open Source Software Composition Analysis (SCA) tools, which otherwise are out of reach for small and mid-sized organizations.
- **Enabling a decentralized validating entity**

The natural immutability of blockchain technologies provide an excellent foundation for building a decentralized validating entity for Software Bill of Materials.

Our goal is to allow public validation of SBOMs by comparing their cryptographic checksums published in the blockchain, along with their date stamps and relationships to preceding SBOMs in the supply chain.
- **Facilitating SBOM interoperability**

The SBOM ledger offers a decentralized, format-agnostic registry of Software Bills of Materials and their relation to their parents. This allows easily tracing and validating software components across the supply chain.

We aim at solving the traceability issue of software components across the supply chain without any depth limitations.
- **Abstraction layer for Blockchain transaction fees**

Our goal is not only to develop and maintain the decentralized SBOM ledger, but also to provide easy access to our sponsors by eliminating the need of cryptocurrencies to cover transaction fees.

We are committed to providing our sponsors with free access to the SBOM ledger in the blockchain.

Software Transparency Foundation

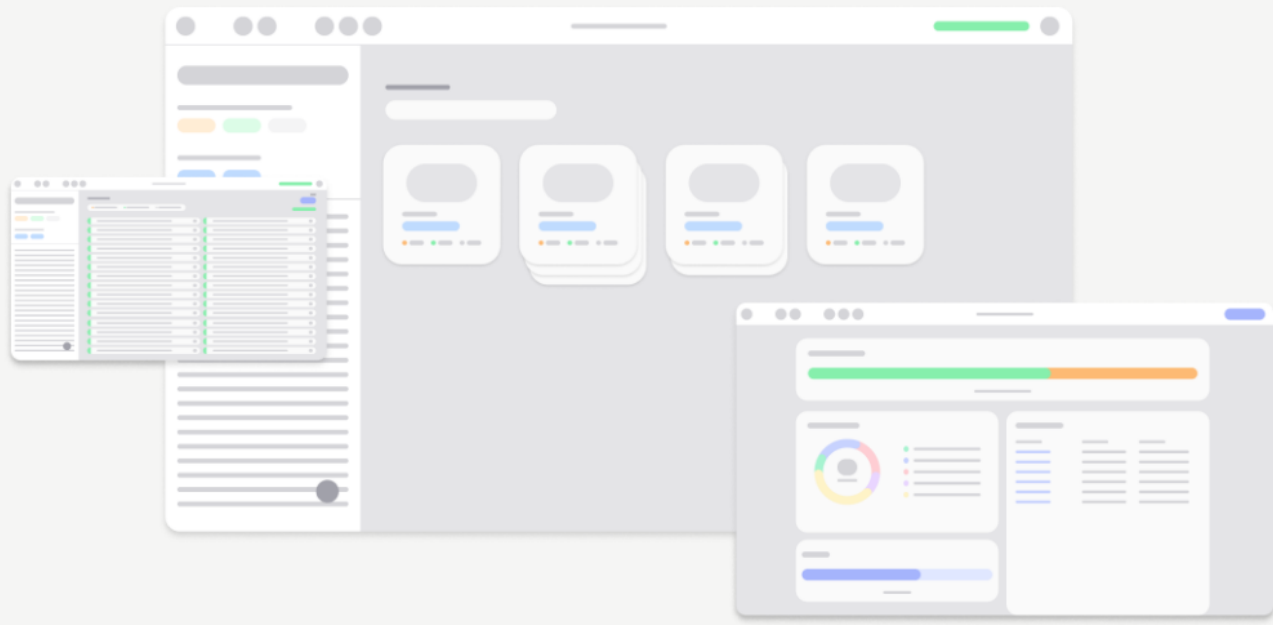
- SCANOSS 측으로 부터 기여받은 OSSKB를서비스로 제공

Open Source Knowledge Base Public Service

This public and free Open Source identification service allows anyone to generate their own Software Bills of Materials and to validate Open Source license compliance.

Thanks to our sponsors, we have launched a public API service for Open Source Inventorying which offers snippet level detection of known Open Source components. The service is available at osskb.org. At the moment we have 136,958,197 URLs indexed.

We are constantly working to grow our OSSKB, [check it out here](#).



