CERIAS

The Center for Education and Research in Information Assurance and Security

Secure Chain: A Knowledge Graph for Resilient, Trustworthy, and Secure Software Supply Chains

Yifeng Di*, **Hadi Askari†**, **Shushan Arakelyan‡**, **Xiangyu Zhang***, **Xiang Ren‡**, **Muhao Chen†**, **Tianyi Zhang***

**Purdue University*

†University of California, Davis

‡University of Southern California

Introduction

Software is now integral to critical U.S. infrastructures, with software supply chains supporting rapid development but also increasing risks. Bugs, vulnerabilities, or unauthorized changes in upstream components can propagate downstream, posing significant threats.

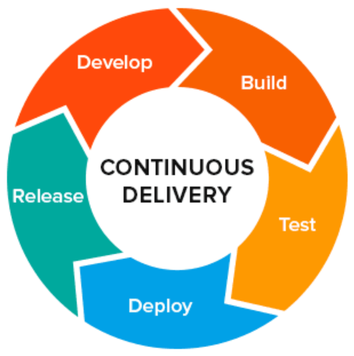
We propose a comprehensive knowledge graph that models the relationships between software, hardware, vulnerabilities, and other entities in software supply chains. It captures rich, up-to-date information about software components in heterogeneous software ecosystems to support secure and transparent management of software supply chains.

Motivation

Exploit vulnerabilities in 3rd party components to affect downstream applications and users

Third-party Software Components

Download & Integrate



In-house Software Development

Data Sources

Software
- All libraries from Conan, an open-source C and C++ package manager
- All Packages in the official Debian distribution
- Top 1000 C/C++ GitHub repositories

Hardware
- Common Platform Enumeration (CPE)

Vulnerability
- Common Vulnerabilities and Exposures (CVE)

Vulnerability Type
- Common Weakness Enumeration (CWE)

Knowledge Collection

Extract 8026 Conan library versions and their dependencies

Extract 475847 Debian package versions and their dependencies

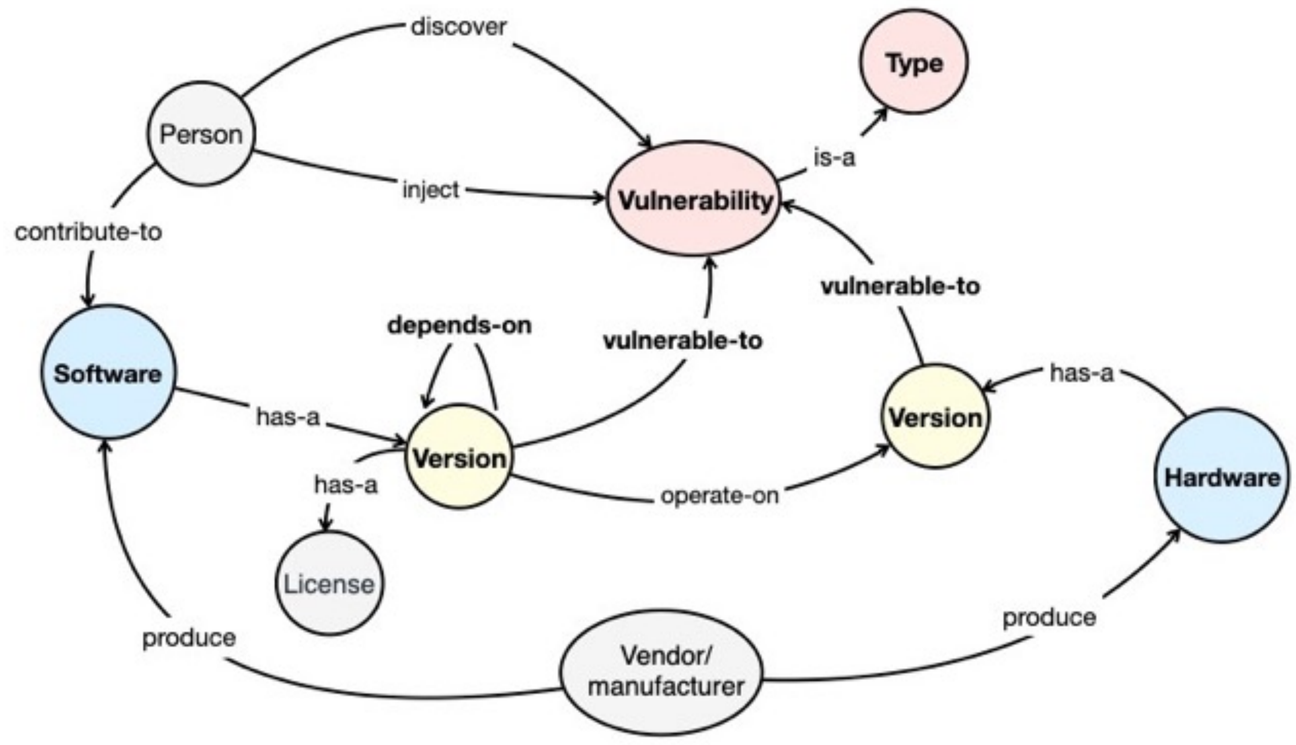
Extract 259334 CVEs and link them to software & hardware

Extract 1446 CWEs and link them to CVEs

Extract 44772 version releases of the top 1000 repos and use CCScanner to extract their dependencies

Extract 58610 hardware versions from CPE Dictionary

Ontology



KG Statistics

- 11168 software and 91467 versions
- 54369 hardware and 58610 versions
- 1653393 dependency edges
- 259334 vulnerabilities in 1446 types
- 473318 vulnerability edges
- 30898 developers
- 22002 vendors and manufacturers
- 21 licenses

Use Cases

IT managers and security analysts can query the KG to check for dependencies on vulnerable 3rd-party software.

End-users can check if software on their computers is affected by recent vulnerabilities.

Developers can check library security and use the KG to generate a detailed Software Bill of Materials.