Artificial Intelligence Threat Reporting & Incident Response System

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

Artificial Intelligence Threat Reporting & Incident Response System

IRIS

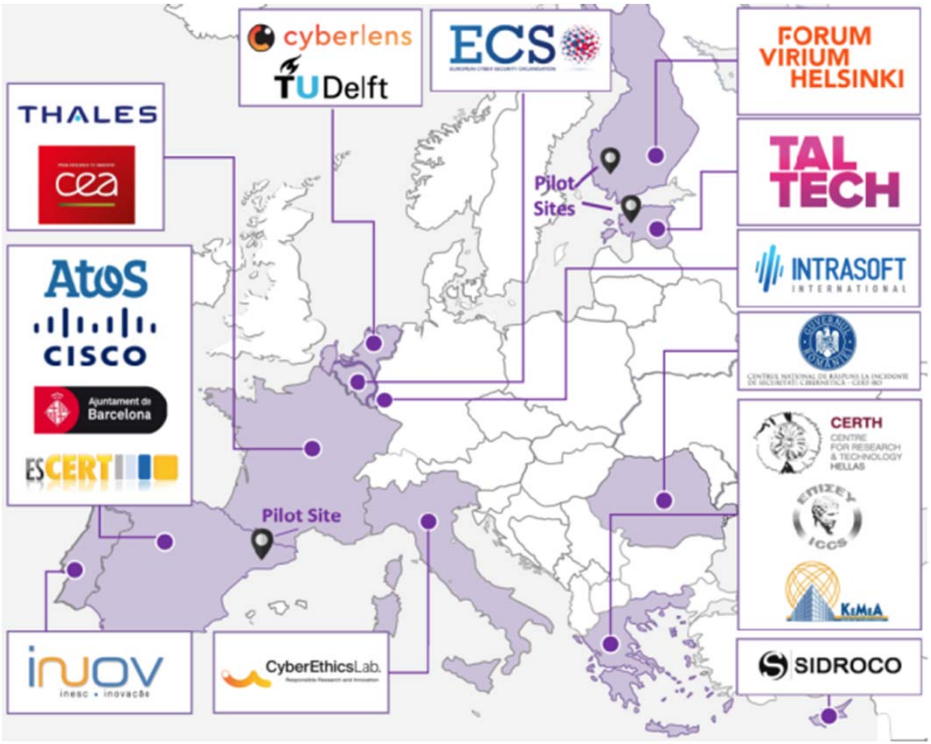
A collaborative CERT/CSIRT platform to combat cyber-threats in **IoT and AI-driven systems**

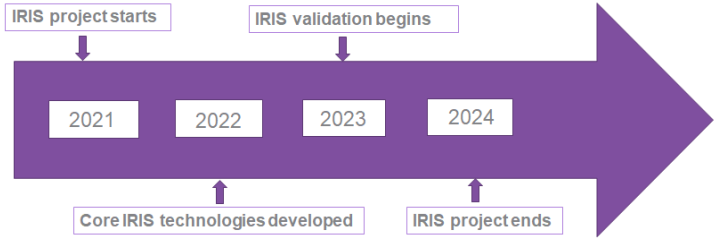
Duarte Nascimento (INOV)

PUC 2 | 27/03/2024

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

Project at a Glance





Call Identifier: 2020-SU-DS-2020

Topic: SU-DS02-2020 Intelligent security and privacy management

EC Funding: 4 918 790.00 EUR

Duration: 36 months (Sept 2021-Aug 2024)

Consortium: 19 partners

Coordinator: INOV - Instituto de Engenharia de Sistemas e Computadores, Inovacão, (INOV), Portugal

Learn More: www. iris-h2020.eu

Join us: @iris-h2020

IRIS H2020 Project

Consortium

- **6 Public organizations**
- **3 SMEs**
- **4 Large ICT industries**
- **6 Research institutions & Universities**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

IRIS Motivation

As existing and emerging **SMART CITIES** continue to **expand their IoT and AI-enabled** systems, **novel and complex threats are introduced**.

