Guard Your AI+Software Supply Chain with Sectrend

Wenhui Jin
Cybersecurity Expert, Sectrend

Table of

CONTENTS

Sectrend Stories

About Sectrend

Sectrend is a leading provider of AI + software supply chain governance tools and services.

We specialize in building AI + DevSecOps tool chain and are dedicated to providing world-class application security products.

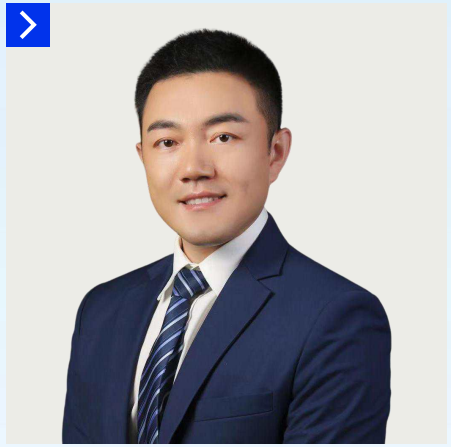
Founded:
June 2021

FTE: 80
(70 SWEs)
R&D team are mostly from Alibaba, Huawei, ZTE, OPPO, etc.

HQ:
Shanghai, China

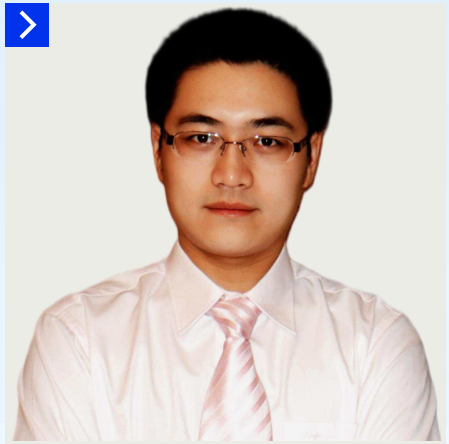
Offices:
Beijing, Shenzhen

Management and Top Experts



Alex Xue | CEO, Founder

- Former Checkmarx Greater China GM
- Former Greater China Business Leader of SIG at Synopsys
- Leader of OpenAtom DevOps Group
- Council Member of KAIYUANSHE
- Draft China DevSecOps standards



King Gao | SVP, Co-founder

- China's Top Open Source Expert
- Former Huawei Open Source Governance Expert
- Extensive experience in open source governance, compliance, and supply chain security across international and domestic initiatives
- Actively engaged in open source communities and contributed to industry standard development

Xianman Zhu | Dir. of Arch.

- Former ZTE Head of Open Source Governance
- Former Lead of Open Source Governance Implementation & Practice at OPPO
- Expert in R&D, Product Development, and Open Source IP Strategy

Safeguard Multi-Dimensional AI + SW Supply Chain

01 CleanSource
SW Composition Analysis

- Generate complete Software Bill of Materials (SBOM)
- Identify known vulnerabilities in 3rd party components
- Manage security and compliance risks in OSS
- Source code and binary (Powered by CleanBinary)

02 PureStream
AI Risk Management

- Generate complete AI Bill of Materials (AI BOM)
- Address privacy, copyright, and content security risks in AIGC
- Promote AIGC compliance and audit
- Enhance AIGC transparency in critical infrastructure sectors

03 CleanCode
Static Analysis

- Identify critical code vulnerabilities and weaknesses
- Standards Compliance (MISRA, AUTOSAR, 21434, etc.)
- OWASP Top 10 & CWE Top 25
- Ensure code quality and security

04 Consulting
Open Source Governance

- Assist in building open source governance teams
- Help establish open source management systems
- Assist in embedding supply chain security and compliance tools into R&D systems
- Goals: Security and Compliance lead to Client Success

Our Story Starts from 2021.06

SECTREND

2021.06

Our dream starts...

2021.12

CleanSource SCA
First GA Release

Largest SCA Investment in China

2021.12

Angel Investment(Millions USD)

2021.12

First Agreement with Tencent

Tencent 腾讯

Solution Delivered to Top Enterprises

2022.06

Second Year of Sectrend Delivered to Large Enterprises

Baidu 百度 · 阿里巴巴 Alibaba.com · BMW · BOSCH · 零束 Z-ONE · HONOR

2022.10

Pre-A Investment(Millions USD)

NIO 蔚来 · 紫光展锐 UNISOC · HUAWEI

2023.04

Pre-A+ Investment(Millions USD)

Comprehensive Penetration in High Tech、Automobile、TeleCom、Internet....

vivo · HIKVISION

The Highest Delivery Standard in China

2024.01

CleanSource SCA V2.0 Released

2024.06

Pre-A++Investment(Millions USD)

2024.10

Awarded “Niche Expertise and Innovation Leaders” in Shanghai

2024.11

Strategical Cooperation with ZTE & OPPO

ZTE 中兴 · oppo · ByteDance 字节跳动

2024.12

PureStream AI Risk Management Platform Officially Announced

2025

Our Dream will Continue with Open Source Stretegy

Contribution to the Global Open Source Community

Continuously Contributing to Global Open Source Community

Linux Foundation

- Member of Linux Foundation
- Member of OpenSSF

OpenChain

- Member of OpenChain and is actively promoting OpenChain adoption in China.
- Participated in the development of the SPDX 3.0 standard and previously incubated projects such as AI BOM maintainer and gopi.