Software Transparency Foundation

❑ Blockchain을 이용한 SBOM 원장 제공

SBOM Ledger

The SBOM Ledger is decentralized, immutable and stored in the Blockchain, which allows integrity validation and relation between SBOMs and their parents, enabling completeness of visibility into the supply chain.

Test Ledger

SBOM Broadcasting

Parent SBOM (optional)

Add Hash +

Add TXID +

SBOM Data*

SBOM hash

Hash

Hash type

SHA224

SHA256

SHA384

Add Hash +

Your Token Id

Token Id

Publish to Blockchain

Your generated JSON file

```
{
  "type": "sbom",
  "parent_sboms": [],
  "sbom_data": {
    {
      "checksum": "",
      "algorithm": "SHA224"
    }
  }
}
```

SCANOSS 활용

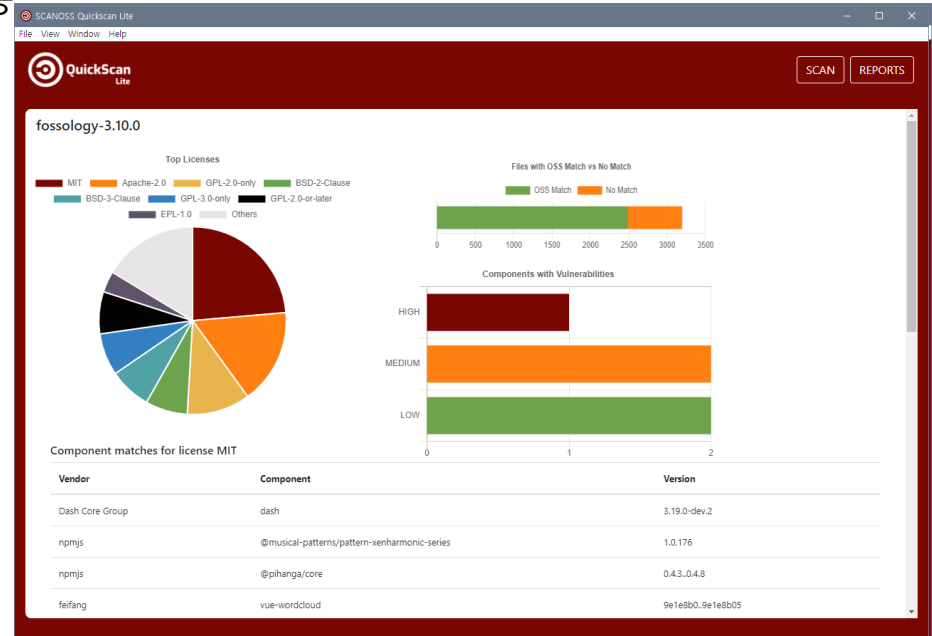
SCANOSS Quickscan Lite

□ SCANOSS 플랫폼을 이용한 Scanning Tool

- Windows, Linux, Mac OS용 실행파일 지원
- GPL-2.0으로 공개됨 : <https://github.com/scanoss/quickscan>
- 현재 지원은 중단되고, Audit Workbench 사용을 권장

□ 동작 방식

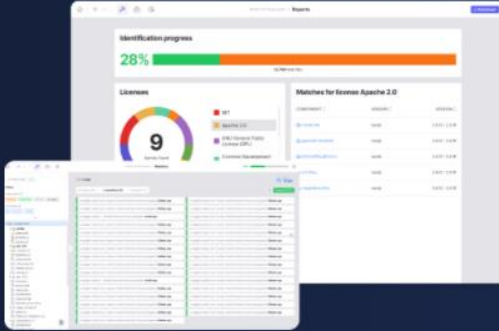
1. 스캔 대상 폴더 선택 → 스캔
2. (자동) Fingerprint 생성
3. (자동) 스캔을 위해 OSSKB API에 Fingerprint 전송
 - Source Code는 OSSKB로 전송되지 않음
4. (자동) 도식화된 License 정보를 보여줌
 - 파이 차트에서 특정 License를 클릭하면 해당 License가 검출된 Component를 보여줌



SCANOSS Audit Workbench

❑ SCANOSS 플랫폼을 이용한 Scanning 및 Audit Tool





- Windows, Linux, Mac OS용 실행파일 지원 : <https://www.scanoss.com/product>
- GPL-2.0으로 공개됨 : <https://github.com/scanoss/audit-workbench>



**The first
multiplatform
OSS auditing
app.**

Auditing your source code for license compliance has never been easier. Simply download the SCANOSS Audit Workbench and scan your source code directory to find and identify open source components. Generate your SPDX-Lite software bill of materials (SBOM) with the press of a button.