Architecture and behaviour of emerging IoT and AI technologies are **not currently well understood** by security practitioners, such as CERTs and CSIRTs.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

IRIS Vision

The **H2020 IRIS project** aims to deliver a framework that will support European CERT and CSIRT networks detecting, sharing, responding and recovering from **cybersecurity threats and vulnerabilities of IoT and AI-driven systems.**

Complement the existing MeliCERTes open platform and tools.

The **IRIS Platform** will be made available, to the European national CERT and CSIRTs, by the end of the project.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

Key Takeaways

- Smart Cities => **novel**, cutting edge AI/IoT-driven technology
- This implies **Emerging Threats** ! High risks!

- Currently, **lack of experience as well as of tools** for incident management that tackle IoT & AI attack vectors
- **IRIS** will enhance the capabilities (knowledge, toolset, training) of CERTs/CSIRTs, to address these challenges.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

What to expect from this Pilot

1) Knowledge of advanced tools to address IoT and AI threats, in Smart Cities.

2) **YOUR FEEDBACK MATTERS !**
   - **What are your drivers, needs, and pain points?**
   - **Are there social acceptance benefits or issues?**

3) Join the **Stakeholder Community**, and influence the IRIS roadmap.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

Agenda

Timeslot (duration)	Session	Partner/ Facilitator
10:00 – 10:05 (5 minutes)	**Introduction to IRIS**	**INOV**
10:05 – 10:15 (10 minutes)	**Pilot Use Case: Introduction**	**TALTECH**
10:15 – 10:20 (5 minutes)	**Societal Acceptance of Technology (SAT)**	**CEL**
10:20 – 10:25 (5 minutes)	**Pilot Use Case: Relevance for Transportation**	**KEMEA**
10:25 – 10:50 (25 minutes)	**Demonstration**	**TALTECH**
	****** Social Acceptance assessment ********	
10:50 – 11:00 (10 minutes)	**Survey**	**CEL**
11:00 – 11:10 (10 minutes)	**Discussion Group**	**CEL**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

Artificial Intelligence Threat Reporting & Incident Response System

IRIS PUC 2 – Introduction

PUC 2 | 27/03/2024 | TALTECH

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

IRIS – Pilot Use Case 2 Tallinn

- Aim:
  - Evaluate the capability of the IRIS Platform for cyber incident response in AI and IoT environments.
  - Assess the enhanced ability, delivered by IRIS, of System Administrators and CERTs/CSIRTs to work collaboratively in cyber incident response.
  - Enhance the security of AI-based systems

- Benefits:
  - Enhanced capability for cyber incident response including cooperation with CERTs/CSIRTs
  - Improved security for AI based Systems

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

IRIS – Pilot Use Case 2 Tallinn

- Tallinn Pilot Use Case focuses on evaluating selected tools from the IRIS platform (Nightwatch (Probe and RRR), ATIO, SiHoneyPot (Indirectly), DPA, CTI(Indirectly)) to support cyber attack detection and incident response in an AI-based system and supporting infrastructure.

- Within Tallinn, the AI-based system focusses on Autonomous Transportation, mainly systems and data of an autonomous vehicle shuttle and smart city planning.

- The Tallinn Pilot Use Case that will be demonstrated today focusses on a cyber attack scenario where telemetry from the autonomous vehicle is manipulated to affect the integrity of decision-making of the Urban Operating Platform (UoP), a system used for transportation planning.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

IRIS – Pilot Use Case 2 Tallinn

- The aim of the Pilot is to present to relevant stakeholders the IRIS platform, demonstrate its capability and effectiveness and engage with you in feedback.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

User story – Malformed Data Injection

Threat actor aims to sniff network traffic of communication between autonomous vehicle shuttle and the smart city management platform (UoP) to then manipulate the telemetry of the autonomous vehicle to cause the UoP to make incorrect city and transportation planning decisions.

Transportation System Administrator keeps track of the infrastructure operational environment.

SiHoneyPot is a honeypot which uses synthetic data of the LiDAR sensor to replicate its telemetry data footprint. The LiDAR data is streamed to the UoP.

Nightwatch probe extension will receive the autonomous vehicle telemetry data which is being fed to the UoP. The probe will parse the data to the **Nightwatch Cortex** to detect anomalies in the data. A report is then sent to the ATIO.

ATIO will enrich the threat information and manage the workflow for cyber incident response with the RRR and EME.

