MINERVA

KI für Korrelations-Plattform mehrfacher Bedrohungsinformations-Quellen (Threat-Data-Feeds)

Programm / Ausschreibung	KIRAS, F&E-Dienstleistungen, KIRAS-K-Pass-KMU Innovation AKUT KIA F&E Dienstleistungen (FED KIA_2023)	**Status**	abgeschlossen
Projektstart	02.04.2024	**Projektende**	31.01.2025
Zeitraum	2024 - 2025	**Projektlaufzeit**	10 Monate
Keywords	KI, Cybersicherheit, Security, Pretrainer; Generative KI		

Projektbeschreibung

Die Motivation für die Entwicklung dieses KI-basierten Ansatzes liegt in der Erweiterung des bestehenden Cybersicherheitsprodukts von IKARUS Security Software (IKARUS TIP – Threat Information Platform), um die Nutzung und Bedienbarkeit des Produktes zu vereinfachen und es Dank der implementierten KI in die Lage zu versetzen, riesige Bedrohungs-Informationsmengen zu korrelieren und zu verarbeiten, was für menschliche Cybersicherheitsanalyst:innen derzeit eine große Herausforderung darstellt und enorm zeitaufwändig ist. Das aktuelle Produkt erfordert von Cyber Sicherheitsanalyst:innen einen sehr hohen Spezialisierungsgrad,um es bedienen zu können (Expertensystem) und ist aktuell nur mit der sehr zeitaufwändigen und darüber hinaus noch sehr eingeschränkten Abfragen bedienbar.

Unsere Cybersicherheitsteams stehen dabei sowohl aufgrund der Komplexität und des Umfangs bzw der Menge der Bedrohungsinformationen vor immensen Herausforderungen.
Der Mangel an qualifizierten Expert:innen im Bereich der Cybersicherheit schließt ein auf Analysten gestütztes Skalierungssystem (ebenso wie aus wirtschaftlichen Gründen) einfach aus.
MINERVA zielt darauf ab, diese Lücken zu schließen, indem eine KI-Lösung entwickelt wird, die als virtueller Assistent für Cybersicherheitsexpert:innen fungiert und sie bei ihren täglichen Aufgaben unterstützt, indem sie ihnen die Möglichkeit bietet, Informationen mithilfe von natürlicher Sprache stark vereinfacht zu suchen und zu korrekt zu korrelieren.

Der dafür entwickelte KI-Assistent soll große Mengen an verschiedensten Bedrohungsinformationen aus dem Cyber Raum aus unterschiedlichsten Datenquellen (Dokumente, Berichte, Datenbanken, Feeds, News, Standards, etc.) verarbeiten können und analysieren. Auf diese Weise unterstützt er menschliche Expert:innen bei der Korrelation von Informationen aus verschiedenen Quellen, was den Prozess effizienter und effektiver macht und die Suche in diesen Informationsquellen beschleunigt. Das erwartete Ergebnis ist eine erhebliche Verringerung des Zeit- und Arbeitsaufwands für die Analyse von Bedrohungen und die Korrelation von Informationen, so dass sich die menschlichen Expert:innen auf komplexere Aufgaben konzentrieren können. Gleichzeitig soll sichergestellt werden, dass die KI-Lösung den rechtlichen Anforderungen wie dem EU-AI Act entspricht und die Antworten erklärbar sind.

Das auf Open-Source-LLM-Modellen basierte KI-Assistenz-Tool soll eine effiziente Datenanalyse auch für OFFLINE-Systeme ermöglichen und damit Cyber Security Spezialist:innen in kritischen Infrastrukturen unterstützen. Diese skalierbare und anpassungsfähige Lösung ist für kritische Infrastrukturen vor Ort von entscheidender Bedeutung, da sie sicherstellt, keine sensiblen internen/ kritischen Daten im Zuge der Korrelationsprozessen an OpenAI, Microsoft, Google oder andere preis zugeben.

