Exercise - Open Source SW Engineer

⚙️ Description

Create an LLM-driven tool, that looks for code vulnerabilities in C/C++ code files.

📦 Deliverables

All the following deliverables should come in an archived local git repository.

CLI Tool

A CLI tool that gets a C/C++ file, and output the potential vulnerabilities and security flaws in it.

Below is an example for a possible output (you can choose whatever output format you see fit):

```
# analyzer vulnerable_code.c

Line 20: Possible UAF due to...

Line 30: Possible stack based buffer overflow due to...

...
```

Readme.md

Add a `Readme.md` file explaining:

- The project architecture and design
- A “Quick Start” section explaining how to use the tool

- Other things you see fit

Report.md

Add a detailed report of your work process, containing your thought process and decision making in solving this exercise.

◼ Guidlines

- You can write this tool in any language you see fit
- Use Microsoft's `phi4` model as the LLM that looks for the vulnerabilities (in case you need a smaller model, you can also use `gemma3:1b`)
- The tool should be working locally, and without any internet connection
- Use git commits to show the progress of the tool's development
- As part of the exercise, you got two C/C++ files that contain several vulnerabilities. You can use those files to test the tool.

➕ Bonus

- Docker - Create a Docker image (using a Dockerfile) with the tool pre-installed and configured for easy-to-use access
- Fix - Include a possible fix for the vulnerability

💬 Disclaimer

The exercise is defined in high-level terms intentionally, there are many different ways to solve it, and we wish to see how do you approach this kind of problems without any restrictions.