EDF-2021-CYBER-R: Cyber threat intelligence and improved cyber operational capabilities

Proposals are invited against any of the following topic:

EDF-2021-CYBER-R-CDAI: Improving cyber defence and incident management with artificial intelligence

<u>Budget</u>

The Union is considering a contribution of up to EUR 13 500 000 to support proposals addressing any of the abovementioned topic and its associated specific challenge, scope, targeted activities and main functional requirements.

Several actions, addressing different solutions, may be funded under this call.

The Commission will pay particular attention to the other R&D and dual-use on-going initiatives at Union level to avoid unnecessary duplication.

The ability to detect and respond to security incidents suffers from several challenges, including: the ever increasing amount of data that needs to be analysed in order to detect and fully understand security incidents; the number of false alarms generated resulting in, for instance, erroneous prioritisation and alarm fatigue amongst operators and analysts; lack of (human) resources to sufficiently analyse all potentially malicious activity; the decreasing effectiveness of traditional defence measures based on known set of rules (e.g. a priori known signatures and/or network traffic profiles) due to the increase of encrypted network traffic and their inadequacy against advanced persistent threats and zero-day attacks (including malware that exploits unknown vulnerabilities, targeted phishing attacks, low-rate data exfiltration, abnormal user behaviour, etc.); choosing appropriate measures in response to attacks in a timely manner, when the scope is uncertain and the situation develops faster than a human being may follow without advanced decision-making support, and while the compromise potentially have or will extend over weeks, months or years.

The use of Artificial Intelligence (AI)[1] seems promising in order to address many of these challenges – and AI has recently shown great results in areas such as playing strategic games and analysing text.

This call seeks proposals that help increase the level of automation in incident management and cyber defence activities through the use AI. In this setting, the engagement of state-of- the-art AI methods should be used to automate incident management and cyber defence activities, including incident detection and response, carried out by security operation centres (SOCs), and cyber defence teams (or similar entities) when they detect and analyse events and determine what actions to take.

Modern SOCs are sometimes equipped with security orchestration, automation and response (SOAR) capabilities that allows human operators to respond to attacks with predefined

[1] e.g. machine learning, deep learning, decision trees, statistical outliers, probabilistic networks, genetic algorithms, reinforcement learning, situation calculus, ontologies, symbolic reasoning, etc.

“playbooks” designed to mitigate ongoing attacks, e.g. by disabling user accounts or reconfigure firewalls, where AI-based solutions seem applicable.

AI can, for instance, be used to complement rule-based detection methods (e.g. through deep learning), to enhance alarms from detection systems using threat intelligence feeds, extract actionable intelligence from the enormous amount of monitoring data and events, correlate alarms with other information to identify attack patterns, automatically respond to events based on the analysis, and recommend actions to human operators. Recent studies unveil that more than two thirds of the organisations included in the studies acknowledge that they are not able to respond to critical threats without AI.

Specific challenge

Creating an AI-based solution that automates larger parts of incident management and cyber defence processes involves several technical challenges. These include (among others): the selection and pre-processing of appropriate data sources; creating and applying models and techniques for analysing the output of sensors to assess if an attack may be happening, including selection and tuning of algorithms and parameters; mapping ongoing attacks to known threats (e.g. using threat intelligence); assessing if the consequences of implementing a particular response outweigh the risks associated with not doing so; and creation/selection of appropriate datasets for training and testing the models. Moreover, transparency and configurability are key requirements, especially in the military domain, which is lacking for most commercial Security Information and Event Management platforms (SIEMs).

While AI-based solutions may be required to address these challenges in this ecosystem, incorporating AI also introduces a new set of technical and non-technical challenges. Questions addressing technical challenges: At which point should an alarm be raised and when should it be elevated? How can AI reason about trust with respect to the use of external information (e.g. threat intelligence)? How can the possible consequences of a cyber-attack, and the consequences of implementing different mitigating means, be assessed before and during response, both in real-time, near real-time/short term or as part of a medium or long term defensive strategy? Many incidents will (at some level) require human interaction and human decision-making – how does an AI based system communicate the results and the underlying explanation and reasoning leading to the result? Non-technical challenges include: Which decision rights and processes can, and should, be delegated to an AI-based system and which should remain manual? How should AI be utilised at strategic, tactical and more technical levels? What is the difference between communication at a technical, tactical, operational and strategic/political level? How can humans work together with AI based systems at the different levels? Military systems increasingly use, and depend on, the private sector and civilian infrastructure – how does incident response differ between military and civilian sectors and what are the challenges in a combined military/civilian setting? What are the implications for AI?