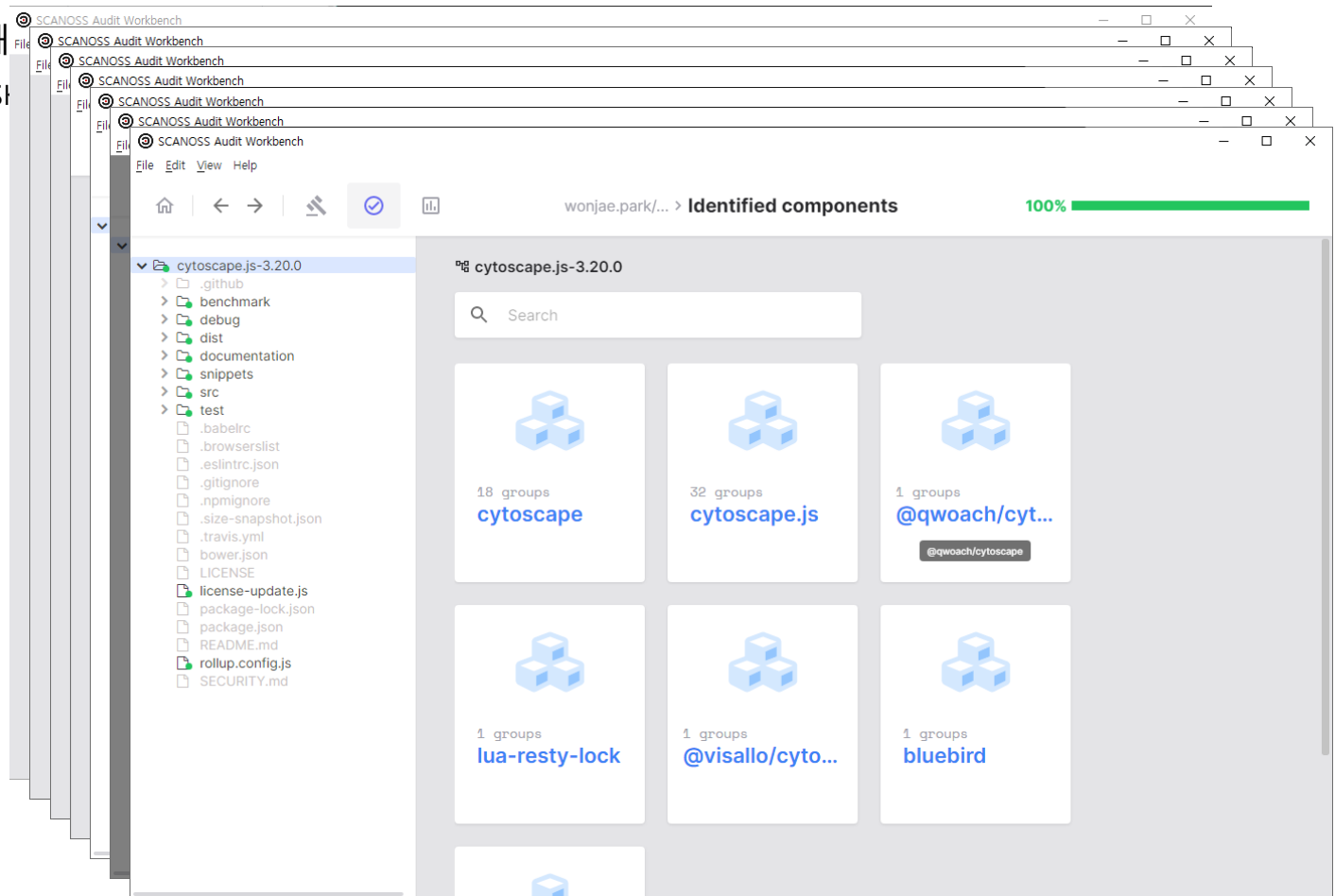
[Download Workbench \(beta\)](#)

