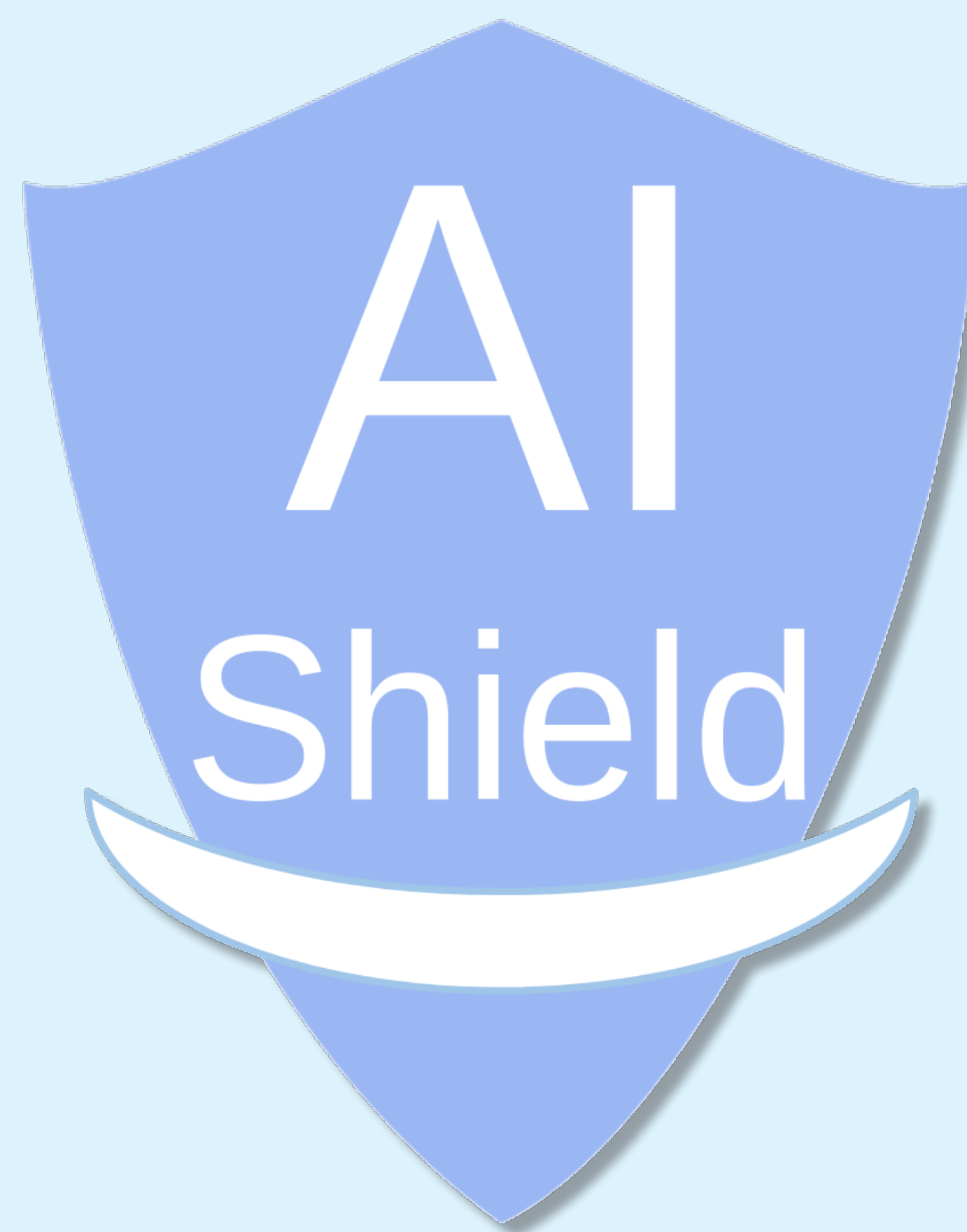


AI Shield

An interdisciplinary research project that is being developed by the Universities of Twente and Groningen and the Hanze University of Applied Sciences

Technical goal

To design a GenAI system based on a *foundational model* capable of automatically analysing the threat landscape and identifying potential attack paths and patterns.



Ethical and legal goals

To ensure *full compliance* of the system to be developed with ethical and legal standards that have been laid down in the applicable ethical and legal frameworks.

Technical aspects of the project

- » The model will be trained on *attack knowledge* from different sources, such as MITRE ATT&CK, CVEs, CWEs, CTIs, CAPEC, threat intelligence blogs and articles;
- » It will significantly improve *attack-path analysis* by leveraging GenAI;
- » Both *the training* and *the application* of LLMs for cybersecurity purposes will be evaluated and improved;
- » In addition, possible *mitigation techniques* will be determined;
- » The output of the model will be *directly translated* to the IT environments of users.

Project's ethical and legal aspects

The project adopts a *human-centric approach* respecting European ethical values and principles and ensures human oversight, technical robustness and safety. It is structured with careful attention to the adherence to such *legal frameworks* as:

- » The General Data Protection Regulation (measures to protect personal information);
- » The Cybersecurity Act (cybersecurity certification schemes);
- » The NIS2 Directive (obligations of entities in critical sectors);
- » The Cyber Resilience Act (products with digital elements);
- » The AI Act (development and use of artificial intelligence solutions).