RRR will enable options for risk-based recovery and self-healing.

EME Dashboard will enable visualisation of the threat information.

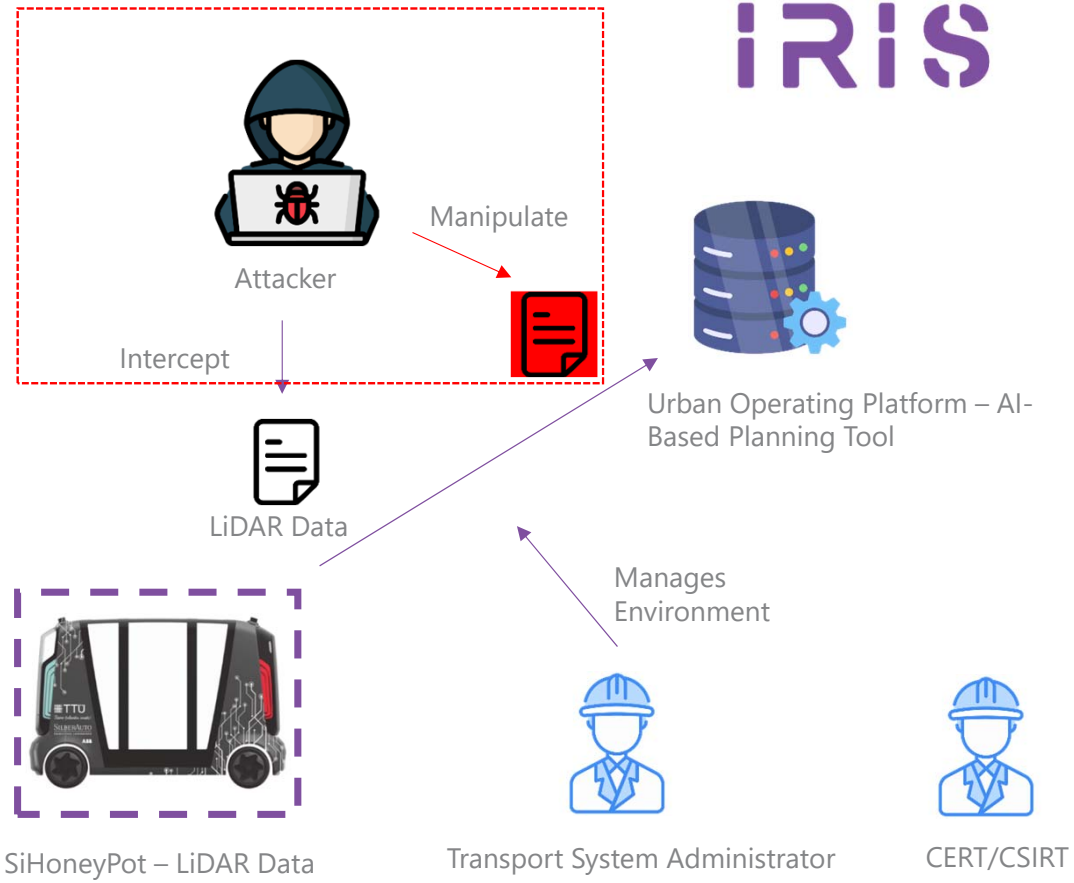
CERTs/CSIRTS will interact with the Transportation System Administrator to manage the cyber incident response.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

User story PUC 2

- The goal of IRIS PUC 2 is to evaluate IRIS platform tools for cyber incident response and recovery for an autonomous transportation case study.

- We will demonstrate:
  - **Detection of Attack that seeks to manipulate the telemetry data of an AV which is used as an input to a smart city planning system.**
  - **Workflow of cyber incident response and recovery, demonstrating capabilities for threat information enrichment and sharing.**



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

User Story: Why is this user story important?

- **Integrity of AI-based systems**: Threat actors may compromise the integrity and security of AI-based systems which affect the decision-making of the AI system and impact users of the system and citizens.

- **Data Breach**: Sensitive data may be sniffed by an attacker which can use the data to craft further attacks and/or reverse engineer the AI model.

- **Device Manipulation**: Attackers might tamper with the underlying infrastructure which supports the AI-based system.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

User story – Data Protection and Accountability

User deletes a workflow

ATIO produces log, adds orchestration workflow metadata and posts log.