Also available on    

SCANOSS Audit Workbench

□ 동작 방식

1. New Project → 스캔 대상 폴더 선택
 - License, API, Ledger Token
2. (자동) Fingerprint 생성
3. (자동) 스캔을 위해
 - Source Code는 OSS
4. 검출 결과 확인
5. Identification



SCANOSS Audit Workbench

```

1 file=99d8e41eb76fff9bf9ald67f4bd
2 4=b03faabe
3 5=23bfe641
    
```

```

11930     "server": {
11931     "version": "4.4.0",
11932     "kb_version": {
11933     "monthly": "22.01",
11934     "daily": "22.01.17"
    
```

cytoscape.js/collection-astar.js

github.com cytoscape.js/collection-astar.js at v3.20.0 · cytoscape/cytoscape.js

maxkfranz Add test for A*-Dijkstra consistency #2880 ✓

2 contributors

268 lines (234 sloc) | 10 KB

```

1 var expect = require('chai').expect;
2 var cytoscape = require('../src/test.js', cytoscape);
    
```

https://osskb.org/api/file_contents/de56ed915a7fb5e02f3a7542112fceb7

```

var expect = require('chai').expect;
var cytoscape = require('../src/test.js', cytoscape);

describe('Algorithms', function() {
  describe('eles.aStar()', function() {

    var cy;
    var nodes;
    var edges;

    beforeEach(function(done) {
      cytoscape({
        elements: {
          nodes: [
    
```

| | A | B | C | D | E | F | G | H | I | J |
|-----|-----------------|-------|--------------------|------------------|----------------------|-------------------|---------------------------|----------------------|---------|------------|
| 1 | inventory_usage | notes | identified_license | detected_license | identified_component | detected_componer | path | purl | version | |
| 239 | 34 file | n/a | MIT | MIT | cytoscape | cytoscape | /src/extensions/r | pkg:npm/cytoscape | 3.18.1 | |
| 240 | 35 file | n/a | MIT | MIT | cytoscape | cytoscape | /src/extensions/r | pkg:npm/cytoscape | 3.15.1 | |
| 241 | 36 file | n/a | MIT | MIT | cytoscape | cytoscape | /src/extensions/r | pkg:npm/cytoscape | 3.12.1 | |
| 242 | 37 file | n/a | MIT | MIT | cytoscape | cytoscape | /src/extensions/r | pkg:npm/cytoscape | 3.10.1 | |
| 243 | 38 file | n/a | MIT | MIT | cytoscape | cytoscape | /src/style/index.js | pkg:npm/cytoscape | 3.16.1 | |
| 244 | 39 file | n/a | MIT | MIT Unlicense | cytoscape.js | cytoscape.js | /test/collection-astar.js | pkg:github/cytoscape | 3.9.2 | |
| 245 | 39 file | n/a | MIT | MIT Unlicense | cytoscape.js | cytoscape.js | /test/collection-astar.js | pkg:github/cytoscape | 3.9.2 | |
| 246 | 39 file | n/a | MIT | MIT Unlicense | cytoscape.js | cytoscape.js | /test/lib/chai-include.js | pkg:github/cytoscape | 3.9.2 | |
| 247 | 40 snippet | n/a | MIT | MIT Unlicense | cytoscape.js | cytoscape.js | /test/collection-astar.js | pkg:github/cytoscape | 3.3.0 | |
| 248 | 41 file | n/a | MIT | MIT Unlicense | cytoscape.js | cytoscape.js | /test/collection-astar.js | pkg:github/cytoscape | 3.2.0 | |
| 249 | 41 file | n/a | MIT | MIT Unlicense | cytoscape.js | cytoscape.js | /test/collection-astar.js | pkg:github/cytoscape | 3.2.0 | /MIT.txt", |
| 250 | 41 file | n/a | MIT | MIT Unlicense | cytoscape.js | cytoscape.js | /test/collection-astar.js | pkg:github/cytoscape | 3.2.0 | |
| 251 | 42 file | n/a | MIT | MIT Unlicense | cytoscape.js | cytoscape.js | /test/collection-astar.js | pkg:github/cytoscape | 3.2.10 | |
| 252 | 43 file | n/a | MIT | MIT Unlicense | cytoscape.js | cytoscape.js | /test/collection-astar.js | pkg:github/cytoscape | 3.4.7 | |
| 253 | 43 file | n/a | MIT | MIT Unlicense | cytoscape.js | cytoscape.js | /test/collection-astar.js | pkg:github/cytoscape | 3.4.7 | |
| 254 | 44 file | n/a | MIT | MIT Unlicense | cytoscape.js | cytoscape.js | /test/collection-astar.js | pkg:github/cytoscape | 3.13.3 | |

