|  |
| --- |
| **Liaison Statement** |
| **Title:** | **Liaison statement from ETSI ISG SAI on Securing Artificial Intelligence** |
| Date: | 2019-12-04 |
|  |  |
| **From** (source): | ETSI ISG SAI |
| Contact(s): | isgsupport@etsi.org |
|  |  |
| **To:** | ETSI TC CYBER, 3GPP SA3, oneM2M, ETSI TC SmartM2M, ETSI TC SmartBAN, ETSI ISG CIM, ETSI ISG ENI, ETSI ISG NFV, ETSI ISG PDL, ETSI ISG ZSM, ETSI EP eHEALTH, TC MTS, CEN/CLC Focus Group on Artificial Intelligence, CEN/CLC JTC 13, AI for EU group, ISO/IEC JTC 1 SC 42 and SC 27, GSMA Fraud and Security Group, 3GPP SA, TCCA-SFPG, ETSI TC TCCE-WG6, GSMA TSG (Terminal Steering Group), ENISA, IETF, ITU-T SG17, ISO/IEC JTC1, NIST, CIS |
| **Copy to:** |  |
|  |  |
| Response to:(if applicable) |  |
|  |  |
| Attachments: (if applicable) |  |
|  |

This is to announce that the Kick-off Meeting for the new ETSI ISG on Securing Artificial Intelligence (ISG SAI) was held on 23 October 2019.

The intent of the ISG SAI is 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).

The ETSI ISG SAI aims to develop the technical knowledge that acts as a baseline in ensuring that artificial intelligence is secure. Stakeholders impacted by the activity of this group include end users, manufacturers, operators and governments.

At the first meeting the following New Work Items were agreed:

[***AI Threat Ontology***](https://portal.etsi.org/webapp/WorkProgram/Report_WorkItem.asp?WKI_ID=58856)

*The purpose of this work item is to define what would be considered an AI threat and how it might differ from threats to traditional systems. The starting point that offers the rationale for this work is that currently, there is no common understanding of what constitutes an attack on AI and how it might be created, hosted and propagated.*

*The AI Threat Ontology deliverable will seek 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, an adversarial attacker, and as a system defender*

[***Data Supply Chain Report***](https://portal.etsi.org/webapp/WorkProgram/Report_WorkItem.asp?WKI_ID=58857)

*Data is a critical component in the development of AI systems. This includes raw data as well as information and feedback from other systems and humans in the loop, all of which can be used to change the function of the system by training and retraining the AI.*

*However, access to suitable data is often limited causing a need to resort to less suitable sources of data. Compromising the integrity of training data has been demonstrated to be a viable attack vector against an AI system. This means that securing the supply chain of the data is an important step in securing the AI.*

*This report will summarise the methods currently used to source data for training AI along with the regulations, standards and protocols that can control the handling and sharing of that data. It will then provide gap analysis on this information to scope possible requirements for standards for ensuring traceability and integrity in the data, associated attributes, information and feedback, as well as the confidentiality of these.*

[***Security Testing of AI***](https://portal.etsi.org/webapp/WorkProgram/Report_WorkItem.asp?WKI_ID=58860)

*The purpose of this work item it to identify objectives, methods and techniques that are appropriate for security testing of AI-based components. The overall goal is to have guidelines for security testing of AI and AI-based components taking into account of the different algorithms of symbolic and subsymbolic AI and addressing relevant threats from the work item “AI threat ontology”.
Security testing of AI has some commonalities with security testing of traditional systems but provides new challenges and requires different approaches, due to (a) significant differences between symbolic and subsymbolic AI and traditional systems that have strong implications on their security and on how to test their security properties, (b) non-determinism since AI-based systems may evolve over time (self-learning systems) and security properties may degrade, (c) test oracle problem, assigning a test verdict is different and more difficult for AI-based systems since not all expected results are known a priori, and (d) data-driven algorithms: in contrast to traditional systems, (training) data forms the behaviour of subsymbolic AI.
The scope of this work item is to cover the following topics (but not limited to):
• security testing approaches for AI
• testing data for AI from a security point of view
• security test oracles for AI
• definition of test adequacy criteria for security testing of AI
• test goals for security attributes of AI
and provide guidelines for security testing of AI taking into account the abovementioned topics. The guidelines will use the results of the work item "AI Threat Ontology" to cover relevant threats for AI through security testing, and will also address challenges and limitations when testing AI-based system.
The work items starts with a state-of-the-art and gap analysis to identify what is currently possible in the area of security testing of AI and what are the limitations. The works will be coordinated with TC MTS.*

The ISG is also discussing adoption of a work item on:

***Securing AI Problem Statement***

*This work will define and prioritise potential AI threats along with recommended actions.*

ETSI ISG SAI believes that this work will be of interest to many other technical standards groups and looks forward to engaging with such groups.