## Root Cause Mapping (RCM) Working Group Meeting Notes

**8/14/2024 at 1pm**

## Agenda

* Iterative discussion on the Chris Madden (Yahoo) Presentation; discuss better mapping techniques.
* Open discussion on current issues and questions from the Community on root cause mapping and what they’ve been dealing with.
* Update on Top 25 – engaging Community partners in helping venerate the best possible term analysis.

## Action Items

* The group plans to engage the CNA community to identify and correct potentially incorrect or insufficient CVE mappings, utilizing keyword matchers and searches.
* Emails will be sent to each CNA with text files containing information to help produce new mappings, drawing from a dataset of approximately 31,000 CVEs.

## Meeting Summary

* **Meeting Overview**: The Root Cause Mapping (RCM) Working Group convened on August 14, 2024, at 1 PM to discuss various topics, including iterative discussion on Chris Madden's presentation from Yahoo, current issues in root cause mapping, and updates on the Top 25 term analysis.
* **Key Focus Areas**: The group aims to identify and incorporate the best capabilities for high-quality mapping into vulnerability research practices.
* **Chris Madden's Presentation**: Chris Madden from Yahoo presented his analysis on using grounded LLM technology to enhance mapping quality, including the use of Google’s Gemini system for document management and note-taking.
* **Prototyping Examples**: Prototyping examples discussed included mapping CWE to CVE, refining CWE specifications, and adding API and browser access for better mapping.
* **Community Involvement**: The group plans to engage the CNA community to identify and correct potentially incorrect or insufficient CVE mappings, utilizing keyword matchers and searches.
* **Action Plan**: Emails will be sent to each CNA with text files containing information to help produce new mappings, drawing from a dataset of approximately 31,000 CVEs.

## Meeting Notes

**Key Focus Areas for this Group:**

1. Try and identify in one or several best capabilities around better-quality mapping and figure out a way to incorporate them in the practices of doing many vulnerability research.

**Chris Madden (Yahoo) Presentation**

* Chris has been doing some analysis in using grounded LLM technology to support better quality mapping.
* NotebookLLM – Google’s Gemini (Google AI) System, which allows document imports and lets you takes notes, ask questions, organize your ideas and more.
* Chris has been using LLM technology for various use cases.
  + Curiosity: to see if they could be used to map CWE to CVE or vulnerability descriptions in general.
* Prototyping examples – CWE Map: reviewing initial mockups, trim CWE specifications to mold next versions, get to know CVE/CWE mappings from 2023 Top 25 process, add API and Browser Access.
* CNAs on the call offered their suggestions and additional examples regarding CVE description layout.
* There were some fundamental problems discussed about how people can try and force things into a lower level of abstraction when it shouldn’t be when it comes to CVE mapping.

**CWE Top 25 – Brief summary on this year’s approach**

* Plans to leverage the CNA community by utilizing some of the keyword matcher and searches internally to help identify potentially incorrect or insufficient CVE mappings.
* This allows the matching data to be turned over to the CNAs who have extra information to dive deeper into these mappings and get some help in determining the proper mappings and use case.
* Emails will be sent out to each CNA individually with text files that would have that information to help produce new mappings.
* Approximately drawing from 31,000 CVE data set with over 200 CNAs to send and split.