Local Organization

- Lead the security and devops SIG in OpenAtom Foundation
- Help expand and enhance the security and compliance capabilities of openAtom Foundation

Open Source Community

- CHAOSS Maintainer
- Organize the first CHAOSS meeting in China
- Actively involve in openEuler, openHarmony and openAnolis, help them continuously build the capability

Standard Drafting - ISO / China

- ISO 18974
- 《Cybersecurity technology Software supply chain security requirements》GB/T 43698-2024 - The first software supply chain standard in China

Academia

- Collaborate with universities on new tech research and talent acquisition

  Fudan University
  Shanghai Jiaotong University
  East China Normal University

Our Open Source Future Starts from NOW

A — 2024 Build Customer Base

Over the past three years, we have accumulated a majority of the super-large enterprises in China, gaining a deeper understanding of user needs.

B — 2025 1H Open Service

The community edition serves developers worldwide for open-source governance with premium features:

- snippet scan
- binary scan
- vulnerability reachability
- ...

C — 2025 - 2026 Open Code

The community can enhance the overall quality and security of the code, assimilate a greater wealth of creativity and contributions from developers, accelerate the iteration of features and the implementation of new characteristics, thereby bestowing a broader applicability to the supply chain security solutions.

D — 2026 ~ Open Data

Sectrend will operate as a fully open-source company, opening its core database to enable both open-source and commercial tools to build on this foundation. This approach allows the community to contribute to and enhance the database collaboratively, driving innovation for all users.

OSS Security and Compliance Empowered by CleanSource SCA

CleanSource SCA at a Glance

Scanned Project

OSS Introduction Methods Vary

- **Dependency**
- Pkg Manager
- Source Code

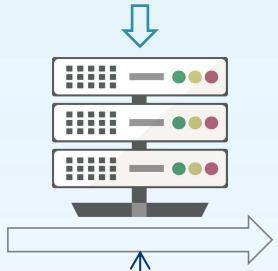
- **Source Code**
- Package
- Partial(File(s))
- Code Snippet

Advanced Technology & Algorithms

Exact Directory Matching
File Matching
Code Snippet Matching

Build Scan
Non-Build Scan

Comprehensive | **Accurate** | **Fast**



Comprehensive and Complete SBOM

SCA KB

- License Risk
- Compatibility Risk
- Tempering Risk
- Cryptography

- Vulnerability
- Reachability Info
- PoC Validation
- Remediation
- Workaround

Compliance Risk Identification & Reporting

License Risk

- Dynamic risk levels based on distribution scenarios, usage and license terms to facilitate timely identification of high-risk OSS

License Compatibility Risk

- Top 110*110 License Matrix
- Risk/no risk/uncertain at a glance, reducing the workload of legal team

License Tempering Risk

- Reduce OSS compliance risk and improve regulatory efficiency by identifying files with tempered license or copyright

Notice File Generation

- Notice file export in different formats to empower open source compliance obligations

SBOM Generation

- Enables traceability and assists in quickly locating vulnerable OSS
- Meet customer or regulatory requirements

Cryptography Identification

- Identify cryptographic algorithms used in open source components to help organizations achieve export control compliance

Vulnerability Identification & Remediation

Up-to-date Vulnerability DB with Comprehensive and Accurate Info Helps Quickly Locate, Fix and Monitor Vulnerability

Vulnerability

- Risk Level
- Affected Entity
- Vulnerability Incident

Remediation

- Clear Function Call Relationship
- Exploitable PoC
- Fix Suggestion(Component/File)
- Patch/Affected Code Snippet
- Vulnerability Reachability

