ANNEX 4

Terms of Reference (ToR) for
ETSI ISG "Securing Artificial Intelligence" (ISG SAI)

Approved by the Director-General on **2 September 2019**, following ETSI Board consultation

Scope

The Securing Artificial Intelligence Industry Specification Group (ISG SAI) will develop technical specifications that mitigate against threats arising from the deployment of AI, and threats to AI systems, from both other AIs, and from conventional sources. As a pre-standardisation activity, the ISG SAI is intended to frame the security concerns arising from AI and to build the foundation of a longer term response to the threats to AI in sponsoring the future development of normative technical specifications.

The underlying rationale for ISG SAI is that autonomous mechanical and computing entities may make decisions that act against the relying parties either by design or as a result of malicious intent. The conventional cycle of risk analysis and countermeasure deployment represented by the Identify-Protect-Detect-Respond cycle needs to be re-assessed when an autonomous machine is involved.

The intent of the ISG SAI is therefore to address 3 aspects of AI in the standards domain:

1. Securing AI from attack e.g. where AI is a component in the system that needs defending.
2. Mitigating against AI e.g. where AI is the 'problem' (or used to improve and enhance other more conventional attack vectors),
3. Using AI to enhance security measures against attack from other things e.g. AI is part of the 'solution' (or used to improve and enhance more conventional countermeasures).

Whilst it is conceivable that the same approach may be used in analysis of each aspect of a Secure AI there are significant distinctions between attack and defence that will require care in addressing. Achieving a common understanding of this duality is key to the successful development of guidance from the ISG. The purpose of the ISG SAI is to develop the technical knowledge that acts as a baseline in ensuring that AI is secure. The stakeholders impacted by the activity of the ISG include all the member groups represented in ETSI and some of the wider societal environments that AIs will be deployed in. This includes end users, manufacturers, operators and governments The activity of the ISG will include gathering concerns of each stakeholder group to ensure that ETSI and the output of the ISG SAI correctly address all of those concerns.

Areas of Activity

The ISG SAI will produce both informative documents (Group Reports) and normative documents (Group Specifications).
The ISG's work will initially focus on the following deliverables which will also be used to help refine the future direction of the ISG work programme. In each case the role of AI as the system, the attacker and the defence will be considered (i.e. addressing modes of securing AI, securing systems against AI, and using AI to secure systems):

AI Threat Ontology

Currently, there is no common understanding of what constitutes an attack on AI and how it might be created, hosted and propagated. The work to be undertaken here will seek to define what would be considered an AI threat and how it might differ from threats to traditional systems
Hence, the **AI Threat Ontology** deliverable seeks to align terminology across the different stakeholders and multiple industries. This document will define what is meant by these terms in the context of cyber and physical security and with an accompanying narrative that should be readily accessible by both experts and less informed audiences across the multiple industries.
Note that this threat ontology will address AI as system, attacker and defence.

Securing AI Problem Statement

This document is modelled on the ETSI NFV "Security Problem Statement" [4] which has been highly influential in guiding the scope of ETSI NFV and enabling "security by design" for NFV infrastructures. It will define and prioritise potential AI threats along with recommended actions.

The recommendations contained in this document will be used to define the scope and timescales for the follow-up work, both within the ISG and also external bodies.

Data Supply Chain Report

Data is a critical component in the development of AI systems, both raw data, and information and feedback from other AI systems and humans in the loop. However, access to suitable data is often limited, causing a need to resort to less suitable sources of data. Compromising the integrity of data has been demonstrated to be a viable attack vector against an AI system.
This report will summarise the methods currently used to source data for training AI, along with a review of existing initiatives for developing data sharing protocols. It will then provide gap analysis on this information to scope possible requirements for standards for ensuring integrity in the shared data, information and feedback, as well as the confidentiality of these.

In addition, the following deliverables may be considered to be produced by ISG SAI:

- Group Report(s) describing the challenges related to securing AI enhanced infrastructures
- Group Specification(s) of the end-to-end security mitigations for AI systems
- Group Specification(s) of the functional architecture and solutions for the provision of a secure interconnection of AIs
- Group Specification(s) including interfaces/APIs/protocols and information / data models to secure the target infrastructures,
- Group Report(s) describing the challenges related to securing infrastructures against AI/ML amplified threats,
- Group Specification(s) of the key use cases and related security requirements in relation to AI/ML threats,
- Group Specification(s) on test methodologies used to validate that threats can be mitigated in the target use cases.
- Group Report(s) describing the opportunities offered by AI/ML security technologies
- Group Specification(s) including interfaces/APIs/protocols and information / data models to produce robust, effective AI/ML security products
- Group Report(s) providing a descriptive Proof of Concept (PoC) framework with minimum requirements, templates, process description,
- Group Report(s) providing gap analysis of the work done in existing standards and open source groups in relation to the agreed AI use cases,
- Group Specification(s) of the business use cases and related requirements (including analysis of outputs from each of ISG ENI and ISG ZSM).

The detailed ISG work plan may be modified as the work and project priorities evolve and will be maintained and made available on the ETSI portal.

Throughout the project, the ISG will provide recommendations to the existing standards groups when impacts on their specifications are foreseen.

Annex (informative): collaboration with other bodies

Close collaboration and coordination with other standard groups is required to ensure that all the organizations together provide complementary solutions. It will also be necessary to identify and agree the roles of the corresponding standardization bodies in filling the identified gaps.

ISG SAI will set-up the appropriate communication channels to the following groups both within and outside of ETSI.

ETSI groups

The ISG initially intends to establish relationships with the following ETSI groups:

- ETSI PP 3GPP
- ETSI PP oneM2M
- ETSI TC SmartM2M
- ETSI TC SmartBAN
- ETSI TC CYBER
- ETSI ISG CIM
- ETSI ISG ENI
- ETSI ISG NFV
- ETSI ISG ZSM
- ETSI EP eHEALTH

and others as identified during the progression of the work.

External groups

The ISG also intends to cooperate with a number of external organizations including:

- ENISA
- IETF
- ITU-T
- ISO/IEC JTC1
- NIST

and others as identified during the progression of the work.