Unser KI-Assistent soll sowohl den IKARUS Experten aber auch den Sicherheitsexperten der Kritischen Infrastrukturen beim schnellen Auffinden, Korrelieren, Identifizieren und Verstehen von Cybersicherheitsbedrohungen helfen. Diese schnelle sowie effiziente Datenverarbeitung und Bedrohungsanalyse ist wichtig um schnell und effektiv auf potenzielle Angriffe und Schwachstellen zu reagieren, die von Angreifer:innen ausgenutzt werden.
Darüber hinaus kann die Lösung, die Open-Source-LLM-Modelle nutzt, so geschult werden, dass sie die Arbeit von Cybersicherheitsanalyst:innen erlernt und sicherstellt, dass das institutionelle Wissen immer innerhalb der Organisation bleibt, trotz Fluktuation oder Ausfall und Wechsel von Mitarbeiter:innen. Vor allem unerfahrene Analyst:innen werden mit dem Werkzeug unterstützt.

In der aktuellen geopolitischen Situation ist der Beitrag der Lösung zur technologischen Unabhängigkeit Österreichs und Europas sehr wichtig. Der „Inhouse“, „On-Premise“ und „Air-Gapped“-Ansatz gewährleistet die nationale Souveränität, indem sensible Daten innerhalb der Organisation behalten werden. Diese Dringlichkeit orientiert sich auch an der signifikanten Zunahme von Cyberangriffen, die auch während der COVID-19-Pandemie zu beobachten war.

IKARUS verfügt bereits über eine funktionierende Lösung zur Verarbeitung und Analyse von Cybersicherheitsbedrohungsinformationen, doch das effektive Durchsuchen und Korrelieren von Bedrohungsinformationen aus verschiedenen Quellen ist zeitaufwändig und erfordert umfassende Kenntnisse der Bedrohungslandschaft und der Funktionsweise der Plattform. Diese schnelle Identifizierung von Bedrohungen ist wichtig, um proaktive Maßnahmen zu ergreifen und die Stakeholder zu informieren, einschließlich der Stakeholder in kritischen Infrastrukturen. KI soll deshalb in die bestehende Lösung von IKARUS integriert werden, um Cybersecurity Expert:innenen bei ihren täglichen Aufgaben zu unterstützen und ihre Effizienz und Entscheidungsfähigkeit zu steigern. Durch die Unterstützung von Expert:innen mit solchen fortschrittlichen Werkzeugen, wird nicht nur die nationale Cybersicherheit gestärkt, sondern auch eine größere technologische Unabhängigkeit zu aktuellen Anbietern außerhalb Europas.
Das Inhouse- und On-Premise-Modell der Lösung, das Open-Source-LLM verwendet, unterstützt dabei das Ziel, Österreichs und Europas Selbstständigkeit in kritischen technologischen Bereichen zu gewährleisten. Es gewährleistet, dass sensible Daten für die nationale Sicherheit sicher im Land verwaltet werden und minimiert die Abhängigkeit von externen Stellen.

Die Erweiterung der bestehenden Plattform um einen KI-Assistenten steht im Einklang mit den Zielen des Schwerpunkts "Effizientes Wissensmanagement". Derzeit besteht die Herausforderung darin, große Mengen an strukturierten und unstrukturierten Bedrohungsinformationen aus dem Cyberraum zu verwalten und wertvolle Erkenntnisse aus diesen Quellen abzuleiten. Dies sind die aktuellen Grenzen unseres traditionellen Informationsmanagementsystems. Gerade deshalb soll mit diesem Projekt die neuesten Fortschritte im Bereich des maschinellen Lernens und der künstlichen Intelligenz integriert werden, einschließlich der semantischen Suche und Open-Source-Sprachmodelle. Dies soll die Art und Weise verbessern, wie Cybersicherheitsexpert:innen ihre Informationen suchen und sammeln, um eine fundierte Entscheidung zu treffen.