DPA receives information from ATIO and stores an off-chain database of logs and uses distributed ledger technology to ensure the integrity and authenticity of the stored CTI data.

Auditors will request audit log from the DPA

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

User Story: Why is this user story important?

- **Integrity and Authenticity of CTI Information**: Threat sharing information and information on the workflow of cyber incident response systems require audit and traceability to assure the integrity and authenticity of information.

- **Innovative Technology**: The use of innovative distributed ledger technology to enhance the integrity of the CTI workflow.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

Artificial Intelligence Threat Reporting & Incident Response System

IRIS PUC 2 – General Context

PUC 2 | 27/03/2024 | KEMEA

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

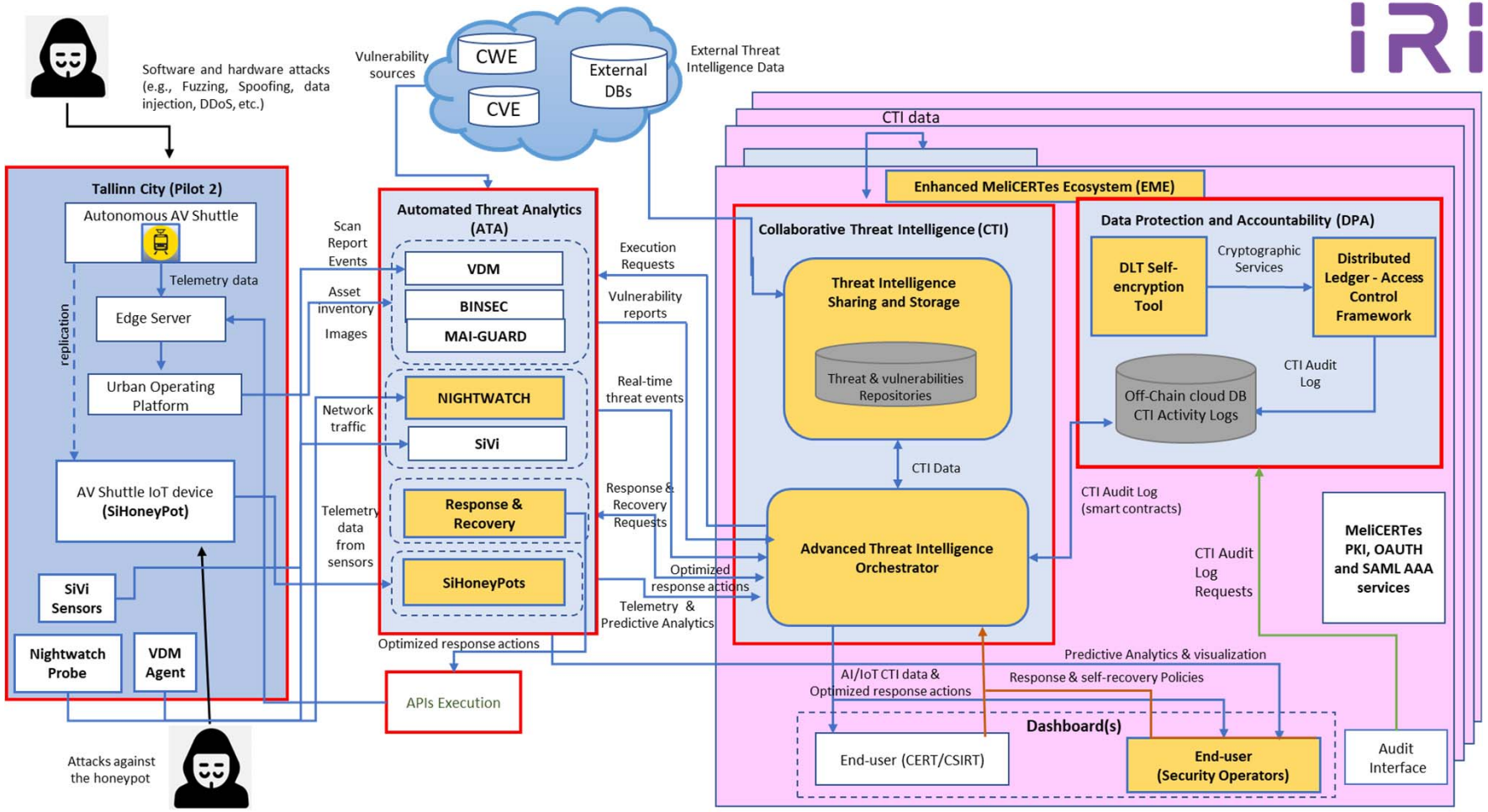
IRIS Pilots - Key Objectives

The demonstration pilots of IRIS target to:

- Secure the **smart city's IoT and control systems** against confidentiality breaches, detecting interception and compromise of IoT systems and reporting efficiently the breach (PUC1 in Barcelona).
- Securing a **smart city's AI-enabled infrastructure of transport systems** (autonomous buses), illustrating IRIS capabilities to identify the threat, self-recover and share the threat intelligence with other system operators and platforms (PUC2 in Tallin).
- Protect the **customer facing components of the smart grid** against threats to control functions defined for the control of the demand, demonstrating IRIS capabilities to manage a cross-border crisis exercise on the Virtual Cyber Range (PUC3 in Helsinki/Tallin).
- IRIS platform focuses on:
  - ✓ Real-time estimation of cybersecurity and privacy risks
  - ✓ Automated threat detection, response and recovery
  - ✓ Threat intelligence sharing
  - ✓ Measures for effective cybersecurity incidents and crises management

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

IRIS General Architecture



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

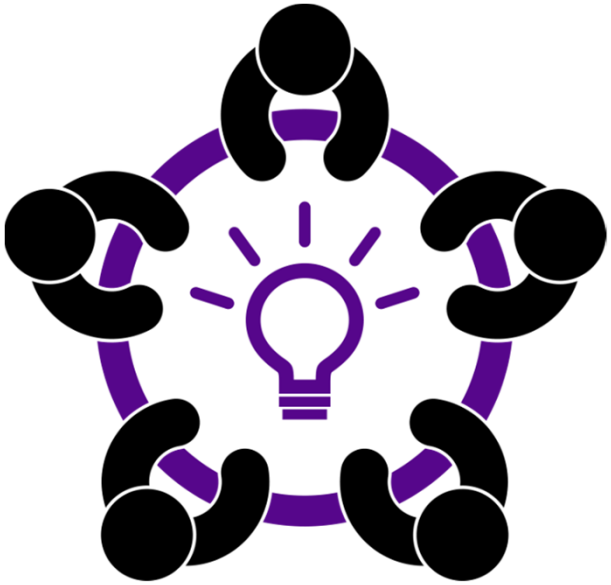
Artificial Intelligence Threat Reporting & Incident Response System