SCANOSS.py

□ SCANOSS 플랫폼을 이용한 Scanning Tool

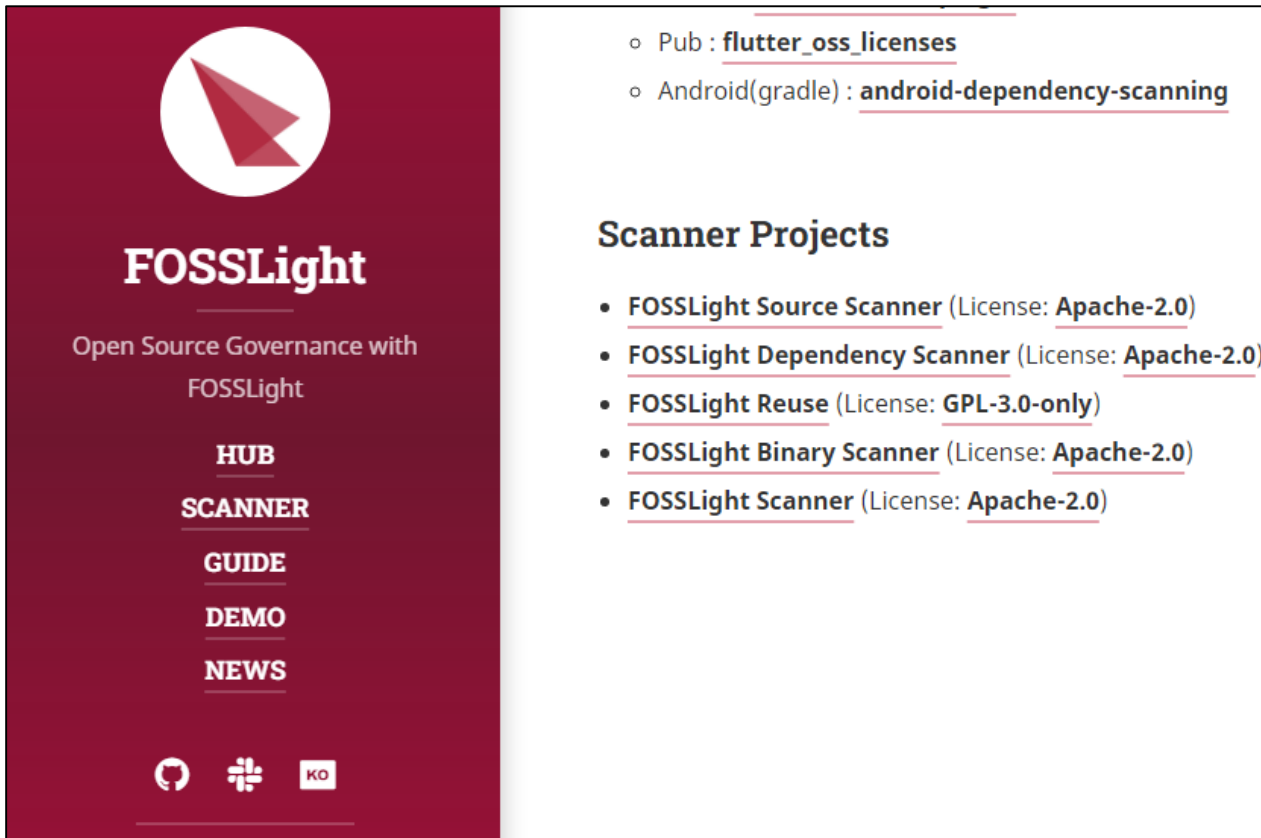
- Pypi를 통한 설치 제공 : <https://pypi.org/project/scanoss/>
- GPL-2.0 -> MIT로 공개됨 : <https://github.com/scanoss/scanoss.py>
- Python 3.7 이상 지원
 - Python 3.6 에서 설치 가능하나, 과거 버전만 지원(~ v0.6.11)

```
(temp3.7) [[woniae /home/woniae/scanoss_test]$ find . -type f | wc -l
702
(temp3.7) [[woniae /home/woniae/scanoss_test]$ du -hs cytoscape.js-3.20.0/
24M   cytoscape.js-3.20.0/
(temp3.7) [[woniae /home/woniae/scanoss_test]$ scanoss-py scan -o scanner-output.json cytoscape.js-3.20.0/
```

