TALON

Autonomous and Self-organized Artificial Intelligent Orchestrator for a Greener Industry 4.0

Deliverable

D5.4 Final TALON Platform Setup, Operation, Continuous Integration & Maintenance Report

Actual submission date: 03/10/2025

Funded by the European Union

Project Number:	101070181		
Project Acronym:	TALON		
Project Title:	Autonomous and Self-organized Artificial Intelligent Orchestrator for a Greener Industry 4.0		
Start date:	October 1st, 2022	**Duration:**	36 months
D5.2 Initial TALON Platform Setup, Operation, Continuous Integration & Maintenance Report			
Work Package:	WP5		
Lead partner:	INTRA		
Author(s):	Magda Foti (INTRA), All technical, All pilot partners		
Reviewers:	Sophia Karagiorgou (UBI) Nikolaos Ntampakis (MINDS)		
Due date:	30/09/2025		
Deliverable Type:	R	**Dissemination Level:**	PU
Version number:	1.0		

Revision History

Version	Date	Author	Description
0.1	19/02/2025	INTRA	Table of Contents release
0.2	29/07/2025	Technical Partners	Content added in Section 2
0.3	16/09/2025	INTRA	Content added in Section 5
0.4	17/09/2025	UBI	Content added in Section 3
0.5	18/09/2025	Demo Partners	Content added in Section 4
0.6	19/09/2025	INTRA	Version ready for review
0.7	23/09/2025	MINDS	Review
0.8	23/09/2025	UBI	Review
0.9	27/09/2025	INTRA	Review comments addressed. Version ready for quality check.
1.0	03/10/2025	ENG	Final coordinator review before submission

Table of Contents

List of Figures

List of Tables

Definitions and acronyms

AI	*Artificial Intelligence*
API	*Application Programming Interface*
AR/VR	*Augemented Reality / Virtual Reality*
CA	*Consortium Agreement*
C2C	*Command and Control Center*
CD	*Continuous Deployment*
CI	*Contrinuous Integration*
CNC	*Computer Numerical Control*
DoA	*Description of Action*
EC	*European Commission*
EU	*European Union*
GA	*Grant Agreement*
JSON	*JavaScript Object Notation*
KPI	*Key Performance Indicator*
ML	*Machine Learning*
MS	*Milestone*
NG-SDN	*Next Generation – Software Defined Network*
OEE	*Overall Equipment Effectiveness*
PC	*Project Coordinator*
PPE	*Personal Protection Equipment*
PLC	*Programmable Logic Controller*
SLOs	*Service Level Objectives*
TC	*Technical Coordinator*
TrL	*Trust Level*
UC	*Use Case*
UI	*User Interface*
UATV	*Unmanned Aerial Tactical Vehicles*
UxV	*Unmanned aerial/ground/underwater Vehicle*
WP	*Work Package*
XAI	*eXplainable Artificial Intelligence*

Disclaimer

This document has been produced in the context of TALON Project. The TALON project is part of the European Community's Horizon Europe Program for research and development and is as such funded by the European Commission. All information in this document is provided 'as is' and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability with respect to this document, which is merely representing the authors' view.

Executive Summary

The objective of TALON is to design, develop and deploy advanced edge-to-cloud services that enhance the performance, adaptability, explainability, trustworthiness, and transparency of Industry 5.0 applications through the instantiation of fully automated AI approach both from a business and technical viewpoint. The approach we adopted infuses intelligence at the edge in a manner that is flexible, adaptable, and efficient in terms of energy and data usage, while keeping compute-heavy processing task for the cloud. To achieve this, TALON develops a suite of interoperable, human-centric components for AI-driven orchestration spanning the cloud–edge continuum, integrating them into real-life applications and demonstrations. The successful delivery of these demonstrations requires the integration of diverse, heterogeneous technologies along with a structured evaluation methodology that addresses technical, integration, and user-centric considerations.

The key objective of the deliverable 'D5.4 - Final TALON Platform Setup, Operation, Continuous Integration & Maintenance Report', is to offer a comprehensive final report on the TALON platform and the usage of different TALON components across the demonstration sites and pilots upon project completion. The deliverable encapsulates platform and component-level technical verification (including unit tests, quality, and vulnerability reporting), presents the Integration Traceability Matrix, and conveys the final Key Performance Indicator (KPI) evaluations as well as end-user evaluations on the usage of the platform for the four pilot cases. It is noteworthy to mention that the different pilot sites, UC1: Automatic UATV Coordination; UC2: I5.0 Automation & Planning; UC3: AR/VR for Training & Maintenance; and UC4: Human–Robot Collaboration, have utilised diverse functionalities from the TALON components and the TALON platform as a whole. Besides some deployments have been performed directly at pilot sites and their on-premise edge-to-cloud infrastructure, while some other have been facilitated through the centralised infrastructure and populated via the TALON Dashboard. Finally, the deliverable reports the instances where KPI targets were achieved or partially achieved, documents remaining issues and corresponding mitigations, and captures lessons learned alongside recommendations to facilitate improvements and real-world implementation.

1 Introduction

1.1 Objective of the Deliverable

The primary objective of Deliverable D5.4 is to provide the final, consolidated evaluation and record of the TALON platform's setup, operation, integration and maintenance up to project completion. The deliverable documents the final technical verification of the platform and its components (unit testing, quality and vulnerability reporting), presents the Integration Traceability Matrix and integration-testing evidence, reports the final KPI measurements and end-user evaluation results for each pilot, and captures residual issues with their mitigation status plus lessons learned and concise recommendations to support further refinement and real-world deployment.

1.2 Relation to other Work Packages

This is the fourth and final deliverable of "WP5 - Integration, Validation and Demonstration" series of deliverables. The work reported in this document is part of demonstrators' tasks, namely: Task 5.3: Automatic UATV Coordination, Task 5.4: I5.0 Automation & Planning, Task 5.5: AR/VR for Training & Maintenance and Task 5.6: Human Robot Collaboration. These tasks, and the output of deliverable D5.4, rely on input from D2.1, D3.1, D3.2, D3.3 and D4.1.

Completing D5.4 is part of achieving Milestone 8 (MS8) – Final release of TALON Integrated platform and the current deliverable marks the end of the WP5 Tasks, namely Task 5.3, Task 5.4, Task 5.5 and Task 5.6.