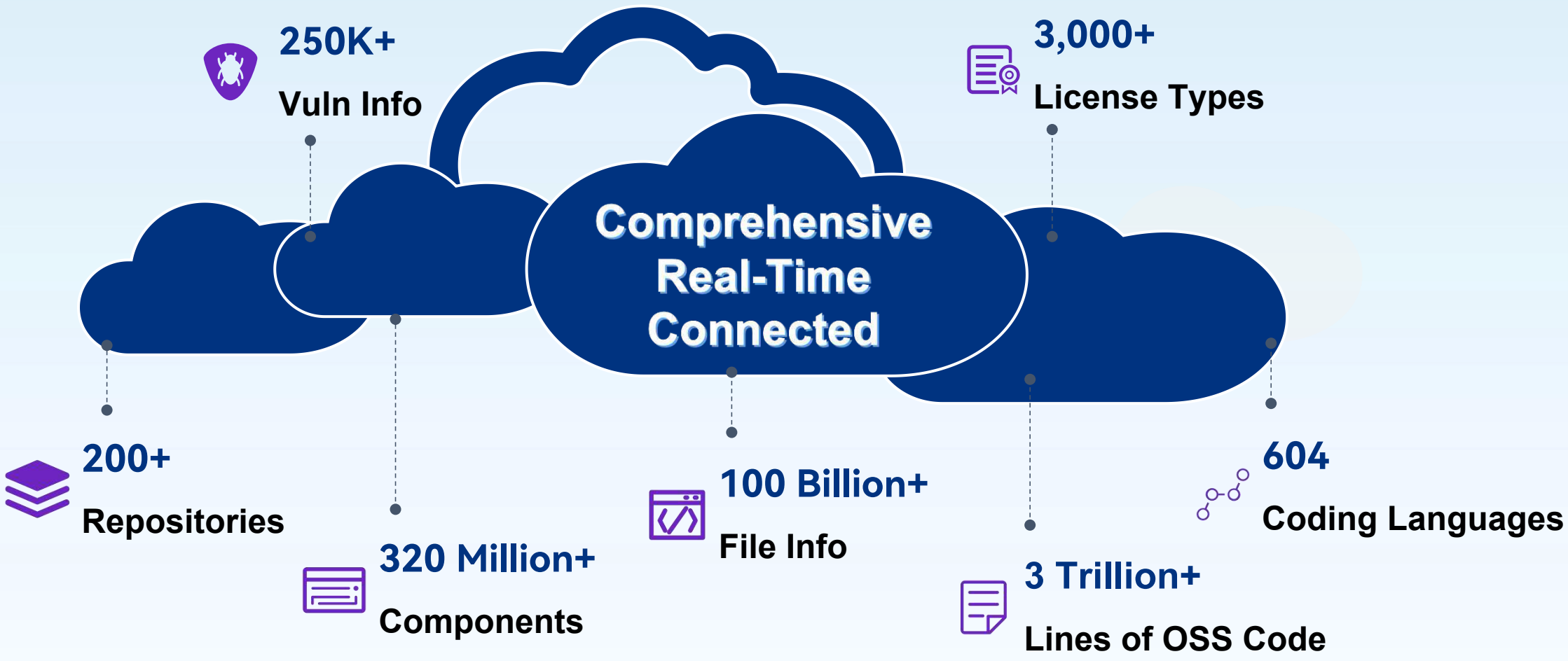
Awareness

- Real Time Updated Databse
- Vulnerability Alerts to Improve Awareness
- Continuous Monitoring Based on SBOM

End-to-End DevOps Integration

...more to come in each release

Comprehensive OSS Component Database



Best Practice – CleanSource SCA Helps Xiaomi TOC with a Eased Software Supply Chain Management

SECTREND

Pain Points

Cross BU Collaboration:
Phone, Auto, IoT, Cloud

OSS Introduced via Different Ways

Lack of Guidance on the Best Practices

Super Large Files with OS Failed or Speed Limited

Code Exposure & Cross Border Data Exchange

Soultions

Deep integration into development platform with API and DevOps plugins

All forms of OSS identification including component, code snippet, dependency, etc

Performed project decomposition with streaming scan to reduce resource consumption and balance scan efficiency(400GB)

CleanSource SCA(On-Prem)

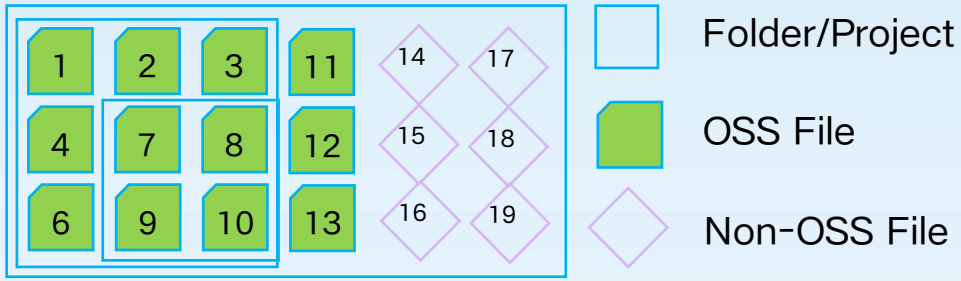
Expert consulting services on open source governance to ensure efficient deployment of tools

Achievements

- Protected Software Supply Chain from the Bottom Up
- Improved OSS Security and Compliance
- Reduced Intellectual Property Risk
- Integrated into Development Process
- Enhanced Brand Image

