LG Electronics Open Source Program Office

Open Compliance Summit 2024 Review

2024. 11. 26

박원재, LG 전자

wonjae.park@lge.com

LG Open Source

CONTENTS

- Open Compliance Summit
- Open Source Project의 Licensing 정책
- AI Licensing
- Open Source 분석 도구
- Open Source Database

Open Compliance Summit

Open Compliance Summit

- Linux Foundation이 주최하는 연례 행사

- Open Source License 및 보안, 수출 규제 등
  Software 공급망에서의 전반적인 Compliance를 주제로 함

- 매년 겨울, 일본 도쿄/요코하마 등지에서 개최 됨

THE LINUX FOUNDATION

https://events.linuxfoundation.org/open-compliance-summit/

Open Compliance Summit

Photo by Linux Foundation

Open Compliance Summit 2024

Wednesday, October 30

08:00 JST Registration & Badge Pick-up — MAIN HALL FOYER (5F)
09:10 JST Keynote: Welcome + Opening Remarks - Shane Coughlan, The Linux Foundation — HALL B 1-2 (4F)
09:20 JST Keynote: Global Compliance Trends - Jimmy Ahlberg, Ericsson — HALL B 1-2 (4F)
09:25 JST Keynote: Compliance in Japan - An Overview by Ayumi Watanabe — HALL B 1-2 (4F)
09:35 JST Keynote: The Dark Ages of GPL Compliance 20 Years Ago - Harald Welte, sysmocom — HALL B 1-2 (4F)
10:00 JST Keynote: License Changes: Why They Happen and What to Do - Heather Meeker, Tech Law Partners LLP — HALL B 1-2 (4F)
10:25 JST Keynote: Product Regulation for Software. Where Is the World Going? - Ciarán O'Riordan, OpenForum Europe & Catharina Maracke, Software Compliance Academy — HALL B 1-2 (4F)
10:50 JST Coffee Break — HALL B 1-2 (4F)
11:20 JST Sponsored Keynote: Fujitsu's OSS Standards Conformance and AI Management System Standardization Participation - Tadayuki Osaki & Dr. Yuchang Cheng, Fujitsu Limited — HALL B 1-2 (4F)
11:30 JST Open Source as a Strategic Asset in M&A: The Evolving Landscape of Transactional Risk Management - Andrew Katz & Stephen Pollard, Orcro; Heather Meeker, Tech Law Partners; Lewis Parle, Lockton; Byron Frost & Ayako... — HALL B 1-2 (4F)
12:05 JST Lunch — HALL B 1-2 (4F)
13:30 JST Protecting Open Source and Companies' Intellectual Property Rights - Keith Bergelt, Open Invention Network — HALL B 1-2 (4F)
13:55 JST Managing Compliance Artifacts as Code with OSCAL and Compliance-Trestle - Chris Butler, Red Hat — HALL B 1-2 (4F)
14:20 JST OSS-Rules: An Open Framework for Automated Open Source License Compliance at Scale - Diego Jorquera & Oscar Valenzuela, Amazon — HALL B 1-2 (4F)
15:00 JST Coffee Break — HALL B 1-2 (4F)
15:25 JST Rapid Handling of Vulnerabilities in the Supply Chain with SBOM and VEX - Yoshihisa Morizumi, Wang Mingyu & Lei Maohul, Fujitsu Limited — HALL B 1-2 (4F)
15:50 JST AI Panel: Jonathan Torres, Meta & Heather Meeker, OMM — HALL B 1-2 (4F)
16:30 JST Unpacking AI Model Licensing: Navigating Intellectual Property Challenges - Ye Tao, Grandall Law Firm — HALL B 1-2 (4F)
16:55 JST AI Governance in China - A Quick Overview of Developments To Date - King Gao, SecTrend