1.3 Structure of the Document

This document is structured in five (5) chapters:

Chapter 1 introduces the deliverable objectives and positioning within the project.

Chapter 2, titled "Technical Evaluation of TALON Components", delves into the evaluation of each TALON component. To this end, the results of unit testing per component are presented, as well as quality reporting and vulnerability reporting per component, presenting in table format the different results achieved for each of them.

Chapter 3, titled "Integration", documents the integration steps that were needed for the different microservices applications inside the platform. Integration testing has been performed to assess inter-process communication across different components of the TALON platform. The traceability matrix is being presented, which maps each of the components that were integrated to one another, as well as describing the integration testing that was performed by means of their APIs functional assessment (i.e., the API is up and running according to the technical specifications) and results validation (i.e., the component produces the expected results).

Chapter 4, titled "Final TALON Evaluation Results", reports on the final KPI evaluations as well as the end-user assessments for each one of the 4 pilots. Moreover, an internal survey's results are presented as graphs, in order to perform an overall evaluation of the different TALON aspects, in terms of each process that has been utilized both by technical and non-technical users.

Chapter 5 briefly concludes the deliverable.

2 Technical Evaluation of TALON Components

2.1 Unit Testing

The Unit Testing section documents the automated functional verification performed at component level. It explains the coverage metrics used (Line Rate, Branch Rate, Lines Covered/Valid, Branches Covered/Valid, Functions Covered/Valid), presents coverage snapshots for the 1st and 2nd reporting releases per partner/component, and highlights coverage gaps that require additional test development. For comparison purposes, we include both unit testing performed in the 1st release of each component at M25 and the 2nd and final release. We clarify that components deployed in the 2nd release are only reported in the tables regarding the 2nd release unit testing figures. A detailed explanation of each metric as well as the Unit Testing table, can be found below:

- **Line Rate:** The total number of valid code lines that can be covered.
- **Branch Rate:** The percentage of branches covered.
- **Lines Covered/Valid:** How many of those lines were valid and executed from tests.
- **Branches Covered/Valid:** How many of those branches were valid and executed from tests.
- **Functions Covered/Valid:** How many of those functions were valid and executed from tests.

1st Release		**Lines Valid**	**Lines Covered**	**Line Rate (%)**	**Branches Valid**	**Branches Covered**	**Branch Rate (%)**	**Functions Covered**	**Functions Valid**	**Functions Rate (%)**
MINDS	*Anonymisation*	314	264	84.08	-	-	-	-	-	-
	XAI-trl3-4	563	347	61.63	-	-	-	-	-	-
	Federated learning	193	182	94.30	-	-	-	-	-	-
	XAI	459	372	81.05	-	-	-	-	-	-
SID	*Api*	413	350	84.74	86	56	65.11	76	88	86.36
	App	-	-	90.33	-	-	74.25	-	-	88.6
	Chaincode	-	-	64.9	-	-	-	-	-	-
	Client	171	142	83.04	21	9	42.85	28	34	82.35
UBI	*Talon-angular*	1245	704	56.54	285	46	16.14	227	417	56.54
	Network-intelligence	555	464	84	-	-	-	-	-	-
KU	*Augmentation (Reports_ku)*	2,819	2,661	94	-	-	-	-	-	-

2nd Release		**Lines Valid**	**Lines Covered**	**Line Rate (%)**	**Branches Valid**	**Branches Covered**	**Branch Rate (%)**	**Functions Covered**	**Functions Valid**	**Functions Rate (%)**
MINDS	***Anonymisation***	314	268	85.35	-	-	-	-	-	-
	XAI-trl3-4	563	358	63.58	-	-	-	-	-	-
	Federated learning	193	182	94.30	-	-	-	-	-	-
	XAI	459	384	83.66	-	-	-	-	-	-

SID	*Api*	413	380	85.22	88	76	67.13	96	89	87.36
	App	-	-	90.33	-	-	75.32	-	-	89.6
	Chaincode	-	-	64.9	-	-	-	-	-	-
	Client	171	298	84.02	23	15	49.75	38	34	86.35
UBI	*Talon-angular*	1270	880	69.3	300	110	36.7	290	430	67.4
	Network-intelligence	590	504	87.7	-	-	-	-	-	-
KU	*Augmentation (Reports_ku)*	2,817	2,670	95	-	-	-	-	-	-
ENG	*Smart Policy Manager*	241	*233*	96	-			27	29	97
	Orchestrator-ui	217	*207*	95				8	8	100
UPV	*Few-shot model*	259	212	81.85	-	-	-	2	2	100
TEI	*Image Anonymizer*	1274	930	82	-	-	-	14	14	100
8BELLS	*Smart Pricing Simulator*	345	292	84.6	42	33	78.6	9	9	100

2.2 Quality Reporting

The Quality Reporting section summarises static-analysis and maintainability indicators that describe code health and technical debt. Reported metrics include Reliability Rating, Complexity, Cognitive Complexity, Open Issues, Bugs, Security Rating, Code Smells and Duplicated Lines. For each component the deliverable presents 1st- and 2nd-release values, interprets their impact on maintainability and future evolution, and identifies hotspots where refactoring, simplification or targeted bug resolution is advised. A detailed explanation of each metric as well as the Quality reporting table, are presented below:

- **Reliability Rating -** A numeric score (1 is best, higher is worse) indicating how reliable the code is, based on detected bugs.
- **Complexity** - The cyclomatic complexity, which reflects how "complicated" the code is.
- **Cognitive Complexity** - A metric indicating how difficult the code is to understand, considering nested logic and readability.
- **Open Issues** - The total number of unresolved issues
- **Bugs** - The number of problems identified that could lead to faulty or unexpected behaviour when the software runs.
- **Security Rating -** A measure of how secure your code is, usually ranging from 1 (best) to 5 (worst), based on detected vulnerabilities.
- **Code Smells** - Patterns in the code that aren't necessarily bugs but may indicate maintainability problems or technical debt.