Abstract

The motivation for the development of this AI-based approach lies in the extension of the existing cyber security product from IKARUS Security Software (IKARUS TIP - Threat Information Platform) to simplify the use and usability of the product and, thanks to the implemented AI, to enable it to correlate and process huge amounts of threat information, which is currently a major challenge and enormously time-consuming for human cyber security analysts. The current product requires cyber security analysts to have a very high degree of specialization in order to use it (expert system) and can currently only be operated with the very time-consuming and, moreover, very limited queries.

Our cybersecurity teams face immense challenges due to the complexity and the scope and volume of threat information. The lack of qualified cybersecurity experts simply precludes an analyst-based scaling system (as well as for economic reasons).
MINERVA aims to fill these gaps by developing an AI solution that acts as a virtual assistant for cybersecurity experts, supporting them in their daily tasks by providing them with the ability to search and correctly correlate information using natural language in a highly simplified way.
The AI assistant developed for this purpose should be able to process and analyze large amounts of different threat information from cyberspace from a wide variety of data sources (documents, reports, databases, feeds, news, standards, etc.). In this way, it supports human experts in correlating information from different sources, making the process more efficient and effective and speeding up the search in these information sources. The expected result is a significant reduction in the time and effort required to analyze threats and correlate information, allowing human experts to focus on more complex tasks. At the same time, the aim is to ensure that the AI solution complies with legal requirements such as the EU AI Act and that the answers are explainable.
The AI assistance tool based on open-source LLM models is designed to enable efficient data analysis even for OFFLINE systems and thus support cyber security specialists in critical infrastructures. This scalable and adaptable solution is crucial for on-site critical infrastructures as it ensures that no sensitive internal/critical data is exposed to OpenAI, Microsoft, Google or others during the correlation processes.

Our AI assistant is designed to help both IKARUS experts and critical infrastructure security experts to quickly find, correlate, identify and understand cybersecurity threats. This fast and efficient data processing and threat analysis is important to respond quickly and effectively to potential attacks and vulnerabilities exploited by attackers.
In addition, the solution, which uses open source LLM models, can be trained to learn the work of cybersecurity analysts and ensure that institutional knowledge always remains within the organization, despite staff turnover or attrition. Inexperienced analysts in particular are supported with the tool.
In the current geopolitical situation, the solution's contribution to Austria's and Europe's technological independence is very important. The "in-house", "on-premise" and "air-gapped" approach ensures national sovereignty by keeping sensitive data within the organization. This urgency is also based on the significant increase in cyberattacks, which was also observed during the COVID-19 pandemic.

IKARUS already has a working solution for processing and analyzing cybersecurity threat intelligence, but effectively sifting through and correlating threat intelligence from multiple sources is time-consuming and requires extensive knowledge of the threat landscape and how the platform works. This rapid identification of threats is important in order to take proactive measures and inform stakeholders, including those in critical infrastructures. AI will therefore be integrated into the existing

IKARUS solution to support cybersecurity experts in their daily tasks and increase their efficiency and decision-making ability. Supporting experts with such advanced tools will not only strengthen national cybersecurity, but also provide greater technological independence from current providers outside Europe.
The in-house and on-premise model of the solution, which uses open source LLM, supports the goal of ensuring Austria's and Europe's independence in critical technological areas. It ensures that sensitive data for national security is managed securely within the country and minimizes dependence on external agencies.

The addition of an AI assistant to the existing platform is in line with the objectives of the "Efficient knowledge management" focus area. Currently, the challenge is to manage large amounts of structured and unstructured threat information from cyberspace and derive valuable insights from these sources. These are the current limitations of our traditional information management system. This is precisely why this project aims to integrate the latest advances in machine learning and artificial intelligence, including semantic search and open source language models. This should improve the way cybersecurity experts search and gather their information to make an informed decision.

