

Homework 1 Report

Robert Wallace – rwallac1@nd.edu

Design Decisions

I use the NVD REST API to extract information from NVD, using the requests library. I did not end up using an API key, though I do have one, because I never got rate limited and so did not notice that I wasn't using it in the first place. I added it in the end to make sure that the TA can easily add the API key

To parse the pom files, I used the lxml library to parse the pom files and navigated and to find the dependencies after removing comments.

To match CVE information to dependencies in a pom file, I select all configurations and relevant columns from each row where the configurations is LIKE the artifactID and either the descriptions or the reference is LIKE the groupId's vendor property. Under this, this should find the vulnerabilities that reference the dependency in question, though it may not cover vulnerabilities in that dependency, since those will not mention the groupId, much less be guaranteed to have a similar artifactId

My database is simply a 263,161-row table containing every CVE from 1999 to 2024, last updated September 17, 2024, 7:42 PM, with columns as shown in Figure 1

Knowledge Base Statistics

For example: "The database includes the tables shown in Figure 1. It contains 1 table, containing all the data that I thought relevant when first importing the data. Further revisions would clean this up and remove unused data, though it may be useful for automatically updating or similar purposes.

Vulnerabilities	
cve_id	TEXT
descriptions	TEXT
source_identifier	TEXT
published	TEXT
lastModified	TEXT
weaknesses	TEXT
configurations	TEXT
reference	TEXT
metrics	TEXT

Figure 1

Research Synthesis

- 1)
- 2)