1st Release		Reliability Rating	Security Rating	Complexity	Cognitive Complexity	Open Issues	Bugs	Code Smells	Duplicated Lines
MINDS	*Anonymisation*	5.0	-	387	-	136	4	-	-

	XAI-trl3-4	3.0	-	169	-	16	2	-	-
	Federated learning	3.0	-	61	-	11	5	-	-
	XAI	3.0	-	205	-	35	6	-	-
SID	*Api*	-	-	-	-	-	0	6	19.7
	Client-App	-	-	-	-	-	0	6	19.7
	Dashboard	-	-	-	-	40	0	40	0
UBI	*Talon-angular*	3.0	1.0	39	-	39	22	17	-
	Network-intelligence	1.0	1.0	101	88	47	0	47	-
KU	*Talon (Reports_ku)*	-	-	415	290	-	1	25	21
ENG	*Smart-policy-manager*	-	-	-	-	-	-	11	-
	Orchestrator-ui	-	-	-	-	-	-	8	-

2nd Release		Reliability Rating	Security Rating	Complexity	Cognitive Complexity	Open Issues	Bugs	Code Smells	Duplicated Lines
MINDS	*Anonymisation*	4.0	-	394	-	127	3	-	-
	XAI-trl3-4	**3.0**	**-**	**163**	**-**	**14**	**1**	**-**	**-**
	Federated learning	3.0	**-**	**61**	**-**	**10**	**4**	**-**	**-**
	XAI	**3.0**	**-**	**199**	**-**	**34**	**6**	**-**	**-**
SID	*Api*					**-**	**0**	**3**	**11.2**
	Client-App					**-**	**0**	**3**	**11.2**
	Dashboard					**28**	**0**	**28**	**0**
UBI	**Talon-angular**	2.0	1.0	32	**-**	**20**	**11**	**10**	**-**
	Network-Intelligence	1.0	1.0	80	75	25	0	20	-
KU	*Talon (Reports_ku)*	-	-	5282	5821	-	45	~800	11
ENG	*Resource-manager*	-	-	-	-	-	0	1	-
	Orchestrator-ui	-	-	-	-	-	0	0	-
UPV	*Few-shot model*	2	0	-	-	8	2	6	346
TEI	*Image Anonymizer*	1	3	140	289	142	-	1	34
8BELLS	*Smart Pricing Simulator*	1	3	67	50	-	0	7	12

2.3 Vulnerability Reporting

This section focuses on Vulnerability Reporting, which identifies and quantifies security weaknesses in code and dependencies. Findings are categorised by severity (Low / Medium / High / Critical / Unknown) and reported per component for both reporting snapshots to show remediation progress (both Release 1 and Release 2). The vulnerability tables in the deliverable make it possible to track which components carry the highest risk, document the status of mitigations between snapshots, and prioritise outstanding vulnerabilities so integration and deployment decisions can be informed by concrete risk evidence. The vulnerability reporting table is shown below:

1st Release		Total Vulnerabilities	Low	Medium	High	Critical	Unknown
MINDS	*Anonymisation*	658	250	354	52	2	0
	XAI-trl3-4	309	158	109	29	3	10
	Federated learning	309	158	109	29	3	10
	XAI	309	158	109	29	3	10
SID	*Api*	22	2	20	0	0	0
	Postgres	138	99	27	11	1	0
	Repository	11	4	3	4	0	0
	Talon Client	22	2	20	0	0	0
	Fabric-baseos, Fabric-ca, Fabric-peer, Fabric-ordered	25	24	1	0	0	0
	Talon-explorer	1201	203	439	470	77	12
UBI	*Talon-angular*	2	0	2	0	0	0
	Network-intelligence	4089	865	2244	899	69	12
KU	*Talon (Reports_ku)*	10	6	4	0	0	0
ENG	*Resource-manager*	2	1	0	1	0	0
	Orchestrator-ui	3	2	0	1	0	0

2nd Release		Total Vulnerabilities	Low	Medium	High	Critical	Unknown
MINDS	*Anonymisation*	632	240	344	47	1	0
	XAI-trl3-4	285	146	103	26	2	8
	Federated learning	272	141	97	23	1	10
	XAI	275	154	89	21	1	10
SID	*Api*	20	2	18	0	0	0
	Postgres	135	98	26	11	0	0
	Repository	10	3	4	3	0	0
	Talon Client	20	2	18	0	0	0
	Fabric-baseos, Fabric-ca, Fabric-	19	18	1	0	0	0

	peer, Fabric-ordered						
	Talon-explorer	1195	259	489	371	76	10
UBI	*Talon-angular*	2	0	2	0	0	
	Network-intelligence	3987	967	2245	706	69	
KU	*Talon (Reports_ku)*	41	23	18	0	0	0
ENG	*Resource-manager*	2	1	0	1	0	0
	Orchestrator-ui	3	2	0	1	0	0
UPV	*Few-shot model*	n/a	n/a	n/a	n/a	n/a	n/a
TEI	*Image Anonymizer*	279	210	45	20	2	2
8BELLS	*Smart Pricing Simulator*	1	0	1	0	0	0

3 Integration

In TALON, we performed comprehensive integration testing of the pods and microservices applications within the Kubernetes cluster to validate the interactions and communication between different services. These tests served as a more efficient alternative to traditional end-to-end tests, which are often resource-intensive. Our approach focused on a smaller test scope, requiring the deployment of the set of dependent services. The integration testing was seamless and independent of our GitLab continuous integration/continuous deployment (CI/CD) infrastructure.

Therefore, we deployed the core dependent subset of microservices and pods into the Kubernetes cluster for testing. Integration and interaction validation tests were executed to verify that services can correctly communicate and exchange data, ensuring their APIs and internal logic are functioning as expected.

We present below the updated Integration Traceability Matrix highlighting in X the integration points for inter-pod communication that have been tested.

Table 1. Integration Traceability Matrix.