Scope

Addressing the identified challenges will require inter- and multidisciplinary approaches, where teams conduct work of both a technical and a non-technical nature. Analysis of technical, tactical, operational, strategic and political considerations are required. On a technical level, proposals should provide proof-of-concept solutions for AI-based incident management and cyber defence, including detection, mitigation and response. Capable intrusion detection systems (IDS) could form a starting point for proposals. However, proposals must not seek to further the analysis capabilities of IDS alone, but in the context of an automated or semi-automated system for handling incidents.

In additional to purely technical solutions, processes and actors of selected enterprises may need to be mapped, modelled and understood to ensure fit-for-purpose solutions and answer questions of a more conceptual nature. Proposals are further expected to consider the interaction between human operators, analysts and decision makers and the automated or semi-automated incident management and response system.

A suitable methodology for building contextual understanding is expected through case studies of selected processes, incidents and cyber-attacks of selected enterprises, and case studies of successful detection approaches and resilience oriented success stories where technical and non-technical challenges can be studied and addressed at different levels. For the development of technical proof-of-concept prototypes, an appropriate development approach, which includes user and stakeholder involvement, should be leveraged.

Targeted activities

The proposals must cover the following activities as referred in article 10.3 of the EDF Regulation:

- Activities aiming to create, underpin and improve knowledge, products and technologies, including disruptive technologies, which can achieve significant effects in the area of defence;
- Activities aiming to increase interoperability and resilience, including secured production and exchange of data, to master critical defence technologies, to strengthen the security of supply or to enable the effective exploitation of results for defence products and technologies;
- Studies, such as feasibility studies to explore the feasibility of new or improved technologies, products, processes, services and solutions
- Design of defence products, tangible or intangible component or technology as well as the definition of the technical specifications on which such design has been developed which may include partial tests for risk reduction in an industrial or representative environment.

All proposed activities ultimately support the creation of fit-for-purpose proof-of-concept prototypes of an automated or semi-automated incident management and cyber defence system, for select phases in the incident management cycle including detection and response. The prototypes may support human operators, analysts and decision-makers at all levels (technical, tactical, operational, strategic and political) and are expected to contribute to enhanced cyber situational awareness, increased military infrastructure resilience and improved protection against advanced cyber threats.

The activities are sorted into three types of tasks:

1. Enhancing contextual knowledge of the enterprises, processes and decision- making where AI should be utilised

2. Developing AI-based techniques supporting specific human operator/analyst tasks

3. Exploring and developing AI as a decision-maker given limited authority in incident management and cyber defence

Feasibility studies drawing upon real-world scenarios should be utilised to ensure that developed proof-of-concepts and techniques are fit-for-purpose. It may also be necessary to create reference systems and appropriate tests cases to generate training data and evaluate the efficacy of different solutions, both with and without human operators interacting with the system.

The proposals[2] must include at least one activity from task 1 and one activity from task 2 and must be coherent with the defined scope as described above.

1. Enhancing contextual knowledge of the enterprises, processes and decision-making where AI should be utilised

1.1. Knowledge building through analysis of real-life situations, use cases and incidents, in order to sufficiently understand and model the enterprise processes and decision-making processes that AI-based incident management and cyber defence systems will interact with. This includes understanding the relevant actors, their enterprises and business/missions, and the threat environment they operate in.

Proposals may include processes and work flows involving human operators, analysts and decision makers at all levels. Proposals may also cover elements such as information requirements, dealing with uncertainty, strategic objectives, mission objectives, the role of ICT, risk analysis, risk appetite, incident and crisis communication to different stakeholders etc.

1.2. Exploring the boundaries for AI-based autonomous or semi-autonomous response. The playbooks of many SOCs and similar entities describe how to respond to given attacks. However, in depth understanding of the broader context is necessary in order to avoid inappropriate measures. To take into account the broader context in order to avoid inappropriate measures at least the following questions may be addressed: Can such playbooks be automated and can AI reasoning capabilities automate such intuition, experience and contextual understanding? When most machine learning and deep learning algorithms require a vast amount of data to learn from, which will not be available for incident response, can one-shot/few-shot learning (e.g. human-style learning) be utilised in this setting to learn how operators respond to incidents? Can symbolic approaches work in conjunction with machine learning (e.g. neuro-symbolic AI) to automate playbooks?

2. Developing AI-based techniques supporting specific human operator/analyst tasks

2.1. Creation of AI-based techniques for detecting and understanding adversarial activity. This may include analysing and triaging alarms, conducting forensics, utilising external information with varying levels of trust (e.g threat intelligence), leveraging behavioural analytics, performing kill-chain detection and analysis, assessing potential attacker intentions, monitoring applications and communication activities, analysing malware, etc.