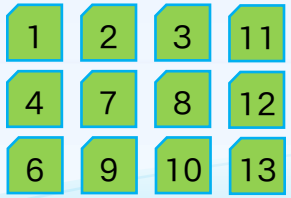
Tech Highlight 1: Exact Directory

Input



Before

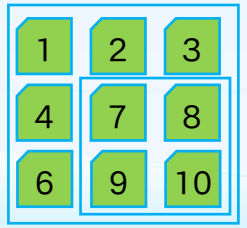
File Match



Code Snippet Match

After

Exact Directory



File

Snippet

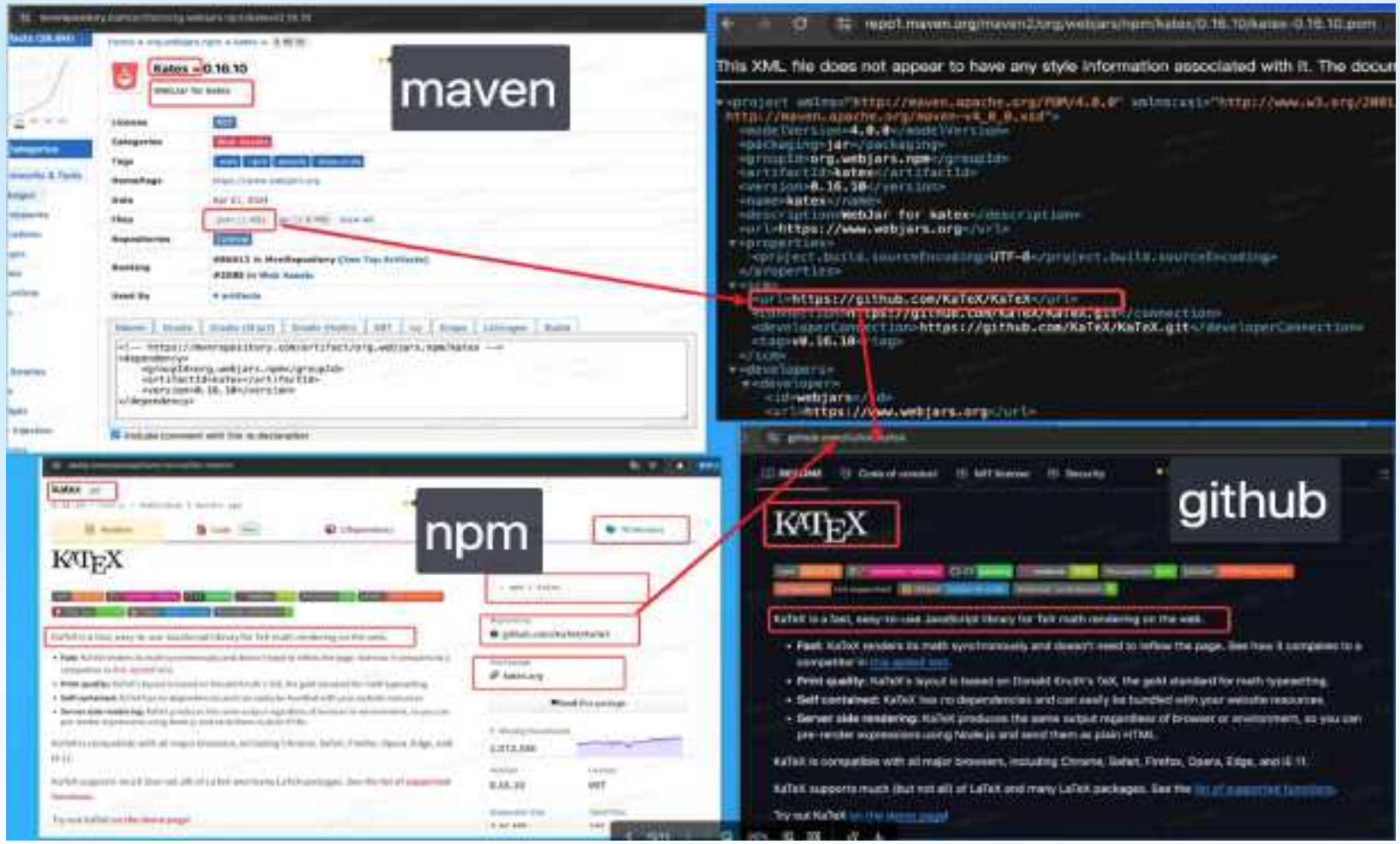
With exact directory scan, we can get:

- More converged and accurate results
- Inter-components lineage relationship
- 30%+ speed improvement

Tech Highlight 2: Component Identification System

WHAT

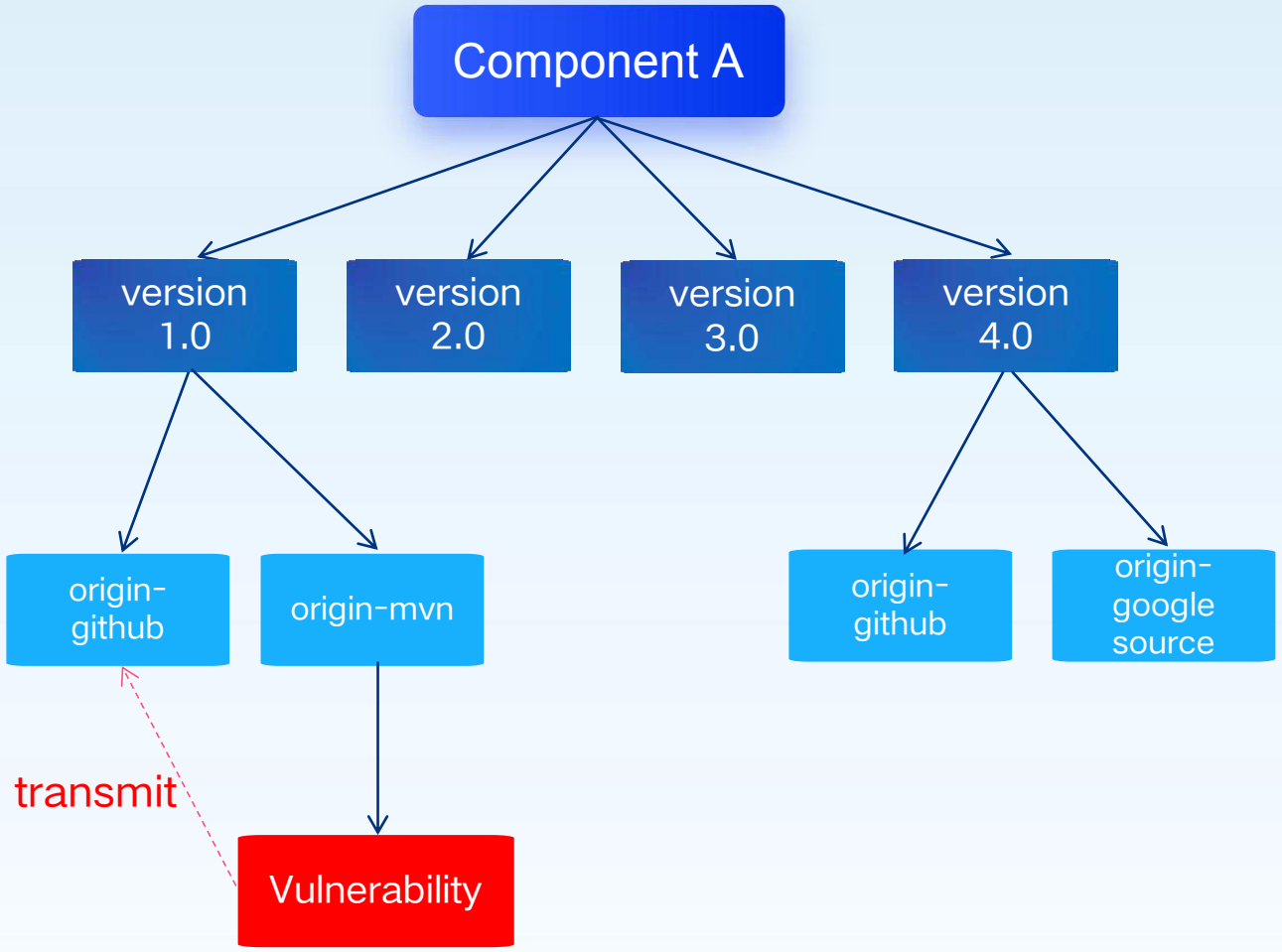
- Leveraging AI capabilities to merge different distributions of the same component into a unified CID.
- This process involves building a knowledge graph to perform knowledge inference on attributes such as licenses, vulnerabilities, and EOL.



