

SENTRA

Scalable Eventdrive Non-stop-learning for Trusted Retrieval & Analytics

Programm / Ausschreibung	Bundesländerkooperationen TP, Oberösterreich, Security Technologies & Solutions	Status	laufend
Projektstart	01.12.2026	Projektende	30.11.2028
Zeitraum	2026 - 2028	Projektlaufzeit	24 Monate
Barwert der Projektförderung	€ 364.944		
Keywords	Cybersecurity; Machine Learning; Online Fraud Detection; OSINT; LLM Evaluation		

Projektbeschreibung

Ausgangssituation und Motivation

Betrugsphänomene im Internet – darunter Fake-Shops, Scam-Maschen, Crypto-Handelsbetrug und der Vertrieb gesundheitsgefährdender Produkte über Online-Kanäle – betreffen Bürgerinnen und Bürger zunehmend. Gleichzeitig verlagern sich Informations- und Kaufprozesse auf KI-basierte Such- und Empfehlungssysteme (Large Language Models, Agenten), deren Ergebnisse neue Angriffsflächen für Betrug und Manipulation eröffnen. Bestehende Monitoring- und Analysewerkzeuge sind entweder plattformabhängig, intransparent oder auf einzelne Bedrohungsklassen beschränkt. Eine integrierte, souveräne Forschungsinfrastruktur, die von der Datenerhebung bis zur evidenzbasierten Handlungsempfehlung reicht, fehlt bislang.

Ziele und Innovationsgehalt

SENTRA erforscht eine skalierbare, ereignisgetriebene Infrastruktur zur souveränen Gewinnung, Analyse und Bewertung sicherheitsrelevanter Daten aus großen Online-Plattformen (Very Large Online Platforms, VLOPs). Die zentralen Innovationen umfassen:

- (1) adaptive, plattformunabhängige Methoden zur Datenextraktion, die Änderungen in Webstrukturen automatisiert erkennen und kompensieren;
- (2) eine KI-gestützte Analysepipeline mit integrierter Drift-Erkennung und automatisierten Retraining-Triggern (Continuous Learning);
- (3) ein neuartiges LLM-Evaluierungsframework, das KI-Systeme erstmals als potenzielle Vektoren für die Verbreitung betrügerischer oder gefährlicher Inhalte systematisch vergleicht;
- (4) die durchgängige Verknüpfung von Datenextraktion, Klassifikation und Insight-Generierung zu einer auditierbaren, nachvollziehbaren Pipeline.

Die gesamte Infrastruktur wird in Oberösterreich konzipiert und betrieben (digitale Souveränität).

Angestrebte Ergebnisse und Erkenntnisse

Die erforschten Methoden werden anhand von drei Forschungsszenarien validiert:

(1) KI-Erkennung manipulativer Online-Angebote

Automatisierte Identifikation von betrügerischen und irreführenden Online-Inhalten (z. B. Crypto-Scams, Fake-Shops, verbotene Gesundheitsversprechen) auf Social Media.

(2) Automatisiertes Produktscreening

Erkennung gefährlicher, verbotener oder falsch deklarerter Produkte durch Analyse von Produkttexten und Inhaltsstoffen.

(3) Trendanalyse & Frühwarnung bei Desinformation

Früherkennung von Desinformation, riskanten Trends oder gefährlichen Konsummustern durch Analyse öffentlicher Online Daten.

Angestrebte Ergebnisse sind Datenextraktion aus mindestens fünf unterschiedlichen VLOP-Quellentypen mit einer Extraktionsqualität (Precision) über 90 %, eine Klassifikationsgenauigkeit (F1-Score) über 0,85 für die Fake-Shop-Erkennung sowie die systematische Evaluierung von mindestens drei LLM-Systemen. Die Erkenntnisse bilden die Grundlage für operative Anwendungen in Marktüberwachung, Verbraucherschutz und Sicherheitsanalyse und stärken Oberösterreich als Standort für KI-basierte Sicherheitsforschung.

Abstract

Initial Situation and Motivation

Fraud phenomena on the internet, including fake online shops, scam schemes, cryptocurrency trading fraud, and the sale of health-endangering products via online channels, are increasingly affecting citizens. At the same time, information and purchasing processes are shifting toward AI-based search and recommendation systems (large language models, agents), whose outputs create new attack surfaces for fraud and manipulation. Existing monitoring and analysis tools are either platform-dependent, opaque, or limited to individual threat categories. A sovereign, integrated research infrastructure spanning the full pipeline from data collection to evidence-based recommendations for action is still missing.

Objectives and Degree of Innovation

SENTRA investigates a scalable, event-driven infrastructure for the sovereign acquisition, analysis, and assessment of security-relevant data from major online platforms (Very Large Online Platforms, VLOPs). The key innovations include:

- (1) adaptive, platform-independent data extraction methods that automatically detect and compensate for changes in web structures;
 - (2) an AI-supported analysis pipeline with integrated drift detection and automated retraining triggers (continuous learning);
 - (3) a novel LLM evaluation framework that, for the first time, systematically compares AI systems as potential vectors for the dissemination of fraudulent or dangerous content;
 - (4) the end-to-end linking of data extraction, classification, and insight generation into an auditable, traceable pipeline.
- The entire infrastructure will be designed and operated in Upper Austria (digital sovereignty).

Targeted Results and Insights

The researched methods will be validated using three research scenarios:

(1) AI-based Detection of Manipulative Online Offers

Automated identification of fraudulent and misleading online content (e.g., crypto scams, fake shops, illicit health claims) across social media and publicly accessible online platforms.

(2) Automated Product Screening

Detection of dangerous, prohibited, or incorrectly declared products through the analysis of product descriptions, ingredient lists, and related metadata from web shops and large online platforms.

(3) Trend Analysis & Early Warning for Disinformation

Early detection of disinformation, emerging risk patterns, and harmful consumption trends through semantic and structural analysis of publicly available online data.

Target outcomes include data extraction from at least five different VLOP source types with extraction quality (precision) above 90%, a classification performance (F1 score) above 0.85 for fake-shop detection, and the systematic evaluation of at least three LLM systems. The findings will form the basis for operational applications in market surveillance, consumer protection, and security analysis, and will strengthen Upper Austria as a location for AI-based security research.

Projektkoordinator

- FH OÖ Forschungs & Entwicklungs GmbH

Projektpartner

- X-Net Services GmbH
- Österreichische Agentur für Gesundheit und Ernährungssicherheit GmbH