[2] Proposals are not expected to include all listed tasks

The techniques may be intended for both real-time and non-real-time detection and analysis, involve multi-disciplinary approaches, use data from endpoints, networks and the cloud, and leverage distributed computing and data processing for real-time scalability.

2.2. Creation of AI-based techniques for building knowledge about own protected ICT systems (e.g., a "cyber record" with current and historical information). This must include collecting, linking and fusing different kinds of information about the system hardware, software, and the relationship between them. Information that may be collected is, for instance, architecture and configuration data, hardware location and specifications, installed applications, network information, services, protocols in use, connected peripheral devices etc. A variety of sources may be leveraged for acquiring information, including hardware and software configuration management systems, documentation, vulnerability scanners, SIEMs, asset discovery tools, etc., and techniques such as reverse engineering may be utilised.

2.3. Creation of AI-based techniques for analysing enterprise systems to appraise the value of assets and the potential consequences of different responses (e.g. configuration changes). This must include both static values manually assigned or derived from fixed factors, and dynamic values that must be seen in relation to ongoing and changing business operation or military missions.

2.4. Creation of explainable AI-based techniques. Many of the most promising machine learning systems are not considered to be "intelligible" or "explainable", which has resulted in a sub-field coined explainable AI (XAI). This task should address how XAI can be utilised to explain detection, analysis and responses at different levels to different actors? Can, for instance, machine learning (e.g. deep learning) be combined with more traditional symbolic AI to make the analysis transparent?

3. Exploring and developing AI as a decision-enabler given limited authority in incident management and cyber defence

3.1. Creation of AI-based information collection and storage systems that dynamically adapts its collection and storage strategy to the situation as continuously analysed and perceived by the system. This includes what is collected, where it is collected and the granularity (e.g. increasing the level of detail of collected information, such as full packet capture, after an initial compromise is detected). As it is not feasible to collect everything and everywhere, such dynamic big data analytics and data lake systems could help the issue of insufficient data due to limited data collection.

3.2. Creation of AI-based decision systems which are risk and impact aware. They should be able to analyse and understand the impact of security incidents on desired mission performance, identify associated risks, generate different response options to maintain requisite cyber resilience and mission assurance, and potentially select and execute a response option if permitted. The analysis of impact, risks, different response options and potential execution should be explainable.

Functional requirements

The proposal must address:

- At least one framework for mixed AI-human cyber defence and incident response suitable for either military or civilian contexts, or both.
- At least one proof-of-concept prototype of an automated or semi-automated incident management system, for select phases in the incident management cycle including detection and response.

Requirements are set across the following categories:

- Architecture and design
- Integration with existing systems
- Utilization of AI/ML for certain tasks
- Reasoning, learning and validation

Architecture and design

- The proposal should include an architecture that is open, modular, scalable, resilient and highly available.
- The proposal could include graphical and programmable interfaces to communicate with both operators and other programs.
- The proposal could include secure cooperative training and sharing of models between different organisational units and nations, e.g. through federated learning.

Integration with existing systems

- Proof-of-concept prototypes should be able to collect telemetries or raw data from existing sources (e.g., sensors systems), including information such as endpoint metrics and logs, network traffic information (e.g. Netflow), and events and logs from both security appliances and application servers.
- It is desirable for proof-of-concept prototypes to interact with, and leverage, existing rule-based detection and classification solutions for coordinated utilisation of both AI and non-AI techniques.

Utilization of AI/ML for certain tasks

- Proof-of-concept prototypes should engage AI/ML in order to improve incident detection rates (e.g. accuracy and recall), compared with existing rule-based solutions.
- Proof-of-concept prototypes should engage AI/ML to automatically propose and potentially effectuate mitigating actions.

Reasoning, learning and validation

- Proof-of-concept prototypes should be able to explain their reasoning and decision making to human operators. It is desirable that algorithms and models are transparent, documented and configurable by the user.
- Proof-of-concept prototypes should be possible to train and configure with information that is readily available or straightforward to produce in contemporary enterprises.
- Proof-of-concept prototypes must be validated in a relevant environment (e.g. reference system) when exposed to relevant test cases (e.g. contemporary network attacks).

Expected impact

- Knowledge on the use case(s) for automated and semi-automated incident response systems, including their practical feasibility and usefulness, limitations, integration into manual processes and interaction with human operators.
- Proof-of-concept for AI-based capabilities for incident detection, analysis and response of selected attacks and scenarios.
- Advanced preparedness of cyber defence operators and improved cyber operational capability, which contributes to EU cyber defence posture.