FOSSLight Source Scanner

❑ ScanCode, SCANOSS.py를 이용한 Scanning Tool

- Pypi를 통한 설치 제공 : <https://pypi.org/project/fossilight-source/>
- Apache-2.0으로 공개됨 : https://github.com/fossilight/fossilight_source_scanner
- Python 3.6 이상 지원
 - Python 3.6 에서는 SCANOSS 별도 설치 필요



The screenshot shows the FOSSLight website interface. On the left, there is a dark red sidebar with a white circular logo containing a red triangle. Below the logo, the text "FOSSLight" is displayed in white. Underneath, it says "Open Source Governance with FOSSLight". A vertical list of navigation links is provided: "HUB", "SCANNER", "GUIDE", "DEMO", and "NEWS". At the bottom of the sidebar are icons for GitHub, a community icon, and a Korean flag (KO). The main content area on the right has a white background. It features a list of scanner projects with their respective licenses: "flutter_oss_licenses" (License: Apache-2.0) and "android-dependency-scanning" (License: Apache-2.0). Below this, a section titled "Scanner Projects" lists five items: "FOSSLight Source Scanner (License: Apache-2.0)", "FOSSLight Dependency Scanner (License: Apache-2.0)", "FOSSLight Reuse (License: GPL-3.0-only)", "FOSSLight Binary Scanner (License: Apache-2.0)", and "FOSSLight Scanner (License: Apache-2.0)".

FOSSLight Source Scanner

```
(temp3.6) [[wonjae /home/wonjae/scanoss_test]$ fosslight_source --help
```

```
FOSSLIGHT SOURCE
```

Usage: fosslight_source [option1] <arg1> [option2] <arg2>...

FOSSLight Source uses ScanCode, a source code scanner, to detect the copyright and license phrases contained in the file. Some files (ex- build script), binary files, directory and files in specific directories (ex-test) are excluded from the result. And removes words such as "-only" and "-old-style" from the license name to be printed. The output result is generated in Excel format.

Options:

Mandatory

-p <source_path> Path to analyze source

Optional

-h Print help message

-v Print FOSSLight Source Scanner version

-j Generate raw result of scanners in json format

-m Print the Matched text for each license on a separate sheet (ScanCode Only)

-o <output_path> Output path

(If you want to generate the specific file name, add the output path with file

name.)

-f <format> Output file format (excel, csv, opossu)

-s <scanner> Select which scanner to be run (scancode, scanoss, all)

```
(temp3.6) [[wonjae /home/wonjae/scanoss_test]$ fosslight_source -p cytoscape.js-3.20.0/ -s all
```

```
-Output Directory: /home/wonjae/scanoss_test
```

```
Parsing Log: 'TOTAL FILE COUNT: 55'
```

```
Scan Result: 'true'
```

```
Searching cytoscape.js-3.20.0/ for files to fingerprint...
```

```
Fingerprinting /
```

```
Writing results to /home/wonjae/scanoss_test/scanoss_raw_result.json...
```

```
\canning |#####| 283/283
```

```
SCANOSS Start parsing cytoscape.js-3.20.0/
```

```
|--Number of files detected with SCANOSS: 269
```

```
Writing Output file(FOSSLight-Report_20220207_211300):True
```