Endberichtkurzfassung

Cybersecurity Expert:innen benötigen einen umfassenden Zugang zu Bedrohungsdaten, detaillierten Berichten und den neuesten Nachrichten, müssen sich aber bei ihrer täglichen Arbeit mit mehreren Tools, Datenbanken und externen Ressourcen auseinandersetzen. Diese Komplexität führt häufig zu einer Informationsflut und erhöht das Risiko, entscheidende Zusammenhänge zu übersehen, die zur Korrelation, Verhinderung oder Eindämmung von Cyberangriffen beitragen könnten. Cloud-basierte Sprachmodelle können zwar die Betriebskosten senken, führen aber zu erheblichen Datenschutzproblemen, insbesondere bei kritischen Infrastrukturen, und schaffen Abhängigkeiten von externen Anbietern.

Im Gegensatz dazu bietet der Einsatz von LLMs On-Premise die Möglichkeit, die Kontrolle über sensible Daten und die lokale Infrastruktur zu behalten, obwohl dies erhebliche Rechenkapazitäten erfordert. Kleinere Modelle kommen mit weniger GPUs aus, eignen sich aber nur für einfachere Abfragen, während größere Modelle komplexere Aufgaben lösen können, insbesondere leistungsstarken GPUs, ausgestattet sind.

Unser erstes MVP nutzte ein kleineres Llama 7B Modell, um eine Bedrohungsdatenbank abzufragen und verschiedene Bedrohungsberichte zu verarbeiten. Erste Ergebnisse deuteten darauf hin, dass ein größeres Modell zu genaueren und aussagekräftigeren Antworten führen würde. Um dies zu prüfen, setzten wir zwei lokal gehostete NVIDIA L40 GPUs mit insgesamt 80 GB VRAM ein, um ein Llama70B Modell mit INT4 quantization zu betreiben. Obwohl dieser Schritt die Leistung bei strukturierten Abfragen und der Dokumentensuche verbessern sollte, erwies sich die präzise SQL-Generierung aus natürlicher Sprache als wesentliche Herausforderung. Die Komplexität der Datenbankschemata – oft geprägt von umfangreicher Geschäftslogik und komplexen Beziehungen – erschwerte es dem Modell, zuverlässige SQL-Abfragen ohne zusätzlichen Kontext oder Expertise zu erstellen. Zwar funktionierte die Dokumentensuche gut, doch das Text-zu-SQL Problem wurde zum zentralen technischen Engpass, was uns dazu veranlasste, unsere Herangehensweise zu überdenken.

In der Folge änderten wir unsere Strategie. Statt von Hardwareeinschränkungen auszugehen, testeten wir ein vollständiges Llama 70B Modell über API-basierte Dienste, das etwa 70 % größer ist als die quantized Variante. Auf diese Weise konnten wir die Fähigkeiten des Modells vor einer finalen On-Premise Implementierung validieren. So gelang es uns, klarere Anforderungen für Kunden festzulegen, die lokale LLM-Lösungen umsetzen wollen – von ersten API-Tests bis hin zur vollständig lokalen Einrichtung mit geeigneter Infrastruktur.

Nach der Validierung des Modells richteten wir unser Augenmerk auf die Datenarchitektur. Wir integrierten Daten aus verschiedenen Cybersecurity Datenquellen, darunter Tenable, Qualys, Microsoft Defender und die National Vulnerability Database (NVD), in ein einheitliches Schema. Dadurch können Analyst:innen nun mithilfe natürlicher Sprachabfragen sowohl auf strukturierte Sicherheitsdaten – etwa Ergebnisse von Schwachstellenscans, Sicherheitsalarme auf Endpoints oder CVE-Daten – als auch auf unstrukturierte Daten wie Bedrohungsberichte und Warnhinweise zugreifen. Indem wir die Datenintegration vorab priorisierten und die zugrunde liegenden Datenbanksysteme vereinfachten, entstand eine Lösung, die Datenhoheit wahrt, den effizienten Zugriff auf sicherheitsrelevante Informationen aus verschiedenen Quellen ermöglicht und vollständig ohne die Abhängigkeit von externen Cloud-Anbietern auskommt.

Projektkoordinator

- IKARUS Security GmbH

Projektpartner

- CyberACI GmbH
- Bundesministerium für Landesverteidigung