	Authentication and Authorisation	Data Anonymisation	DLTs for Securing AI/ML models weights	Anomaly Detection	Service Modelling (aka “ Configurations) and Enactment	NG-SDN and Distributed Intelligence	Orchestration	AI Swarm Orchestration	Resource Allocation and Deployment	Definition, Customisation and Monitoring of Metrics	Smart Policy Manager	Data Monitoring, Collection and Aggregation	AI Model Training and SLOs Optimisation	Self-healing and Self-correcting	Hybrid and Optimised Learning	AI Capabilities and Transfer Learning	Data Operations	Digital Twins	XAI, Monitoring and Reporting	Data Lifecycle Management	Visualisation Dashboard
Authentication and Authorisation	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Data Anonymisation	X	X																		X	X
DLTs for Securing AI/ML models weights	X		X																	X	X
Anomaly Detection	X			X																X	X
Service Modelling (aka “Configurations) and Enactment	X				X	X	X	X	X	X	X	X	X								X
NG-SDN and	X				X	X	X	X	X	X	X	X	X								X

Distributed Intelligence																					
Orchestration	X				X	X	X	X	X	X	X	X	X								X
AI Swarm Orchestration	X				X	X	X	X	X	X	X	X	X								X
Resource Allocation and Deployment	X				X	X	X	X	X	X	X	X	X								X
Definition, Customisation and Monitoring of Metrics	X				X	X	X	X	X	X	X	X	X								X
Smart Policy Manager	X				X	X	X	X	X	X	X	X	X								X
Data Monitoring, Collection and Aggregation	X				X	X	X	X	X	X	X	X	X								X
AI Model Training and SLOs Optimisation	X				X	X	X	X	X	X	X	X	X								X
Self-healing and Self-correcting	X													X	X	X	X	X	X	X	X
Hybrid and Optimised Learning	X													X	X	X	X	X	X	X	X
AI Capabilities and Transfer Learning	X													X	X	X	X	X	X	X	X
Data Operations	X													X	X	X	X	X	X	X	X
Digital Twins	X													X	X	X	X	X	X	X	X
XAI, Monitoring and Reporting	X													X	X	X	X	X	X	X	X
Data Lifecycle Management	X	X	X	X										X	X	X	X	X	X	X	X
Visualisation Dashboard	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Integration testing and interoperability validation tests in the TALON platform have focused on dedicated components to ensure seamless and secure collaboration. Analysis of Table 1 shows that only components with direct access to the Authentication and Authorisation and the TALON Dashboard and the Common Polyglot Data Storage (i.e., Data Lifecycle Management) are highlighted in X. The reason is why not all the components interact directly with each other, or some of them interact through the Dashboard. Specifically, Authentication and Authorization has been tested to verify secure access control and identity management across all integrated components, data operations and flows. Also, the Visualisation Dashboard has been tested to confirm coherent and real-time representation of system status, metrics, decisions, and alerts for operators. Last, the component comprising the TALON Orchestrator (i.e., Service Modelling and Enactment, NG-SDN and Distributed Intelligence, Resource Allocation and Deployment, Definition, Customisation and

Monitoring of Metrics, Smart Policy Manager, Data Monitoring, Collection and Aggregation, and AI Model Training and SLOs Optimisation) have been tested to ensure coordinated scaling, and self-healing of workloads across heterogeneous edge and cloud environments.

4 Final TALON Evaluation Results

4.1 Automatic UATV Coordination

Objective

The objective is to orchestrate autonomous Unmanned Aerial Tactical Vehicles (UATVs) to perform real-time surveillance, fire/smoke detection, and situational tracking through intelligent formation flying, managed by an AI-powered orchestrator across edge-cloud infrastructure.

Usage Scenario

Two UATVs fly in formation, scanning and tracking targets (people, cars, fire). They react to sensor triggers (e.g., smoke detectors), dynamically change formation, and send a real-time unified view to the Command & Control Center (C2C). Coordination is achieved via:

- Local GCS (Ground Control Station)
- Cloud platform (Initial plan was the online orchestrator – now we use the local orchestrator on the GCS – assisted by the cloud for heave operations)
- Vision AI platform

AS-IS Architecture (Current Operational State)

Main Characteristics:

- UATVs operated manually or semi-autonomously.
- Each drone has basic AI perception but minimal inter-communication.
- No centralized orchestrator – decisions are either pre-coded or human-in-the-loop.
- High latency due to data sent to remote servers/clouds for AI inference.
- Energy usage is suboptimal due to redundant sensing and uncoordinated flight.

```
[Drone 1]      [Drone 2]
   |              |
[ Basic AI ]   [ Basic AI ]
   \              /
    \            /
   [ Manual GCS ]
        |
 [ Cloud Services ]
   (Analytics, Logging)
```

Figure 1: Use case 1 AS-IS Diagram

TO-BE Architecture (TALON-Orchestrated Coordination)

Enhanced Features:

- Full AI-based orchestration across edge (GCS) and cloud.
- Swarm control (via local control center and AiRFLOW), enabling synchronized formations.

- Realtime vision AI recognition of humans, vehicles, fire/smoke.
- AI Agent for dynamically plans, navigates, and informs of anomalies.
- Efficient energy/data use via compression, streaming in JSON.
- Reduced latency (<150ms) with decision loops closed at edge (GCS).

```
[UATV 1]      [UATV 2]
   |              |
[Edge AI]     [Edge AI]
   |              |
   ---> [GCS / TALOS] <---
          |    ↑    |
          v    |    v
   [ALiSA AI Agent]   [VORTEX AI]
          |    ↓    |
   [Orchestrator] — [Mission Control Center (MCC)]
         ↕            ↕
     [AiRFLOW]    [Dragonfly]
         ↕            ↕
       [Teal Cloud Services]
               |
      [C2C Command View]
```

Figure 2: Use case 1 TO-BE Diagram

Technical Evaluation

System Components

- GCS: Real-time edge inference, formation control.
- Orchestrator: RAM/CPU allocation, telemetry optimization, power analytics.
- Vision AI: Object detection (fire, human, vehicle), alert triggering.
- AI agent optimizer: Adapts to individual drone behavior, manages anomalies.
- Local Control Center: Supports any UxV, enables multi-device orchestration.
- AiRFLOW: Swarm control, environment adaptation.

Data Format

- **Input**: JSON stream over edge/cloud telemetry.
- **Output**: Structured visual logs, statistical reports.

End-User Evaluation

Benefits to Operators

- Low Cognitive Load: Swarm behavior reduces need for micromanagement.
- Real-Time Awareness: Unified C2C dashboard with actionable detections.
- Resilience: Semi-autonomous fallback in case of signal degradation.
- Mobility: Ground teams benefit from quick repositioning of UATVs.
- Trustworthy AI: Transparent decision logic via XAI layer (TALON's vision).