Thursday, October 31

08:00 JST Registration & Badge Pick-up — HALL B FOYER (4F)
09:30 JST Keynote: Opening Remarks - Shane Coughlan, The Linux Foundation — HALL B 1-2 (4F)
09:40 JST Keynote: Lessons Learned from the Integration Journey of 2 OSPOs in LY Corporation - Seoyeon Lee, LINE+ Corporation — HALL B 1-2 (4F)
10:05 JST Keynote: Adjusting in-Production Processes for OSS Management - Mary (Meixia) Wang, Volvo Car Corporation — HALL B 1-2 (4F)
10:30 JST Keynote: OpenChain Practices in Real Use, How OpenChain Helps Nokia in Its Open Source Journey - Eleftheria Stefanaki, Nokia — HALL B 1-2 (4F)
10:55 JST Coffee Break — HALL B 1-2 (4F)
11:25 JST Using Case Studies to Inspire: The Value and Process of Sharing Experience with the Community - Russ Eling, OSS Consultants; Yang Hanbo (Tony), openEuler; Seoyeon Lee, LINE; Masato Endo, Toyota — HALL B 1-2 (4F)
11:55 JST Rethinking GPL Interpretations to Promote Licence Compatibility - James Bottomley, Microsoft — HALL B 1-2 (4F)
12:15 JST Lunch — HALL B 1-2 (4F)
13:30 JST Open Source Compliance: What Is Under the Carpet? - Armijn Hemel, Tjaldur Software Governance Solutions & Oscar Valenzuela, Amazon — HALL B 1-2 (4F)
15:05 JST Open Source Diligence: From Risk Assessment to Post-Close Integration - Jari Koivisto, Freelance — HALL B 1-2 (4F)
15:30 JST Coffee Break — HALL B 1-2 (4F)
16:05 JST Adopt Supply Chain Standards, Create an Ecosystem, and Assist Industry Development - King Gao, openEuler and OpenHarmony — HALL B 1-2 (4F)
16:30 JST From an Open Data Set to Standardized Management Processes. Step One: Cryptographic Algorithms List - Agustin Benito Bethencourt & Julian Coccia, SCANOSS — HALL B 1-2 (4F)
16:55 JST ClearlyDefined: Sharing your Open Source License Scan Data with Others - Tom Bedford, Bloomberg — HALL B 1-2 (4F)
17:10 JST Compliance and Integrity in the Software Supply Chain with Software Heritage: A Call to Action - Roberto Di Cosmo, Director, Software Heritage — HALL B 1-2 (4F)
17:35 JST Keynote: Closing Remarks - Shane Coughlan, The Linux Foundation — HALL B 1-2 (4F)

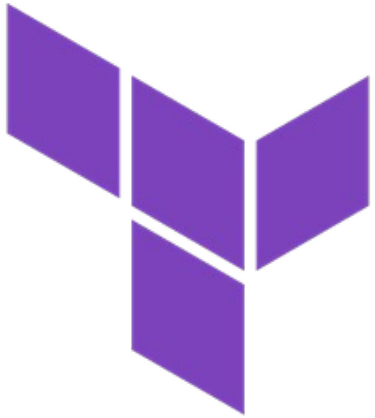
Chatham House Rule

- 정보의 자유로운 공유를 촉진하기 위한 규칙
- 정보는 자유롭게 사용할 수 있지만, 정보의 출처(발언자, 혹은 소속 등)은 공개할 수 없음

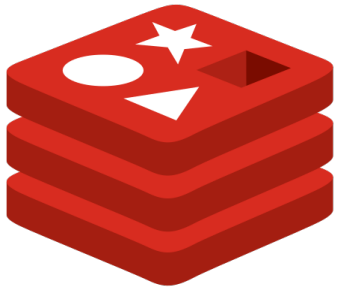
Open Source Project의 Licensing 정책

Open Source Project의 Licensing 정책 (Background)

elasticsearch



HashiCorp
Terraform



redis

Open Source Project의 Licensing 정책

Open Source Project의 License 변경?

Licensing 정책의 변경!

Open Source Project의 Licensing 정책

- 전제 : 기업의 Open Source 프로젝트 중단은 비도덕적 행위가 아님

- 라이센스 변경의 유형
  - Closed -> Open
  - Open -> Closed
  - Open -> Open
    - Permissive -> Restrictive
    - Restrictive -> Permissive

