

Defence-related Research Action - DEFRA

ACRONYM: SIMBA

Title: Smart Interactive Multi-Agent Bias-Resilient Analysis

Duration of the project: 1/12/2025 - 1/03/2029

Key words: Intelligence analysis, Foreign Information Manipulation and Interference (FIMI), Human-centred explainable AI, Multi-agent AI (MAAI), Large language models (LLMs), Disinformation

Budget: 1.619.000 €

of which RHID contribution:

1.210.000 €

PROJECT DESCRIPTION

Context: The intelligence environment presents significant challenges for Belgian defence, marked by an overwhelming volume of information often compromised by sophisticated disinformation and hybrid warfare tactics. Intelligence analysts must discern credible threats from this complex landscape to inform critical decision-making. Existing analytical tools often focus on surface-level analysis, failing to address the underlying argumentation structures and potential biases within intelligence sources. This leaves defence vulnerable to flawed intelligence based on manipulation or cognitive biases. The SIMBA project directly addresses this critical gap, a need formally recognised and supported by the Belgian General Intelligence and Security Service (SGRS/ADIV).

General objectives: The project's main objective is to significantly enhance the capabilities of intelligence analysts by developing an innovative system that assists in identifying and evaluating biases and logical fallacies within textual intelligence sources. To achieve this, SIMBA will:

- Develop a flexible multi-agent AI (MAAI) framework capable of integrating multiple specialised software agents focused on source analysis and the detection of cognitive biases and logical fallacies.
- Integrate advanced natural language processing (NLP) capabilities by adapting Textgain's specialised European language model, CaLICO, to power the analytical agents.
- Design an adaptive, human-centred explanation interface based on explainable AI (XAI) principles to present findings to analysts in a clear, intuitive, and actionable manner.
- Integrate all components into a functional system demonstrator (TRL5-6) and validate its capabilities in realistic, simulated intelligence analysis scenarios relevant to defence.

Methodology: The 36-month project will be executed through a collaborative partnership between KU Leuven (coordination and XAI interface), Sopra Steria (MAAI framework and system integration), and Textgain (NLP/LLM capabilities). The methodology employs an iterative, user-centred approach, ensuring continuous alignment with the needs of intelligence analysts through regular feedback loops. Development will proceed through three main phases: foundation and design (M1-M9), core development and integration (M10-M24), and evaluation, refinement, and final demonstration (M25-M36).

Potential impact of the research on Belgian Defence: SIMBA aims to significantly enhance the speed, accuracy, and reliability of the intelligence cycle. By semi-automating the detection of potential reasoning flaws and biases, the system will enable analysts to focus on higher-level assessment and critical thinking. This directly impacts defence by improving situational awareness, supporting more robust decision-making, and bolstering resilience against information manipulation and cognitive warfare threats. The project is designed for potential future integration into the SGRS/ADIV's Next-generation AI-supported Integrated Intelligence Software Suite (NAIS), ensuring its practical relevance and strengthening Belgian expertise in cutting-edge AI applications for defence and security.

Description of the expected final research results and valorisation perspectives: The primary expected result is a TRL5-6 demonstrator of the SIMBA system, validated in two defence-specific use case scenarios (foreign influence concerning the Ukraine/Russia conflict and NATO perception). A third, public-facing demonstrator will showcase a dual-use application in fact-checking support.

Other key results will include:

- A reusable MAAI framework for bias-resilient analysis.
- Novel NLP/LLM techniques for bias and fallacy detection.
- Validated human-centred XAI methods and user interface components.
- Several peer-reviewed scientific publications.

In the short term, Belgian defence is the primary exploiter, with potential for integration into NAIS and use in analyst training. At medium term, the project partners will valorise the results through enhanced commercial services (Sopra Steria, Textgain) and further academic research (KU Leuven). The project also lays the groundwork for follow-on initiatives aiming to operationalise the technology (TRL7-9) or apply it to new domains such as law enforcement and crisis management.

CONTACT INFORMATION

Coordinator

Katrien Verbert

KU Leuven – Department of Computer Science

katrien.verbert@kuleuven.be

Partners

Cédric Genin

Sopra Steria Belgium – Defence & Security Business Unit

cedric.genin@soprasteria.com

Redouan el Hamouchi

Textgain

redouan@textgain.com

LINK(S)

Not available (yet).