IRIS PUC 2 – SAT

PUC 2 | 27/03/2024 | CEL

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

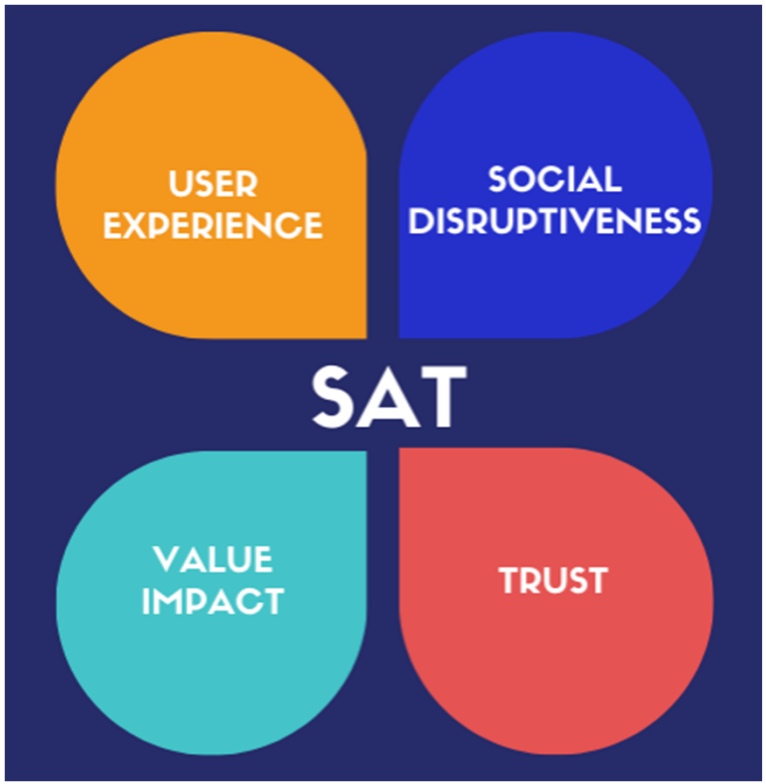
Social Acceptance Assessment



- ✔ This session is also part of our research on the social acceptance of IRIS project
- ✔ The aim is to get a feedback on the presentation based on your perceptions
- ✔ Following the pilot use case presentation we will invite you to fill in a survey

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

Social Acceptance of Technology
IRIS model



BASED ON 3 PRINCIPLES

- ✔ **Human centredness,**
- ✔ **Leave no one behind**
- ✔ **Ethics-by-design**

WHO: CERTs/Csirts, Transport sector stakeholders to assess the IRIS technological solution

WHAT: perceptions on the IRIS Platform.

HOW: demos, survey, focus groups to assess different dimensions of social acceptance