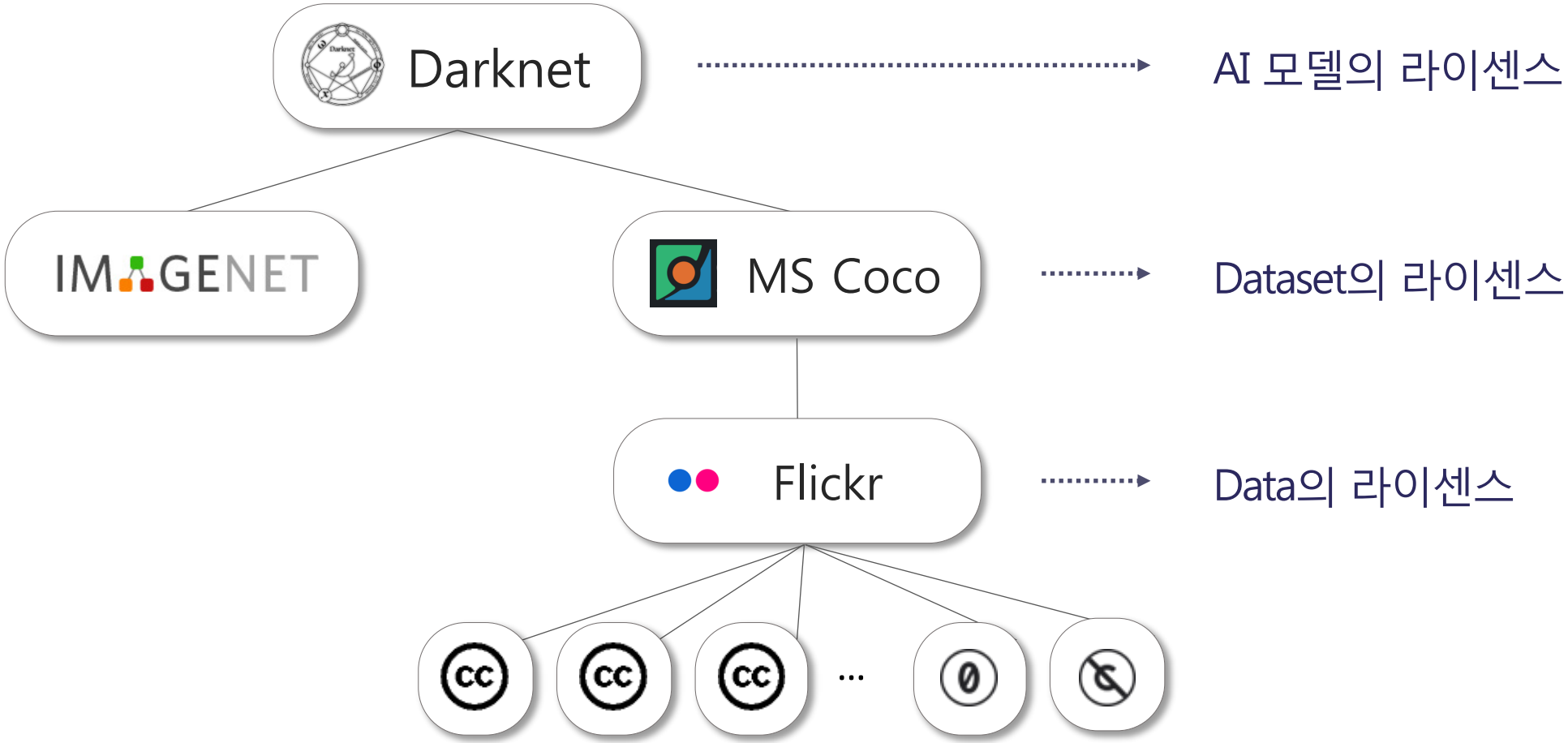
- 기업 관점에서의 Licensing 정책 변경의 요인
  - 창립자의 이직
  - 사모 펀드의 기업 매입
  - 투자자들의 오해 : 무임 승차자(Freeloader)를 고객으로 변모 시킬수 있다는 착각

Open Source Project의 Licensing 정책

- Licensing 정책 변경의 결과
  - 종종 Project Fork 발생
- 투자자를 위한 조언
  - Open Source 비즈니스 모델이 작동하는 이유를 고찰
- Open Source 커뮤니티를 위한 조언
  - License 정책 변경에 대한 비난 자제
- Open Source 사용 기업을 위한 조언
  - Open Source는 비용이 들지 않는다는 가정을 지양
  - 의존하는 주요 인프라 프로젝트를 적극적으로 지원
- Open Source를 발전시키기 위한 Resource에 대한 고민 필요

AI Licensing

AI Licensing (Background)



AI Licensing (Background)

