Defence-related Research Action - DEFRA

ACRONYM:AHOI

Title: Adaptive Human Operator Interaction with Autonomous Systems

Duration of the project: 31/12/2023 - 1/3/2027

Key words: Explainable AI (XAI), Human-Machine Teaming, Situational Awareness, Decision-Making under uncertainty, Intelligent Automation Systems, Human-centred design.

Budget: 1.667.000 €

of which RHID contribution: 1.588.493€

PROJECT DESCRIPTION

In the realm of autonomous systems, the role of human oversight and interaction remains crucial. Machines, while adept at navigating uncertain scenarios, often fall short in accounting for ethical considerations embedded in real-world situations. This is particularly evident in AI-driven systems, as modern machine learning approaches tend to be "black boxes", obscuring the underlying mechanisms of their decision-making processes. A primary risk with the mainstream deployment of such systems is that their decisions may lack transparency altogether and due to that, the humans working alongside them, particularly in highly uncertain and / or unsafe scenarios, may have difficulty in monitoring their performance, understanding their processes, and determining whether they are fulfilling their intended purpose and are operating within the boundaries of the socio-ethical value system. All of these issues can be exacerbated by human biases and heuristics, which can cause a navigator to (1) be overconfident in their navigational skills and knowledge, and ignore or overlook warning signals from an autonomous system (2) dismiss the autonomous system's recommendations that contradict their own beliefs or assumptions (e.g. confirmation bias). The outcome of this process is a breakdown of trust in system recommendations.

There is thus a strong need to investigate how explainability of such algorithms and human trust in systems shape each other. To achieve this, the Adaptive Human Operator Interaction with Autonomous Systems consortium (AHOI) brings together a team of researchers from various research domains (Artificial Intelligence, Explainable AI, behavioural science, maritime personnel training and human machine interaction) to investigate the factors that facilitate or impede the handover of responsibilities (handover points) between human and autonomous systems. Specifically, we will focus on the role of trust, and how it is affected by information sharing and transparency, by dynamically tracking key performance indicators of the system and its human operator within an operational, maritime context. The maritime domain presents unique challenges for autonomous

navigation systems, requiring resilience to a range of internal and external constraints. Our research aims to address these challenges by:

1. Robust Navigation System
Developing a resilient autonomous navigation system that operates seamlessly in dynamic maritime environments, including sensor failures, engine malfunctions, dense traffic, and adverse weather conditions. This system will leverage advanced machine learning techniques to generalize to unknown, unstructured settings with minimal reconfiguration.

2. Explainable AI for Trustworthy Decisions
Investigating the impact of model explanation methods (XAI) on human operators' trust and decision-making processes. We will examine how XAI can enhance human understanding of AI-generated decisions, particularly for operators with varying levels of experience and expertise.

3. Human-Machine Collaboration in Decision-Making
Exploring the relationship between human biases and the perception of transparency in human-machine interaction (HMI). We will conduct experiments to determine the optimal point of contact between humans and machines, enabling them to effectively collaborate in situations with high uncertainty or bias.

4. Advanced HMI for Enhanced Transparency
Designing an advanced HMI tool that not only presents AI-generated decisions but also provides insights into their underlying reasoning. This HMI will dynamically adjust its output based on user profiles, tailoring explanations to individual levels of experience and fostering trust in AI decision-making.

5. Adaptive XAI and Feedback-Driven Optimization
Leveraging new methodologies in XAI and visualization software to create a dynamic and interactive HMI. This HMI will gather user feedback and adapt XAI explanations accordingly, improving AI performance and fostering a continuous learning cycle.

Our holistic approach to autonomous navigation aims to create a system that is both robust and transparent, enabling seamless human-machine collaboration in the complex maritime domain. By addressing the challenges of reliability and explainability, we can pave the way for safe and efficient autonomous operations in the future.

This work is relevant for various defence related applications, where systems might operate independently, but still require human involvement and the right chain of command. Although focused on autonomous vessel navigation scenario, our research outcomes can be generalized to applications such as mine hunting, Intelligence, Surveillance and Reconnaissance (ISR) missions, UAV operations, and UGV operations etc. In AHOI, iMec will conduct research on AI and XAI, while UA and AMA will conduct in-depth study of the biases and profiling of the human operators. MAHI will conduct research on the situational awareness of the autonomous vessels and design the final HMI.

CONTACT INFORMATION

Coordinator
Dirk Van Rooy
University of Antwerp, Product development.
e-mail: : Dirk.VanRooy@uantwerpen.be

Partners
Ali Anwar, José Oramas
IMEC - Internet Data Lab (IDLab) Antwerp
e-mail: Ali.Anwar@imec.be , Jose.Oramas@imec.be

Pieter-Jan Note
MAHI
e-mail: pieterjan.note@mahi.be

Pieter Maes
Antwerp Maritime Academy
e-mail: pieter.maes@hzs.be

LINK(S)

/