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

Artificial Intelligence Threat Reporting & Incident Response System

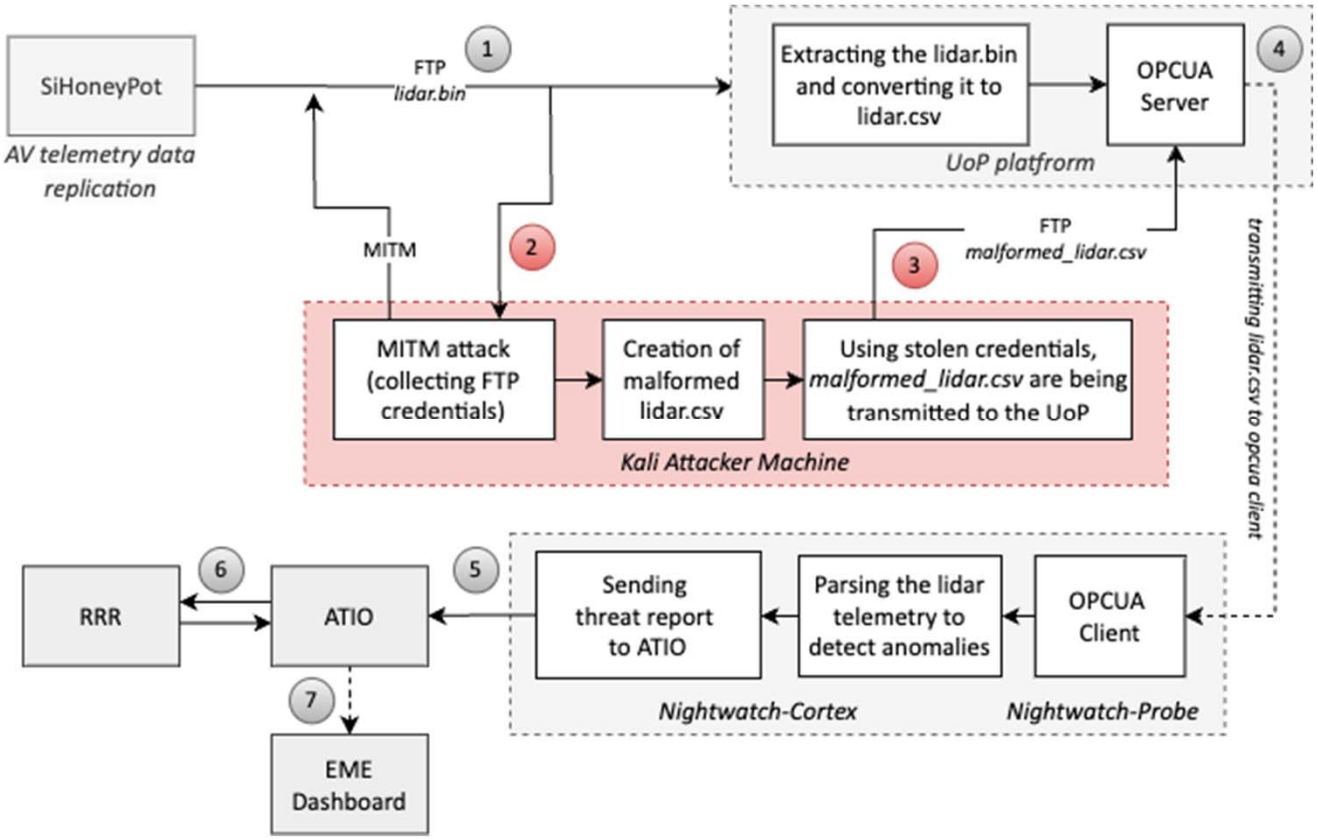
IRIS Platform Demonstration

PUC 2 | 27/03/2024 | CLS, SID, ICCS, INTRA, INOV

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

Attack Detection Workflow

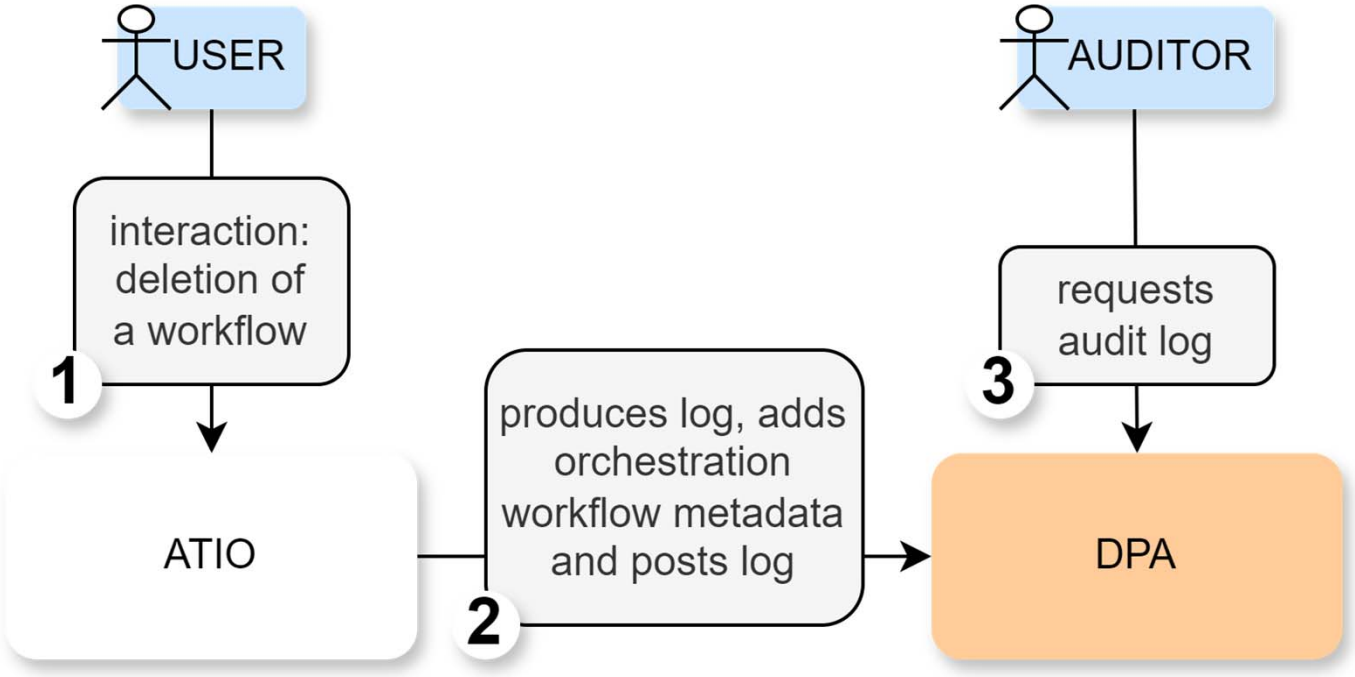
Trigger: malformed data injection attack



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

Critical Change Workflow

Trigger: deletion of a critical ATIO workflow



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

Social Acceptance Survey

Direct Link or QR code to Google Form

https://forms.gle/Kqz7h5TUPjagG52Y6

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 101021727.
This material reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.