연합뉴스

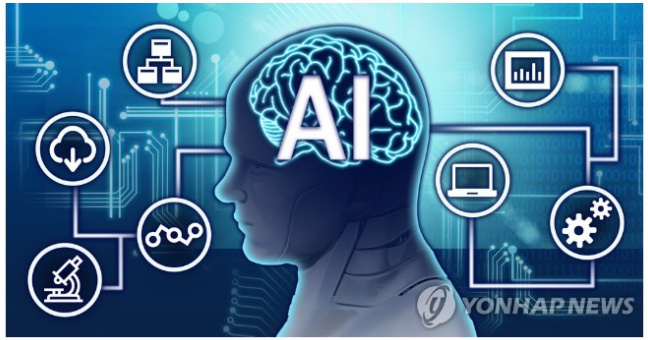
공개 데이터로 훈련하는 AI 제동 걸리나…사상 첫 소송 제기

송고시간 | 2022-11-24 11:53

| 美 유명 프로그래머, MS·깃허브 등 프로그래밍 AI 제작사에 소송

(서울=연합뉴스) 임화섭 기자 = 인공지능(AI)이 대량의 공개 데이터를 퍼가면서 스스로 학습하는 'AI 훈련' 기법에 대해 "사람이 힘겹게 만든 창작물을 무단도용하는 저작권 침해"라고 주장하는 첫 소송이 미국에서 제기됐다.

AI는 음성·이미지 인식, 자율주행 등 산업뿐 아니라 미술, 음악과 같은 예술까지 다양한 분야에 쓰이면서 인간의 일자리를 위협하고 있으나 이 학습 방법이 합법인지 모호한 경우가 대부분이어서 IT 업계는 소송 결과에 촉각을 곤두세우고 있다.



HOME > GLOBAL

게티이미지, 저작권 침해 혐의로 스태빌리티AI 고소

정병일 위원 | 승인 2023.01.18 18:26

| 스테이블 디퓨전 개발한 스태빌리티AI 상대 소송

(사진=셔터스톡)

이미지 판매 사이트인 게티이미지가 이미지 생성 AI도구를 개발한 회사인 스태빌리티 AI에 대해 저작권 위반혐의로 소송을 제기했다.

예술가 3인 '스테이블디퓨전'과 '미드저니' 저작자 대상 소송

사라 앤더슨, '스테이블디퓨전'과 '미드저니' 저작자 대상 소송
"원작 예술가의 동의 없이" 웹에서 스크랩한 50억 이미지 소송
"수백만 명의 예술가"의 권리 침해

Editor. 이호선 기자 | 입력 2023.01.17 16:00 | 댓글 0

AI가 그리는 프로그램인 스테이블리티와 미드저니의 제작자등이 예술가3인에 의해 소송을 당했다. (사진=AI가 그린 그림)

[디지털비즈온 이호선 기자] **예술가 3인방**이 **스태빌리티 AI**(Stability AI)와 **미드저니**(Midjourney), AI 예술 생성기 스태빌리티과 미드저니의 **제작자**, 그리고 최근 자체 AI 예술 생성기인 드림업(DreamUp)을 만든 아티스트 포트폴리오 플랫폼 디비언트아트(DeviantArt)를 **상대로 소송을 시작**했다.

이미 AI에 대한 저작권 침해는 예견된 사항으로 판단되고 있다. **사라 앤더슨**(Sarah Andersen), **켈리 맥커넌**(Kelly McKernan), **칼라 오티즈**(Karla Ortiz)와 같은 예술가들은 이러한 조직이 '**원작 예술가의 동의 없이**' 웹에서 스크랩한 50억 이미지에 대해 AI 도구를 훈련함으로써 "**수백만 명의 예술가**"의 권리를 **침해**했

https://www.yna.co.kr/view/AKR20221124068400009

https://www.aitimes.com/news/articleView.html?idxno=149031

https://www.digitalbizon.com/news/articleView.html?idxno=2331217

AI 패널 토의

❑ Open Source AI의 정의

- AI는 전통적인 SW와 달리 Source Code만으로 구성되지 않고, 모델, Weight, 학습 데이터 등 복잡한 구성 요소를 포함
- 전통적인 SW와 달리 AI에서는 투명성과 재현성(Reproducibility)이 보장되지 않음
- OSI는 표준 수립을 위해 Open Source AI Definition (https://opensource.org/ai/open-source-ai-definition)을 발표

❑ 데이터의 라이센스

- 방대한 양의 데이터와 각 데이터 요소에 대한 저작권 문제
- AI 모델이 생성한 합성 데이터(Synthetic Data)를 해결 방안으로 활용하는데, 이 역시 간접적 저작권 침해 위험이 있음