VALUE

- Accuracy
- Fast
- Efficiency
- Real Time

Component Identification System



CID

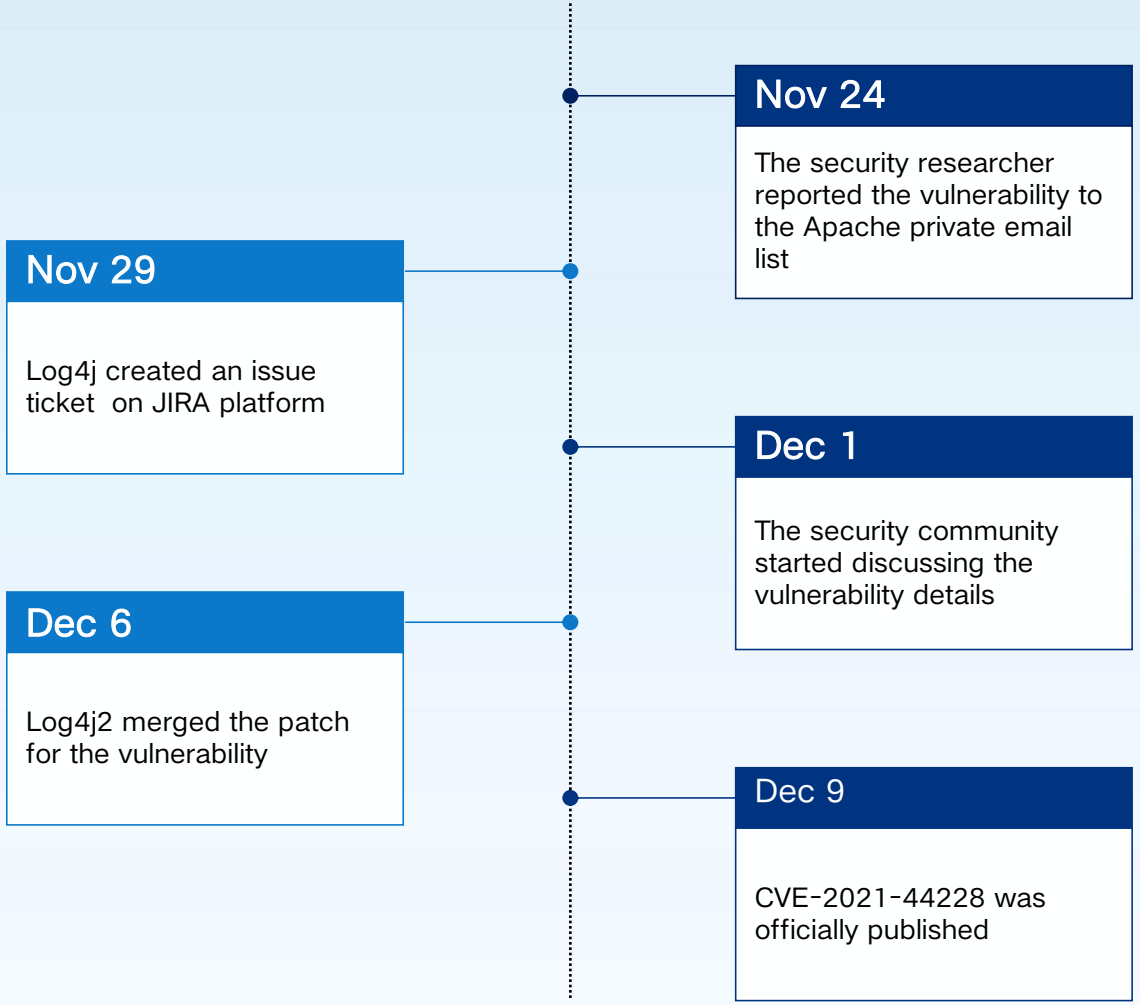
VID

OID

Tech Highlight 3: 0-day Vulnerability Intelligence

Identify potential vulnerabilities before disclosure

- Some open-source vulnerabilities may not be assigned a CVE number
- Relying solely on CVEs or community security advisories lacks timeliness; by monitoring publicly available information in open-source repositories, potential CVEs can be identified in advance
- Researchers or developers may inadvertently disclose vulnerability details in open-source communities, such as GitHub Issues