End-User Feedback

- High usability reported for the MCC interface.
- Operators noted significant latency reduction compared to previous non-AI systems.
- Positive feedback on energy-saving mechanisms and auto-recalibration.

Recommendations / Notes by PROBO

- Extend swarm size from 2 - 6 UATVs to maximize coverage.
- Enhance for multimodal input (e.g., thermal + RGB fusion).
- Integrate real-time anomaly explanation with XAI to build operator trust.
- Prepare for CBRN scenarios by incorporating gas sensors into the UATV payload.

4.1.1 Key Performance Indicators' Final Evaluation

KPI_01: Latency Reduction

Objective: Measure the total round-trip communication time between UATVs and coordination nodes (GCS, Cloud).

Data Collection:

- A **network probing tool** was deployed on both the **UATV onboard system** and the **GCS (TALOS)**.
- For each message sent, **T1** (send timestamp) and **T2** (receive timestamp) were logged at millisecond precision.

Analysis:

- A time-series log of |T2 - T1| was generated over 500 iterations for each drone.
- Outliers and packet drops were filtered using a moving average filter.

Result Visualization:

- **Line graph** showing latency over time.
- **Histogram** of latency distribution pre- and post-optimization.

Conclusion:

- **Baseline**: 50ms
- **Final Result**: 25ms (50% latency reduction)
- Achieved via **edge inference** and local pipeline optimizations in TALOS.

KPI_02: UATV-to-Node Communication Latency

Objective: Specifically measure the latency between UATV and GCS/Cloud during telemetry streaming and command acknowledgment.

Data Collection:

- Utilized the same probing infrastructure but scoped to **low-level telemetry/control signals**.
- Probed via RESTful API call timestamps and socket-level ACK logs.

Analysis:

- Time deltas between command dispatch and reception at drone.
- Benchmarking done under controlled scenarios (e.g., static UATV hover vs. in-flight).

Result Visualization:

- **Dual-axis line chart** showing latency vs. packet size over network quality.
- **Bar graph** comparing latency with and without AI Orchestrator.

Conclusion:

- **Baseline**: 20ms
- **Final Result**: 8ms
- Major improvement came from **direct OS pipeline communication** and **reduced serialization/deserialization overhead**.

KPI_03: Energy and Data Efficiency

Objective: Measure reduction in transmitted data and associated power savings due to compression, filtering, and smart streaming.

Data Collection:

- Each UATV recorded the size of telemetry/data streams per second in JSON format.
- Data was categorized by type (image frames, detections, metadata).

Analysis:

- Before and after implementation of TALON's smart compression (via Dragonfly).
- Tracked via internal counters and periodic log snapshots stored in Teal Cloud.

Result Visualization:

- **Stacked bar chart** of data type volumes over time.
- **Pie chart** comparing pre/post average transmission sizes.

Conclusion:

- **Baseline**: 0.5KB per message
- **Final Result**: ~16 bytes per message
- Achieved through **adaptive frequency control**, **filtering irrelevant detections**, and **aggressive compression**.

KPI_04: Energy Efficiency on Operated Flights

Objective: Measure average power consumption of UATVs during missions.

Data Collection:

- Used **on-chip voltage and current sensors (VCC/A)** with 100ms sampling interval.
- Collected telemetry over multiple missions (>10 min flights).

Analysis:

- Energy = $\sum$(Voltage × Current × Δt)
- Compared between baseline (no optimization) vs. TALON-enhanced orchestration (auto-flight, reduced telemetry).

Result Visualization:

- **Box plots** of power consumption per flight stage.
- **Overlay graphs** of energy over mission time with and without orchestration.

Conclusion:

- **Baseline**: 30+ Watts
- **Final Result**: ~20 Watts average
- ~33% reduction due to **minimized idle time**, **shorter optimized paths**, and **reduced telemetry chatter**.

KPI_05: AI – Valid Detections

Objective: Measure object detection accuracy in real scenarios (fire, people, vehicles).

Data Collection:

- Each detection was compared with ground truth (manually annotated video frames).
- Detection logs were cross verified with real event occurrences using timestamped video.

Analysis:

- Calculated standard precision, recall, and F1 score.
- Considered TP (true positives), FP (false positives), FN (false negatives).

Result Visualization:

- **Confusion matrix**
- **PR (Precision-Recall) curve**
- **ROC curve**

Conclusion:

- **Baseline**: 80%
- **Final Result**: 94.7%
- High accuracy achieved via **custom-trained VORTEX AI models**, optimized for edge inference.

KPI_06: Valid Actions (Triggers) from AI

Objective: Measure how many AI-generated alerts correspond to predefined valid actions (e.g., send alert, change path, hover).

Data Collection:

- AI-generated actions logged and tagged with a UID.
- Logs cross-referenced with predefined action dictionary.

Analysis:

- Counted correct matches vs. total actions triggered.
- Manual validation for 10% of samples for QA.

Result Visualization:

- **Bar chart** of action types vs. frequency.
- **Action validity ratio line graph** over time.

Conclusion:

- **Baseline**: 190 (no filtering)
- **Final Result**: 102 valid actions (from ~110 total)
- Filtering of redundant or non-contextual actions improved **energy efficiency and human interpretability**.

Key performance indicators							
ID	**Name**	**Reference to mentioned use case objectives**	**How will be measured / assessed?**	**Baseline Value**	**Target Value**	**Early Results**	**Final Results**
KPI_01	Latency reduction	Latency reduction	Network probing tool, measured by $\|T2 - T1\|$	50ms	24ms	25ms	25ms

KPI_02	UATV-to-Node communication latency	UATV-to-Node communication latency	Network probing tool, measured by $\lvert T2 - T1 \rvert$	20ms	≤10ms	8ms	8ms
KPI_03	EE and Data Efficiency	Energy Efficiency; and Data Efficiency	Data rate measure tool, measured by $\frac{\Delta size_{before}}{\Delta size_{after}}$	A few KB	A few bytes	16B	16B
KPI_04	EE on operated flights	Energy Efficiency	On-chip VCC/A measurements, measured by $\lvert E2 - E1 \rvert$	30+ Watt	About 20 Watt	20Wattt	20Watt
KPI_05	AI – Valid Detections	Accuracy	Measured based on actual objects on screen	80%t	95%	~94.7%	94.7%
KPI_06	Valid Actions (Triggers) from AI	Data & Energy Efficiency	Match against original defined actions	190	100	~102	102