❑ 법적 프레임워크와 규제

- 일본, 이스라엘 등 일부 국가에서는 AI 학습을 위한 데이터 사용이 법적으로 허용됨
- 미국 등 일부 국가에서는 법적 입장이 아직 불확실함
- 특히, 저작권 표기가 데이터에서 분리될 수 없는 부분이 Open Data Licensing에 문제가 되고 있음

AI 패널 토의

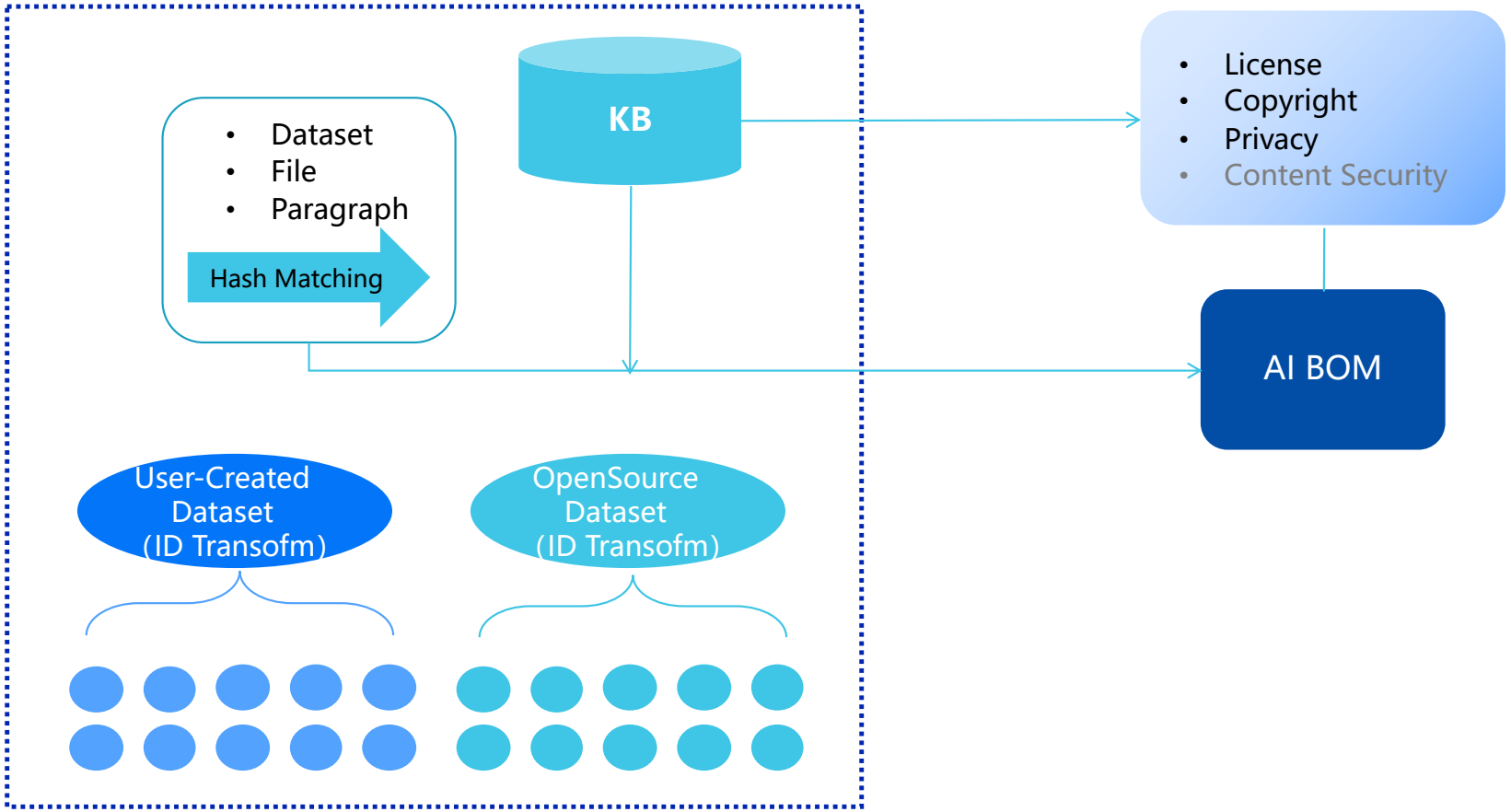
❑ AI 라이센스의 미래

- Source Code 이외의 구성 요소를 포괄하는 라이센스에 대한 논의가 진행 중
- AI 관련 소송이 활발히 진행 중이며, 향후 1년 내 규범적 측면에서 많은 변화가 있을 것
- 웹사이트에서 AI 데이터 사용을 제어하려는 움직임이 확산되면, 법적으로 인정 받을 가능성이 있음 (robots.txt 등)