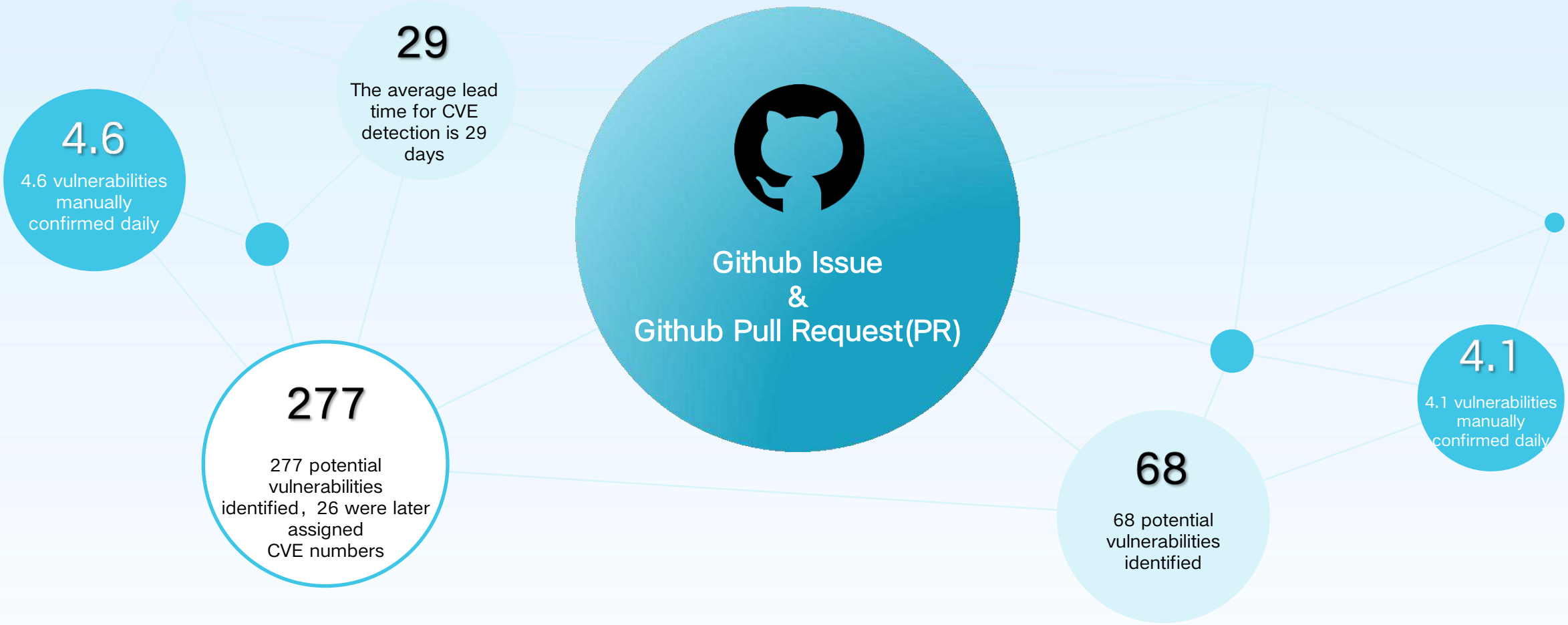
Tech Highlight 3: 0-day Vulnerability Intelligence

Github Issue

Monitored GitHub issues from Nov 2024 - Jan 2025

Github Pull Request(PR)

