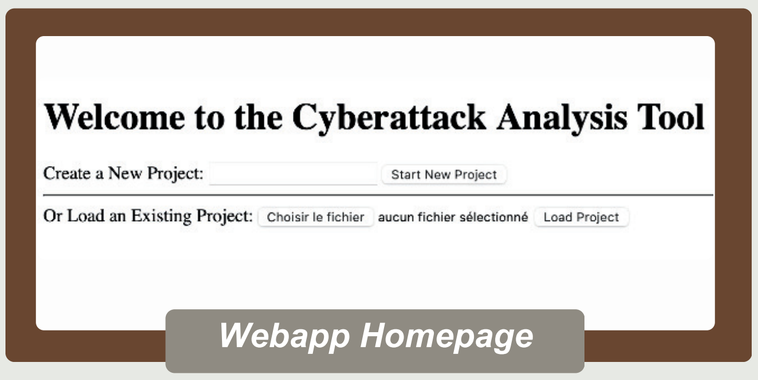
Can AI Analyse Cyberattacks Better Than You Do?

Exploring the Use of LLMs (Large Language Models) in Cyberattack Analysis

Boris Christov, EPFL - The CyberPeace Institute

1 Motivation

- Performing **non-trivial analysis on text** takes **time and effort for humans**
- With the advancements in the field of LLMs, **we can now conceive solutions** with text **understanding and reasoning abilities**
- ***"Automate to Scale"***



Webapp Homepage

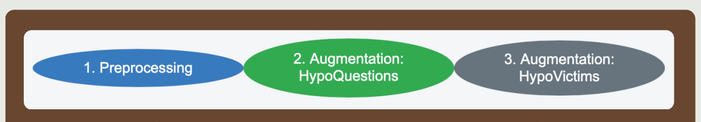
2 Requirements

- **not** a **categorisation program:** an analysis
- **reliable**: quality consistently meets expectations
- **thorough**: minimising bias, considering all the input information
- **understandable**: the steps taken to arrive at the conlusion need to be clear

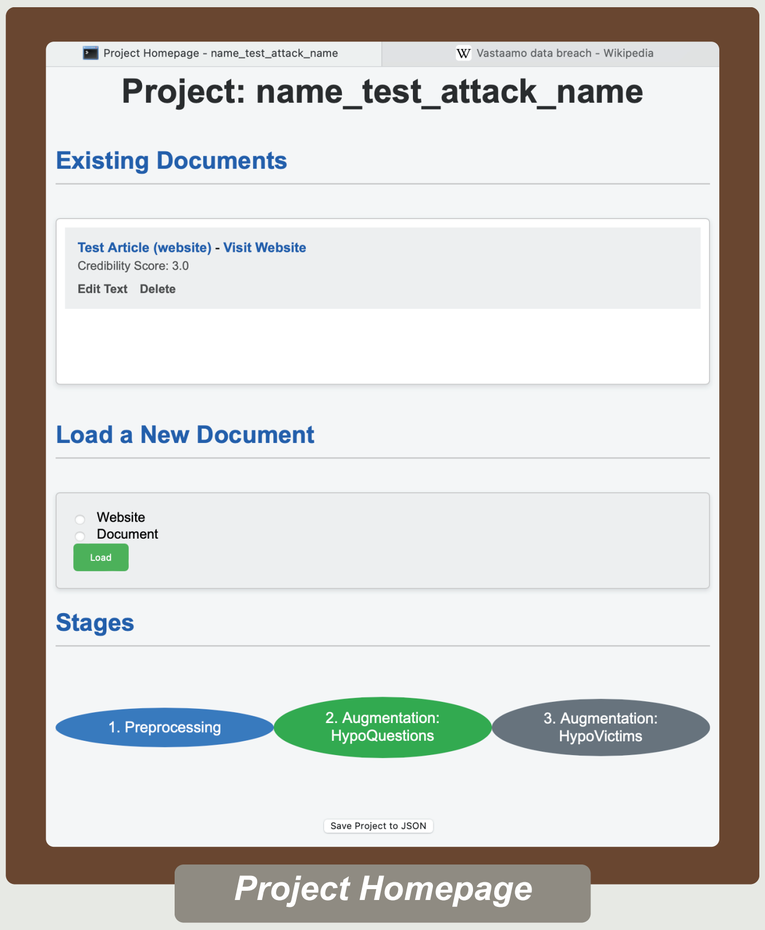
A typical single LLM task looks like this

3 Challenges

- **reliability:** hallucinations, sources of different styles => in different responses;
- **thoroughness:** not all the information is considered, analysis bias, complexity of instruction undermines result quality;
- **understandability:** a balance between automatisation and modularisation in a user-friendly platform must be found



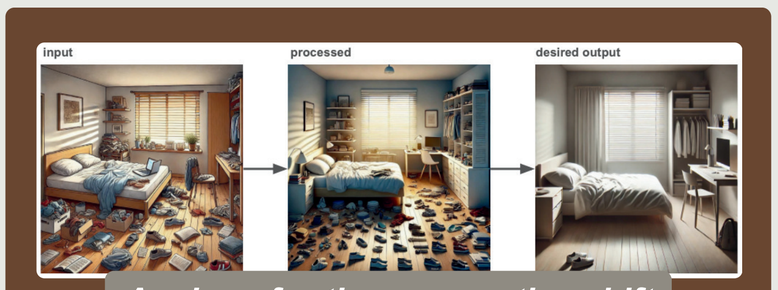
The "Stages" Section



Project Homepage

4 Perspective Shift

- focus on **creating the environment**, not on refining a model
- **divide** the **text** into pieces of **information**
- allow for a **person-in-the loop**



Analogy for the perspective shift

5 Modularity

In the form of "Stages." Implications:

- code **updatable** => **program relevant**
- analyst can **validate** program output => **program reliable**
- programmers can **create different processing pipelines** and/or use usage data to fine-tune models => **program upgradeable**

6 Next Steps

- **Finish the analysis pipeline**
  - Create one working configuration
    - Analyse its approach and optimise it
  - Brainstorm new approaches (different "Body of Knowledge," different "atomic pieces of information," etc.)

7 Conclusion

- "**Better**" is **subjective**--how much do we value speed, quality, work capacity, consistency and with what proportions?
- To create **solution** that is both reliable and thorough is **not trivial**, and one way to do that is to concentrate on creating the infrastructure around the LLMs. More specifically, creating a "macro data processing pipeline" might be **key to solving this problem**.

Contacts

Boris Christov

8 Aknowledgments

- I sincerely thank the CyberPeace Institute, and specifically Ms. Charlotte Lindsey and Mr. Gwyn Glasser for the ongoing trust, support and encouragement.
- I am equally thankful to the Laidlaw Foundation and EPFL for empowering the youth.