AI Scanning Tool

AI Scanning Tool

Utilize AI BOM to Perform Open-Source Dataset Identification and Risk Analysis



Open Source 분석 도구

Open Source 분석 도구

- Open Source Identification은 도구의 역량에만 의존할 수 없음
  분석가의 높은 역량(경험)이 필요함
  - Scanner가 검출에 실패하는 경우
  - Scanner의 Database에 잘못된 정보가 있는 경우
  - 여러 Scanner의 정보를 취합하는 과정에서 발생하는 오류
  - 개별 License에 대한 이해 : 결합 구조에 따른 License 적용

- Tool Vendor들의 문제점
  - Open Source 관리를 위한 상용 도구 구매 강요
  - Database 사이즈를 홍보
  - Tool이 모든 문제를 해결하는 것처럼 마케팅
  - 특허받은 알고리즘이 모든 Open Source를 검출할 수 있다는 주장

활용 가능한 Open Source 도구

- Binary Analysis Next Generation (BANG)
  - Firmware, ELF 등 바이너리 분석에 사용되는 도구. AGPL-3.0
- ScanCode
  - 라이센스, 저작권 등의 검출에 사용되는 Source Code 분석 도구. Apache-2.0
- DeviceCode
  - 하드웨어 정보를 바탕으로 유사 장비에 대한 패턴을 파악하여 발생 가능한 버그 및 취약점 정보를 제공하는 도구. Apache-2.0
- Tracing tools
  - 빌드 프로세스를 추적하여 바이너리를 구성하는 소스 및 바이너리의 정보를 제공. Apache-2.0
- VunerableCode / MatchCode / pulrdb
  - purl 별 metadata를 바탕으로 취약점 정보, file/snippet matching 정보를 제공하는 프로젝트. Apache-2.0
- Linux tools
  - vi, grep, cut, sort, readelf 등

Snippet 검출 필요성에 대한 논란

- Q. Open Source 커뮤니티에서는 Snippet 검출이 불필요하다는 주장이 많은데, 대부분의 Tool Vendor들은 Snippet 검출을 필수라는 주장. 원인이 무엇인가?
  - A. 각종 규제 및 고객사의 요청에 따라 Snippet 검출이 필요하다고 한다.

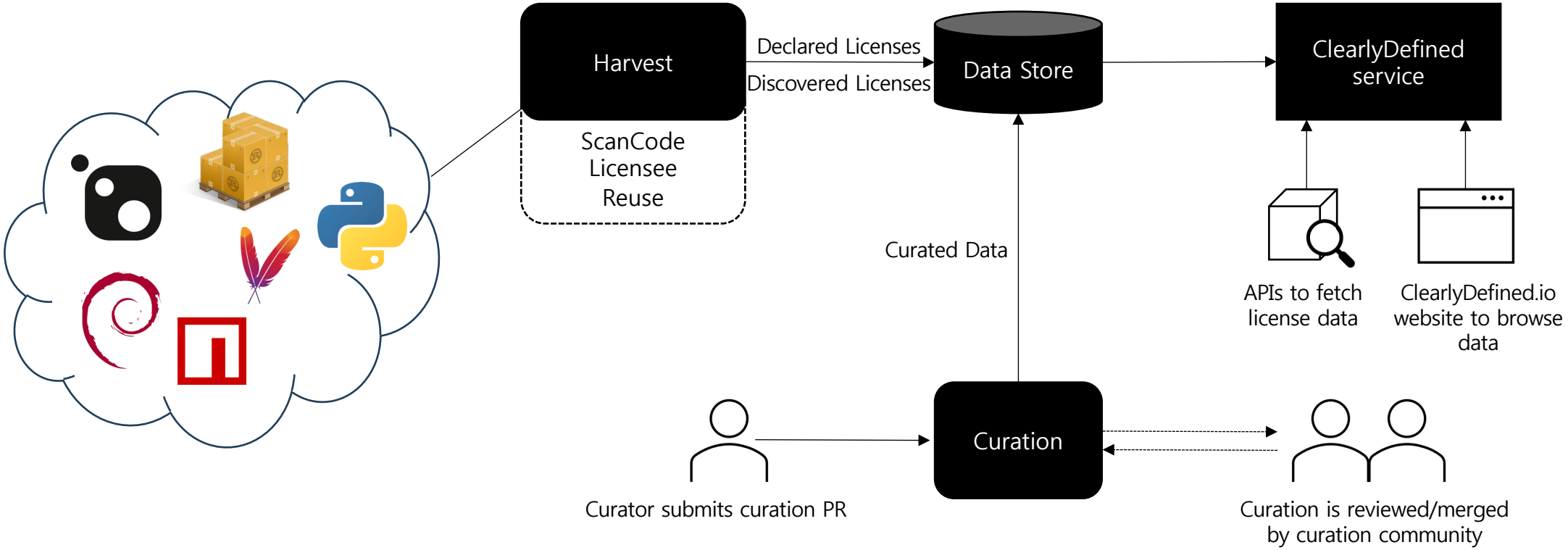
- Q. Snippet이라 함은, 정의에 따라 아주 짧은 Code인데, 이것이 정말 Copyrightable 한 것인가?
  - A. 'Snippet' 의 유형이나 길이 등에 따라 다를 것이다. 결국 준거법에 따라 달라진다.
  - Comment : 큰 기업의 법무팀 입장에서는 Risk 최소화를 위해 Snippet 역시 Copyrightable한 저작물로 판단할 수 밖에 없다.