Monitored GitHub PRs from Feb 2024 - Now

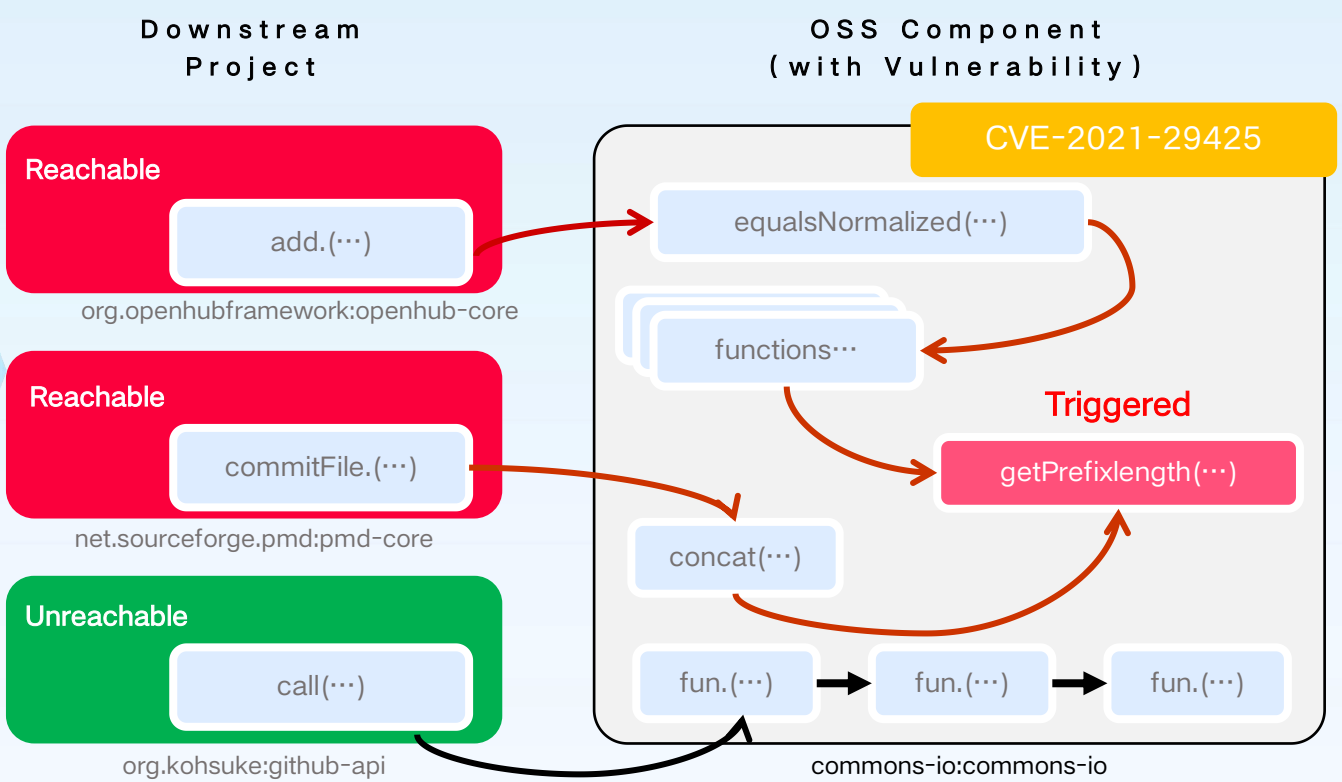


Tech Highlight 4: Vulnerability Reachability

Enhanced Vulnerability Analysis, Accuracy & Fix Efficiency

Function-level vulnerability reachability is of critical importance.

It not only helps reduce security risks in the software supply chain but also enhances the speed and efficiency of responding to and addressing security vulnerabilities.



Dashboard
Tasks
Elements
Policies
Wiki
Report
System
Manage

xstream-CVE-2021-29505 / xstream-CVE-2021-29505 / **defaultVersion**

Distribution: External Stage: Developing Project Licenses 1

Unconfirmed...	License	Total OSS Files	Cryptography	Vulnerability	Alert	License Risk	Self-researc...
0	2	0	0	32	0	0	100.00%

Component | Source | Vulnerability | Policy | Sensitive Information | Report | History

Please Inp... | Please Input C...

Please Select Risk Level

commons-collections 3.1
Comp... https://mvnrepository.com/...
Vulner... 1 1

xstream 1.4.15
Comp... https://mvnrepository.com/...
Vulner... 7 22 1

xstream 1.4.15

30 Known Vulnerability Remediation Suggestion: Upgrade to version 1.4.21

Risk Level: Please Select ... | CVE: Please Input... | Vulnerability Reachability: All | CWE: Please Input...

Fix Status: Please Select ...

Set Fix Status in Batch | 1 Items Selected

Vulnerability	Level	Vulnera...	CVE	CNNVD	Release Date	Operate
XStream 代码问题漏洞	High	Reachable	CVE-2021-21349	CNNVD-202103-1236	2021-03-23 08:15	Fix Status
XStream 资源管理错误漏洞	High	Reachable	CVE-2021-21348	CNNVD-202103-1238	2021-03-23 08:15	Fix Status
XStream 缓冲区错误漏洞	High	Unreachable	CVE-2022-40151	CNNVD-202209-1234	2022-09-16 18:15	Fix Status
XStream 代码问题漏洞	High	Reachable	CVE-2021-29505	CNNVD-202105-1981	2021-05-29 05:15	Fix Status
XStream 安全漏洞	Medium	Unreachable	CVE-2021-39140	CNNVD-202108-1900	2021-08-24 03:15	Fix Status

2 items 1 / 1

30 items 1 50 / page

Dashboard
Tasks
Elements
Policies
Wiki
Report
System
Manage

High **XStream 代码问题漏洞**

CVE-2021-29505 CNNVD-202105-1981 Release Date: 2021-05-29 05:15 Update At: 2023-11-07 11:32