4.2 I5.0 Automation & Planning

Objective

The main objective is to evolve industrial manufacturing toward the zero-defect paradigm and Industry 5.0 by integrating explainable, self-healing, and self-optimising AI systems into production environments. Specifically, by studying and optimising the processes and workplans of industrial manufacturing, it is possible to not only considerably reduce the costs of the company's resources related to the avoidance or even treatment of defective products, but also the refinement of the overall production chain regarding efficiency, safety, adaptability, and flexibility. To achieve this, TALON will evolve the SotA manufacturing line into a qualification system based on a continuous process validation sustained on AI to certificate its manufacturing quality level and guarantee the inalterability of product and process data to its customers.

Usage Scenario

The scenario focuses on transforming current manufacturing workflows into a self-optimised and self-healing environment powered by edge AI. Today, machines like FACTOR's Nakamura 2 lathe rely on periodic quality checks, meaning that if one defective part is detected, entire batches may be scrapped, generating significant waste. In the TALON's AI-enabled system, data from CNC machines are continuously collected and analysed to identify patterns of potential faults in real time, trigger corrective actions automatically, and adapt machine parameters before defects occur. This scenario demonstrates how automation, predictive quality assurance, and human-in-the-loop oversight can drastically reduce scrap, improve availability, and optimise overall efficiency.

AS-IS Architecture (Current Operational State)

At FACTOR's plant, raw material arrives and is stored in the CNC machine (the Nakamura 2) until needed (cf. AS-IS Diagram). Engineers manually configure the machine's operation parameters depending on the part to be manufactured. During machining, the machine and shopfloor are monitored through multiple data sources: (i) the CNC machine itself, which provides temperatures of the engines and alarms through the MTLINKi software; (ii) external sensors measuring factors like turret power consumption, room temperature, humidity, and lubricant temperature; and (iii) quality control data from periodic inspections. However, quality checks are only performed periodically (for example, one part per hour). If the inspected part shows defects, all parts produced during that interval must be re-checked or discarded, generating large amounts of scrap. If the inspected part passes, production continues unchanged, with only minor manual adjustments if deviations are noticed. This periodic, reactive approach means that defects can accumulate unnoticed for long stretches, leading to inefficiency, waste, and higher costs

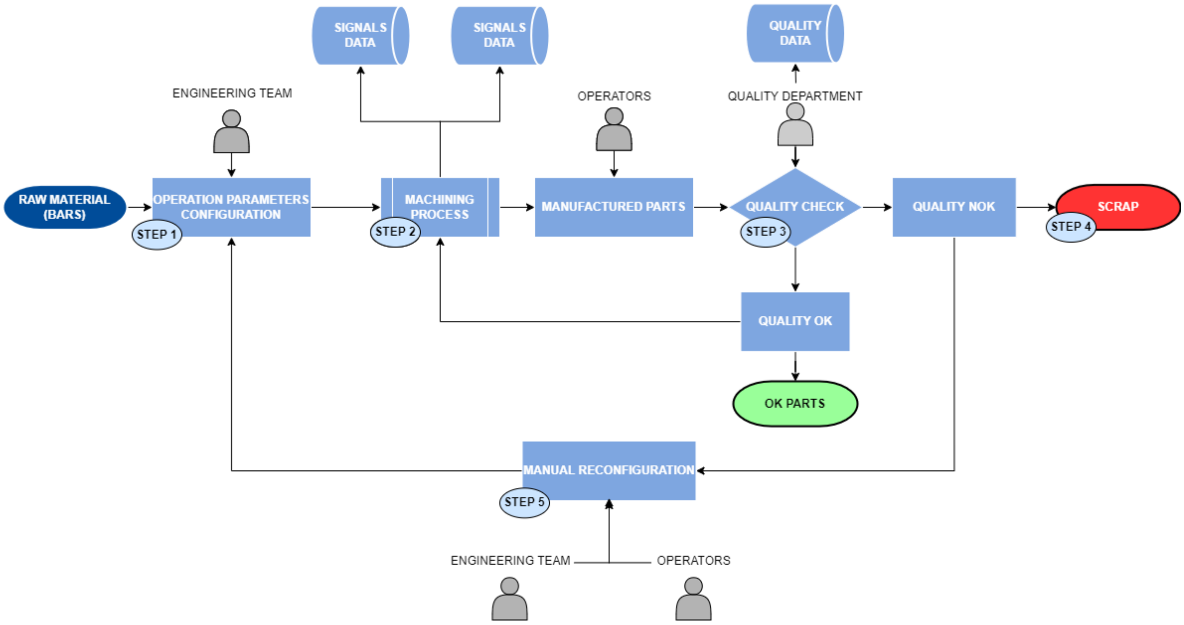

Figure 3. Use Case 2 AS-IS Diagram

TO-BE Architecture (TALON-Orchestrated Coordination)

With TALON, the manufacturing line shifts from periodic quality control to continuous monitoring and proactive optimisation (cf. TO-BE Diagram). Instead of waiting for hourly inspections, data from the CNC machine (Nakamura 2) and external sensors are streamed in real time to edge AI systems. The AI algorithms detect anomalies early, identify patterns that could lead to faults, and trigger corrective actions, such as adjusting machining parameters before defects occur.