Open Source Database

ClearlyDefined

- Microsoft가 시작하여 OSI에 기여된 프로젝트

- 1)라이센스 데이터를 수집, 2)커뮤니티가 Curation, 3)openAPI를 통해 활용



http://clearlydefined.io/

Software Heritage

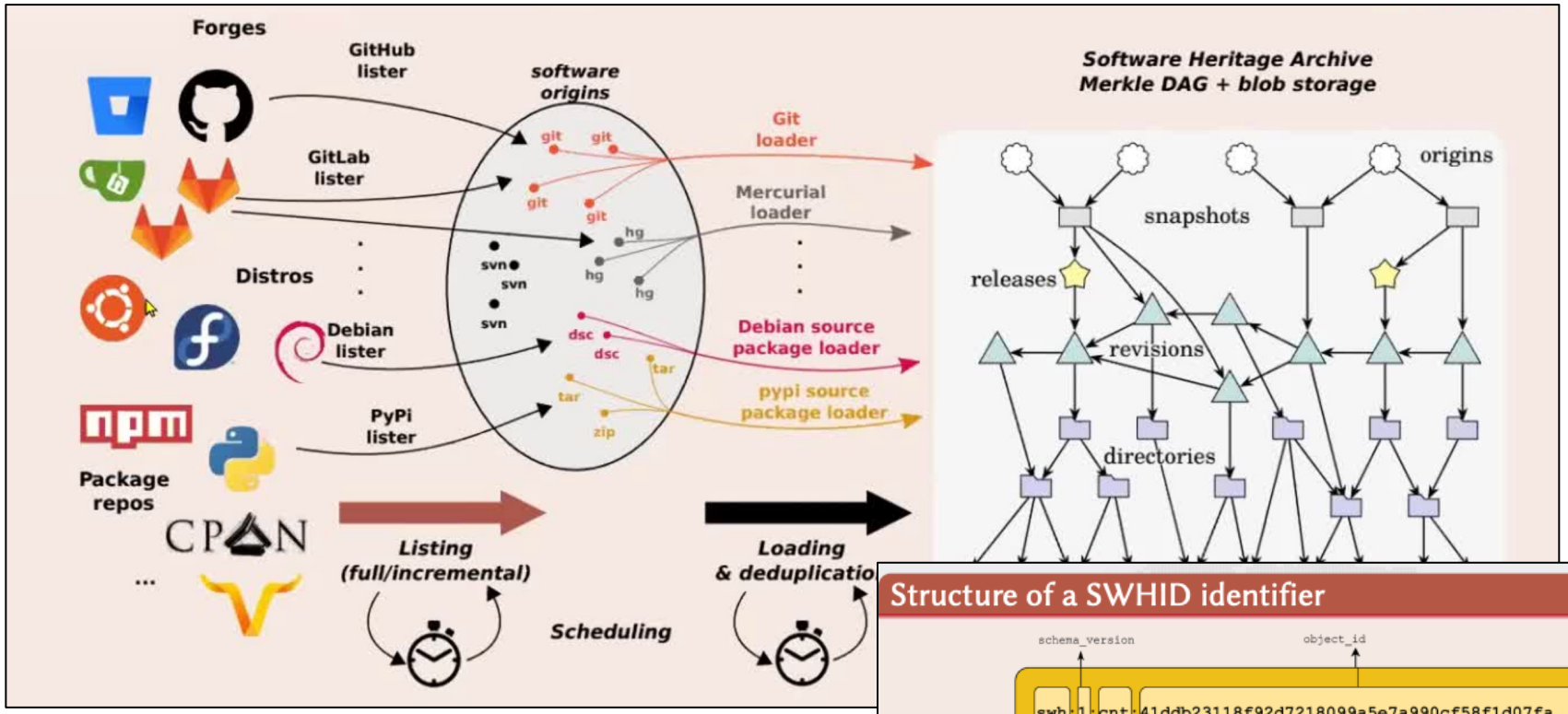
Software Heritage
THE GREAT LIBRARY OF SOURCE CODE