File Path src/main/java/Main.java

src.../Main.java

Line 66

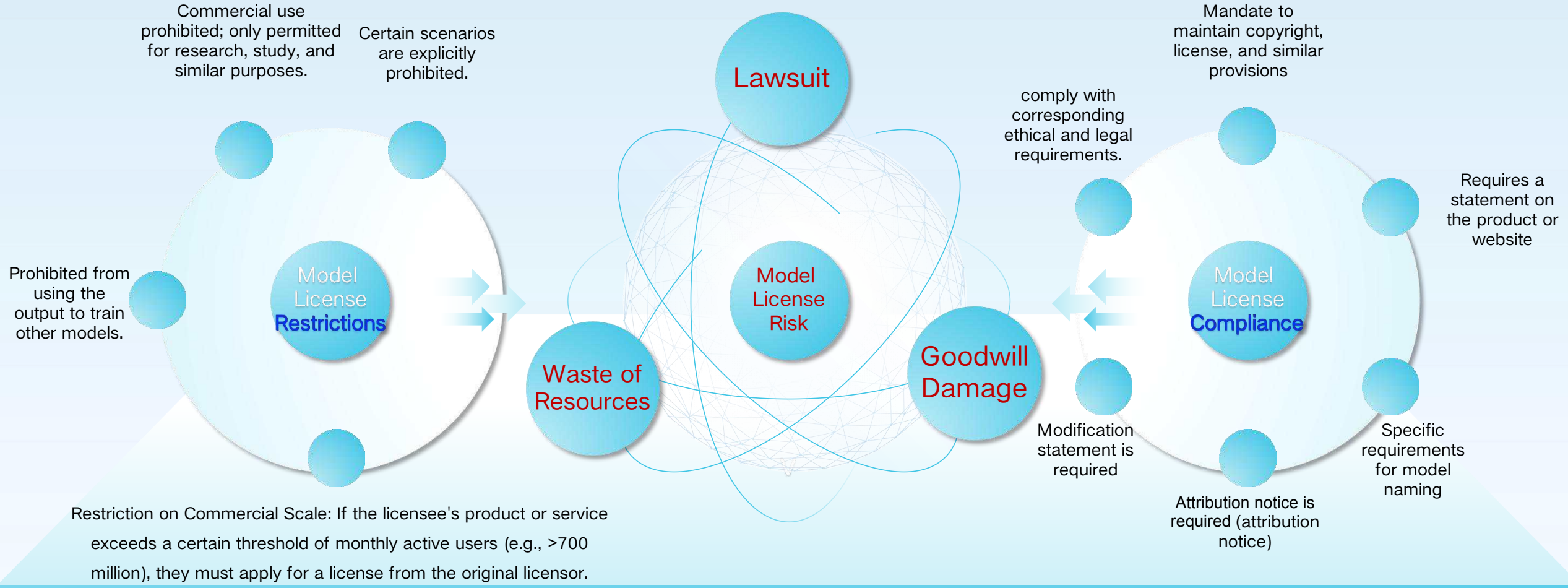
Local Source Code

```
50              "                                        <short>0</short>  \n" +
51              "                                        <boolean>false</boolean>  \n" +
52              "                                    </java.rmi.server.RemoteObject>  \n" +
53              "                                </registry>  \n" +
54              "                                <host>127.0.0.1</host>  \n" +
55              "                                <port>1099</port>  \n" +
56              "                            </ctx>  \n" +
57              "                        </candidates>  \n" +
58              "                    </aliases>  \n" +
59              "                </nullIter>  \n" +
60              "            </sm>  \n" +
61              "        </message>  \n" +
62              "    </value>  \n" +
63              "</javax.naming.ldap.Rdn_-RdnEntry>  \n" +
64              "</java.util.PriorityQueue>  \n" +
65              "</java.util.PriorityQueue> ";
66          xStream.fromXML(xml);
67          //CommonsBeanutils1 cb = new CommonsBeanutils1();
68          //String xml = xStream.toXML(cb.getObject("cmd.exe /c echo lll>99999999.txt"));
69          System.out.println(xml);
70      }
71  }
```

04

AI is Reshaping SCA 2.0

AIGC is Reshaping the World and the Future SCA



Data Leak

Personal privacy, trade secrets, and similar matters.

Copyright Violation

Some training data may not have been legally authorized and could potentially involve copyright-protected data.

Content Security

Resulting in the product being taken down or facing regulatory penalties.

Product(PureStream) Main Features

PureStream clearly lists dataset elements and their sources, helping trace origins, mitigate legal risks, and ensure compliance. It details dataset names, risky content ratios, licenses, copyrights, and sensitive personal information.

PureStream enables precise traceability of models, displaying a list of local model names along with corresponding matched model names, types, official distribution sources, code repositories, licenses, official links, methods, use cases, model architectures, base models, datasets, applicable scopes, Commit IDs, release dates, and more.

PureStream enables model searches by source and type, helping users understand model features, performance, and usage details. Its knowledge base displays model names, types, sources, licenses, architectures, datasets, scopes, and revision records (Commit ID & release date).

Multi-dimensional corpus comprehensive detection

Model risk detection

Model Knowledge Base Selection

Technology Advantages

LLM Recognition ModelCard

Using pre-trained large models' unsupervised learning or named entity recognition (NER) capabilities, key entities and relationships in ModelCards (e.g., model-dataset usage or version-performance metrics) are identified and integrated into a complete, visually structured format for user-friendly understanding.

Multi-Source Dataset Collection

PureStream collects datasets from diverse sources, supporting around 20 mainstream datasets like Wikipedia, Zhihu, and CC100. It detects AIGC-generated content, assessing risks such as copyright issues, privacy concerns, and content safety.

Model Risk Identification

PureStream supports model detection traceability and assists in model selection. It currently includes 5 model sources and enables retrieval of over 1,000 model types. By constructing a model derivative relationship graph, it establishes communities for high-imitation models and untrusted models, analyzes model licenses and copyright risks, and traces official links, code repositories, and download addresses of the used models.

Dimensionality Reduction and Compression

To protect user data, before uploading a model for detection, users can use PureStream's in-house dimensionality reduction and compression engine to locally process the target into an irreversible encrypted fingerprint. The model host remains offline, and the platform never accesses the source code, ensuring data security.

Complex Multi-Scenario Model Detection

PureStream uses feature extraction to analyze fused layers, determining original feature dimensions and representations. It leverages feature space similarity, calculating distances or mapping features to align detection techniques with fused or segmented models. By comparing statistical characteristics of original and modified model parameters, it identifies optimal detection strategies.

Super Large Scale KnowledgeBase(PB-Scale)

Data Set

CulturaX	The Stack	CC100	mC4
Cirrussearch wiki	peS2o	SkyPile	The Pile
Wikipedia	Common crawl	Baidu Baike	Sougou Baike
Zhihu	Quora	wanjuan	C4
Wudao	OLMO	project gutenberg	openwebtext

Data Label

Dataset

Name, License, Confidentiality Level, Sensitive Information, Data Type, Purpose, Language, Size, etc.

File

Name, URL, Language, Copyright Information, Sensitive Information Label, Compressed File Location, Positioning, etc.

Paragraph

Paragraph Positioning, Sensitive Information Label

Data Format

- parquet
- jsonl
- zip
- rar
- zstd
- zst

Thank you!
Feel free to contact us...

King Gao kun.gao@sectrend.com.cn

Feng Wang (Director of Marketing) feng.wang@sectrend.com.cn

SECTREND

Beyond Security, more than open source.

Thank you !