FOSSLight Source Scanner

| 1 | A | B | C | D | E | F | G | H | I | J | K | L | M |
|-----|-----|-------------------------|--------------|-------------|---------------|---|----------|----------------|---------|---------|-----------------------|---|----------------|
| | ID | Source Name or Path | OSS Name | OSS Version | License | Download Location | Homepage | Copyright Text | Exclude | Comment | scanoss_matched_lines | scanoss_fileURL | scanoss_vendor |
| 242 | 241 | src/extensions/renderer | cytoscape.js | 3.11.0 | unlicense,mit | https://github.com/cytoscape/cytoscape.js | | | | | 100% (all) | https://osskb.org | cytoscape |

Dashboard | Project List | 4146_Identify | OSS List

Project Name: testtestest | Created: 박원재 CTO SW센터 (2022-01-10)

3rd party | **SRC** | BIN | BOM | BAT(Optional) | Request

Show Comment History | Comment Edit

Upload Analysis Result | Project Search | Not Applicable

FOSSLight Report

Upload Drag & Drop Files

FOSSLight-Report_20220207_211300.xlsx

2022-02-08 12:17:19

Loaded List

| ID | project Name | project |
|----|--------------|---------|
| | | |

OSS bulk registration | Check OSS Name | Check License | Bulk Edit

| ID | Source Name or Path | OSS Name | OSS Version | License | Download Location |
|-----|--------------------------------|--------------------------------------|-------------|---------------|---|
| 517 | src/selector/expressions.js | cytoscape Unconfirmed open source | 3.3.0 | MIT | https://github.com/cytoscape/cytoscape.js |
| 518 | src/selector/index.js | cytoscape Unconfirmed open source | 3.3.0 | MIT | https://github.com/cytoscape/cytoscape.js |
| 519 | src/selector/matching.js | cytoscape Unconfirmed open source | 3.3.0 | MIT | https://github.com/cytoscape/cytoscape.js |
| 520 | src/selector/parse.js | cytoscape Unconfirmed open source | 3.3.0 | MIT | https://github.com/cytoscape/cytoscape.js |
| 317 | benchmark/a-star.js | cytoscape.js Unconfirmed ve | 3.3.0 | Unlicense,MIT | https://github.com/cytoscape/cytoscape.js |
| 318 | benchmark/add-remove-class.js | cytoscape.js Unconfirmed ve | 3.3.0 | Unlicense,MIT | https://github.com/cytoscape/cytoscape.js |
| 319 | benchmark/add-remove.js | cytoscape.js Unconfirmed ve | 2.7.1 | Unlicense,MIT | https://github.com/cytoscape/cytoscape.js |
| 320 | benchmark/add.js | cytoscape.js Unconfirmed ve | 3.3.0 | Unlicense,MIT | https://github.com/cytoscape/cytoscape.js |
| 321 | benchmark/all-are-neighbors.js | cytoscape.js Unconfirmed ve | 3.1.0 | Unlicense,MIT | https://github.com/cytoscape/cytoscape.js |
| 322 | benchmark/all-are.js | cytoscape.js Unconfirmed ve | 3.3.0 | Unlicense,MIT | https://github.com/cytoscape/cytoscape.js |
| 324 | benchmark/all/index.js | cytoscape.js Unconfirmed ve | 3.3.0 | Unlicense,MIT | https://github.com/cytoscape/cytoscape.js |

FOSSLight Hub - Whale

fossilight.lge.com FOSSLight Hub

There exists another OSS which has same download location. Please click "Change OSS Name" if you want to change to the registered OSS Name.

| Result | Download location | OSS name (now) | Registered OSS name (to be changed) |
|--|---|---|--|
| <input type="checkbox"/> Change Add | https://github.com/openresty/luar-resty-lock | lua-resty-lock Unconfirmed open source | openresty-lua-resty-lock |
| <input checked="" type="checkbox"/> Change Add | https://www.npmjs.com/package/cytoscape | cytoscape Unconfirmed open source | cytoscape.js |
| <input type="checkbox"/> Change Add | https://www.npmjs.com/package/%40visallo/cytoscape | @visallo/cytoscape Unconfirmed open source | npm:%40visallo |
| <input type="checkbox"/> Change Add | https://www.npmjs.com/package/%40qwoach/cytoscape | @qwoach/cytoscape Unconfirmed open source | npm:%40qwoach |

Change OSS Name | Add Nickname



THANK YOU