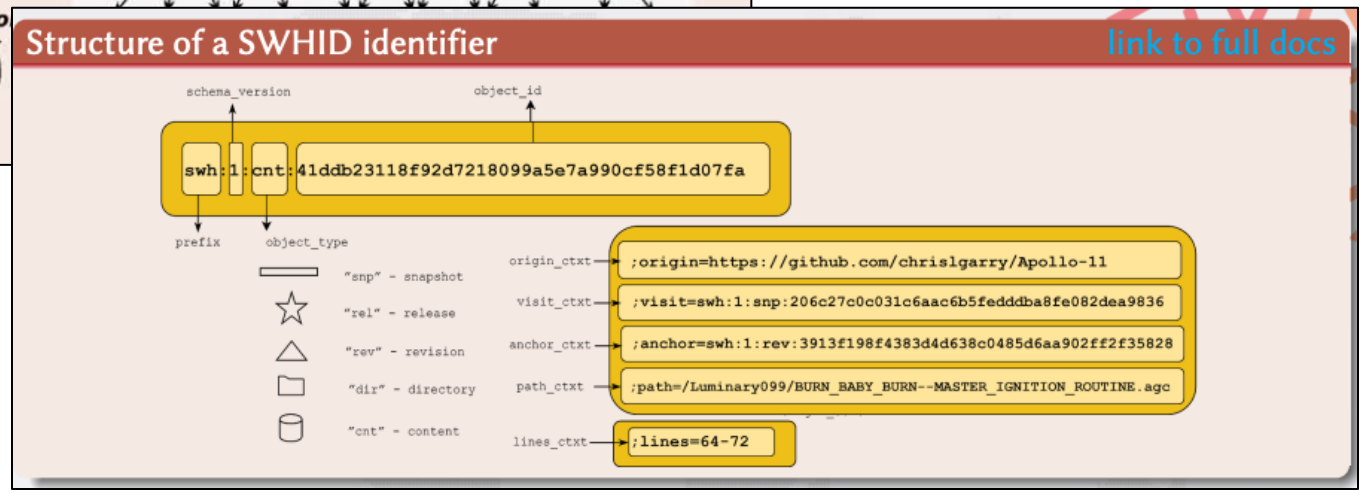
- 2016년 출범
- 미션 : 이 세상 모든 Software의 Source Code를 저장
- 현재 상태
  - 2억개 이상의 프로젝트
  - 200억개 이상의 고유 파일
  - SWHID를 이용해 고유 파일을 찾을 수 있는 API

https://www.softwareheritage.org/

Software Heritage

Software Heritage
THE GREAT LIBRARY OF SOURCE CODE





Slide by R. Di Cosmo / Chauvet Benoît

Software Heritage 사용법

Software Heritage
THE GREAT LIBRARY OF SOURCE CODE

❑ Complete & Corresponding Source Code 제공

- Hoster가 동의하면 법적으로 문제 없음

❑ Cyber Resilience Act (CRA) 준수

- CRA에서 요구하는 장기 가용성 및 무결성 제공

❑ Code Scanning 및 검증

- Pypi 패키지를 이용한 Scanning 가능

❑ 향후 발전 방향

- Hugging Face와 협력하여 Code 생성 LLM 레퍼런스 개발 중

+ @

OSORI

LG전자

kakao

SAMSUNG

한국저작권위원회

감사합니다