

EKID

Einsatz von Künstlicher Intelligenz zur forensischen Analyse von digitalen Daten

Programm / Ausschreibung	KIRAS, Kooperative F&E-Projekte, KIRAS CS Kooperative F&E-Projekte (CS KFE_2025)	Status	laufend
Projektstart	01.01.2027	Projektende	31.12.2028
Zeitraum	2027 - 2028	Projektlaufzeit	24 Monate
Barwert der Projektförderung	€ 537.848		
Keywords	Agentic AI, Multimodale Forensik, Explainable AI, Information Retrieval		

Projektbeschreibung

Strafverfolgungsbehörden stehen bei der forensischen Bearbeitung digitaler Beweismittel vor stark wachsenden, heterogenen Datenbeständen aus unterschiedlichen Quellen wie Messenger-Diensten, E-Mails, Cloud-Speichern, Bildern, Videos oder vollständigen Datenträgerabbildern. Während bestehende Werkzeuge primär Datenextraktion, Filterung und Stichwortsuche unterstützen, sind semantische Verknüpfungen, multimodale Zusammenhänge und strukturierte, nachvollziehbare Analysepfade bislang nur eingeschränkt abgebildet.

Large Language Models (LLMs) und agentenbasierte KI-Ansätze werden international zunehmend diskutiert, ihr systematischer, rechtskonformer und nachvollziehbarer Einsatz in der digitalen Forensik ist jedoch bislang nicht ausreichend wissenschaftlich untersucht. Insbesondere fehlt es an konzeptionellen Rahmenwerken, die technische Leistungsfähigkeit, Transparenz, rechtliche Anforderungen und dokumentierbare Analyseprozesse gemeinsam betrachten.

Ziel des Projekts ist die wissenschaftliche Untersuchung einer modularen agentenbasierten KI-Architektur für die digitale Forensik. Dabei wird analysiert, wie spezialisierte KI-Agenten – etwa für Retrieval, semantische Kontextualisierung, Reasoning, Konsistenzprüfung oder strukturierte Berichterstellung – in einer kontrollierten Forschungsumgebung interagieren können. Ein zentraler Innovationsaspekt besteht in der methodischen Verbindung von Agentic AI und Explainable AI. LLM-basierte Schlussfolgerungsprozesse sollen nicht isoliert betrachtet, sondern als nachvollziehbare, dokumentierte Entscheidungsschritte mit Bezug zu zugrunde liegenden Datenquellen analysiert werden. Die Architektur folgt dabei einem Human-in-the-Loop-Ansatz, bei dem menschliche Entscheidungsverantwortung explizit erhalten bleibt. Rechtliche und regulatorische Rahmenbedingungen – insbesondere Datenschutzrecht und AI Act – werden konzeptionell berücksichtigt und im Projektkontext reflektiert.

Das Projekt zielt darauf ab, eine wissenschaftlich fundierte Grundlage zur Bewertung des Potenzials und der Grenzen agentenbasierter KI-Systeme in der digitalen Forensik zu schaffen. Erwartet werden Erkenntnisse zur Skalierbarkeit modularer Architekturen bei heterogenen Datenlagen, zur Qualität und Nachvollziehbarkeit von LLM-gestützten Reasoning-

Abstract

Law enforcement authorities are confronted with rapidly growing and highly heterogeneous data volumes in the forensic processing of digital evidence. Data originating from messaging services, emails, cloud storage, images, videos, or full disk images are often distributed across different systems and formats. While existing tools primarily support data extraction, filtering, and keyword search, semantic linking of content, multimodal relationships, and structured, traceable analysis workflows remain only partially supported.

Large Language Models (LLMs) and agent-based AI approaches are increasingly discussed at the international level. However, their systematic, legally compliant, and traceable application in digital forensics has not yet been sufficiently investigated from a scientific perspective. In particular, there is a lack of conceptual frameworks that jointly address technical performance, transparency, legal requirements, and documentable analytical processes.

The objective of the project is to scientifically investigate a modular agent-based AI architecture for digital forensics. The project analyzes how specialized AI agents — for example for retrieval, semantic contextualization, reasoning, consistency checking, or structured report generation — may interact within a controlled research environment. A central innovation aspect lies in the methodological integration of Agentic AI and Explainable AI. LLM-based reasoning processes are not considered in isolation but are examined as traceable and documented decision steps linked to underlying evidence sources. The architecture follows a Human-in-the-Loop approach, explicitly preserving human decision-making responsibility. Legal and regulatory requirements — in particular data protection law and the EU AI Act — are conceptually considered and reflected within the project context.

The project aims to establish a scientifically grounded basis for assessing the potential and limitations of agent-based AI systems in digital forensics. Expected outcomes include insights into the scalability of modular architectures when handling heterogeneous data sets, the quality and traceability of LLM-supported reasoning processes, and requirements for transparent and reproducible analytical workflows.

Projektkoordinator

- AIT Austrian Institute of Technology GmbH

Projektpartner

- Bundesministerium für Inneres
- Research Institute AG & Co KG
- UNINET it-consulting GmbH