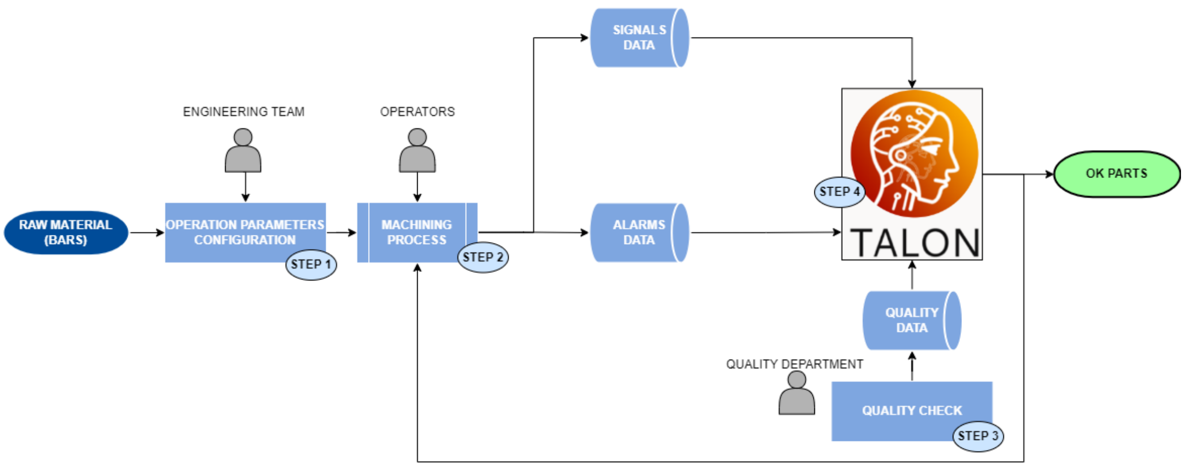

Figure 4. Use Case 2 TO-BE Diagram

Technical Evaluation

System components:

Kubernetes Cluster: Orchestration of microservices across edge–cloud continuum, enabling self-healing and self-correcting mechanisms.
Monitoring & Alerting (Prometheus + Grafana): Collects real-time metrics (CPU, RAM, latency) and triggers alerts based on predefined thresholds.
Node-RED: Executes automated corrective actions (scale, cordon/uncordon, pod rescheduling) through Kubernetes API, ensuring closed-loop resilience.
Few-Shot Learning algorithm: Implements hybrid learning approaches to reduce data collection needs and enable efficient predictive models for tool wear and OEE optimisation.

Data format:

Input: CNC machine nakamura 2 telemetry and process data
Cluster telemetry: infrastructure metrics (CPU, RAM, disk, latencies, pod/node health) collected in real time from the servers.
Output:

- JSON-based alerts and orchestration logs (from Node-RED flows).
- Time-series databases (Prometheus/Timescale).
- Predictive models and reports (tool wear estimation, OEE improvement).

End-User Evaluation

Benefits to operators:

- **Early fault detection** → fewer surprises during production, reducing stress and workload.
- **Reduced scrap** → less rework and wasted time dealing with defective batches.
- **Increased safety** → continuous monitoring reduces risks linked to machine malfunctions.
- **Human-in-the-loop design** → operators remain in control, with smarter assistance.
- **Resilient Data Flows:** Continuous traceability of industrial processes
- **Low Cognitive Load:** Automated orchestration minimises the need for manual cluster management.
- **Real-Time Awareness:** Grafana dashboards with actionable system health indicators.
- **Sustainability:** By maintaining cluster health through self-healing orchestration, the system avoids resource over-provisioning and reduces energy consumption. In parallel, few shot learning model minimise unnecessary tool changes, cutting waste and costs.

End-User feedback:

- Positive usability of the Rancher-Grafana–Node-RED interface for orchestration, monitoring and corrective workflows.
- Clear traceability of corrective actions (alerts, Node-RED workflows, Kubernetes API calls) improved trust in the system.
- Energy efficiency gains were observed, since maintaining cluster health prevented resource overuse and stabilised workloads.

Recommendations / Notes by FACTOR

In the future, FACTOR plans to extend the monitoring and orchestration ecosystem to more CNC machines. We also plan to harvest and collect more monitoring data to improve the model with more precise predictions.

4.2.1 Key Performance Indicators' Final Evaluation

KPI_07: Quality ratio increase

Objective: To improve the proportion of correctly machined parts produced without defects. By optimising tool usage and machining parameters, the system aims to ensure more parts meet quality standards on the first pass, reducing rework and increasing overall product quality.

Data Collection:

The Quality Ratio represents the percentage of produced parts that meet quality standards without requiring rework or being scrapped. This KPI is continuously monitored by the MES, which collects data from inspection systems, quality control stations, and operator inputs.

- How data is collected:
  - Automatic inspection devices (e.g., vision systems, measurement tools) flag non-conforming parts.
  - Manual entries by quality inspectors for rejected or reworked parts.
  - MES aggregates this data in real time and calculates the ratio of good parts to total parts produced.

Conclusion:

- **Baseline**: 98,0 %
- **Final Result**: 99,0%

KPI_08: Reduction of scrap parts

Objective: To minimise the number of defective or unusable parts produced during machining. This is achieved through better tool condition monitoring and predictive maintenance, leading to fewer breakdowns or tool-related issues that cause parts to be discarded.

Data Collection:

When parts are found to be out of tolerance during inspection, either by automated systems or manual checks, they are flagged and sent for further evaluation by the quality department. The quality team then assesses whether each part can be reworked or must be discarded. This decision is recorded in the MES, specifying the outcome—reworked or scrapped—along with details such as the reason for non-conformance, the operator involved, and the associated machine or work order. Only the parts classified as scrap by the quality department are counted toward the scrap KPI, ensuring accurate tracking and traceability of production losses.

Conclusion:

- **Baseline**: 2,2 %
- **Final Result**: 2,0 %

KPI_09: Lower environmental footprint

Objective: To reduce the environmental impact of the machining process by decreasing energy consumption, material waste (e.g., scrap), and tool usage. Extending tool life and minimising rework directly contribute to more sustainable operations.

Data Collection: Number of tool per parts. The quantity of parts that can be manufactured thanks to the optimization on tool consumption is increased by 20%. It generates a 16% reduction in tools consumption. KPI: Tool cost contribution per part has been reduced 16%, so the total environmental impact is reduced, at equal energy and raw material consumption.

Conclusion:

- **Baseline (**Tool cost contribution per part): 1,2 %
- **Final Result**: 1,008% (**16% reduction**).

KPI_10: Increase in availability

Objective: To maximise the uptime of the production equipment. By predicting tool wear and preventing unplanned downtimes, the system ensures that machines are ready and running when needed, improving the availability component of OEE (Overall Equipment Effectiveness).

Data Collection:

Availability measures the proportion of planned production time during which machines are actually running and capable of producing. The MES tracks machine status and production schedules to determine downtime and uptime.

- How data is collected:
    - Real-time signals from machine PLCs or sensors indicating machine status (running, stopped, idle).
    - MES logs unplanned and planned downtimes, including causes (e.g., tool change, maintenance, breakdown).
    - These logs are used to calculate machine availability over defined time periods.

Conclusion:

- **Baseline**: 78 %
- **Final Result**: 80,1 %

KPI_11: Decrease in investigation efforts

Objective: To improve optical network efficiency by minimizing the time and resources required to identify, diagnose, and resolve issues, thereby enhancing system availability, reducing downtime, and lowering overall maintenance costs on the core fibre infrastructure. KPI_11 and KPI_12 have been measured and reported as part of Theodorou et al. scientific publication[1].

Data Collection: The laboratory data collected and made available by the TEI team comprises detailed measurements and diagnostics from the core fibre optical network infrastructure. This data

[1] *Theodorou, G., Karagiorgou, S., Fulignoli, A., & Magri, R. (2024, July). On explaining and reasoning about optical fiber link problems. In World Conference on Explainable Artificial Intelligence (pp. 268-289). Cham: Springer Nature Switzerland.*

includes signal quality metrics, fault detection logs, network performance indicators, and maintenance records, all aimed at facilitating rapid identification and resolution of fibre network issues.

Conclusion:

- **Baseline**: Optical Time-Domain Reflectometer (OTDR) <1 second
- **Final Result**: ML-based detection and reporting ~0.5 seconds

KPI 12: Increase optical fault analysis reliability

Objective: To enhance the accuracy, and reliability of fault detection and diagnosis processes, enabling quicker root cause identification and more effective remediation regarding optical fibre link damages, leading to improved overall network performance and stability.

Data Collection: The laboratory data provided by the TEI team contains comprehensive measurements and diagnostics from the core fibre optical network, including signal quality, fault detection logs, and maintenance records. This data is designed to support accurate identification and explanation of fibre network issues.

Conclusion:

- **Baseline**: measured as Cross Lipschitz Extreme Value for nEtwork Robustness: <0.98
- **Final Result**: 0.97 measured through CLEVER method serving as an indicator of how stable and trustworthy the AI model's predictions and explanations are measuring their explanation robustness and consistency

Key performance indicators							
ID	**Name**	**Reference to mentioned use case objectives**	**How will be measured / assessed?**	**Baseline Value**	**Target Value**	**Early Results**	**Final Results**
KPI_07	Quality ratio increase	Quality ratio increase by increasing the number of manufactured parts	This KPI is measured daily in the company, measured by $(\frac{V2-V1}{\lvert V1 \rvert}X100)$	98%	99% (This improvement in the manufacturing environment represents around 10.000 € / month for the company)	98 %	99 %
KPI_08	Reduction of scrap parts	Reduction of scrap parts	% of wrong parts, measured by $(\frac{V2-V1}{\lvert V1 \rvert}X100)$	2,2%	2% (This improvement in the manufacturing environment represents around 2.000 € / month for the company)	2,2 %	2 %

KPI_09	Lower environmental footprint	Lower environmental footprint	ISO 14001 Carbon footprint index	Under calculation (Dec.2024)	Under calculation (Dec.2024)	Under calculation (Dec. 2024)	Reduction of 16% in tools consumption
KPI_10	Increase in availability	Increase in availability	This KPI is measured daily in the company. It will be measured by $(\frac{V2-V1}{\|V1\|}X100)$	78%	79% (This improvement in the manufacturing environment represents that the whole factory would work 7,2 more hours per month, which represents 11.000 €/month)	78 %	80 %
KPI_11	Decrease investigation efforts	Decrease maintenance time	Compute the time taken by the ML classifier to infer the class of a time-series associated with the optical fibre link fault, measured by $(\frac{T2-T1}{\|T1\|}X100$ and $\|T2-T1\|)$	Focusing on the time it takes to carry out the root cause analysis, in an optical network equipped with integrated Optical Time-Domain Reflectometer (OTDR), the investigation time carried out by expert network engineers can be estimated, from experiences conducted by research laboratories. This is estimated to be approximately a couple of hours compared to the response time of an ML classifier of <1sec	In case of a correct ML-driven inference, decrease the time it approximately takes is estimated to be <1sec	<1sec	0.5secs
KPI_12	Increase optical fault analysis reliability	Improve investigation quality	Measured as Cross Lipschitz Extreme Value for nEtwork Robustness (CLEVER)	Currently, reliability is measured by observation. New measures are being introduced, as XAI is an emerging topic.	$[l_2<0.7,\ l_\infty<0.05]$, Least Likely <1.5, Random Target<1.5, and Top-2 Target<0.98	<0.98	0.97

			for XAI reliability				

4.3 AR/VR for Training & Maintenance

Objective

The aim of this use case is twofold: firstly, to provide a structured and interactive training pathway that enables personnel with no prior domain knowledge to safely gain the required expertise; and secondly, to enhance equipment maintenance processes for on-site technicians by incorporating real-time remote support from experts through augmented reality (AR) interactions.

This demonstrator showcases the utilisation of AI-enabled functions deployed at the edge to support real-time, on-site AR/VR-guided maintenance and crew training. In doing so, it highlights the broader potential of human–AI collaboration in industrial contexts. The system can be used to simulate hazardous environments, where the benefits of reducing risks to personnel are most pronounced.

The overarching goal is to assist maintenance and repair workers in diagnosing and repairing specialised equipment by providing immediate access to remote expertise. Building on the additive value of AI-enhanced AR/VR technologies for industrial applications, the scenario illustrates their integration within the TALON platform to improve reusability across maintenance operations and hazardous tasks. Furthermore, it utilises edge AI in order to reduce latency and computational overheads, thereby enabling safer, faster, and more effective interventions in critical industrial environments.

Usage Scenario

In the first scenario, a virtual reality (VR) training application is employed to familiarise new personnel with workplace-specific knowledge and procedures. Trainees are fully immersed in a virtual environment that replicates a range of operational conditions, including hazardous settings. This approach enables them to acquire critical skills in a safe, controlled, and repeatable manner before entering real-world environments.

The second scenario involves an on-site technician carrying out maintenance tasks on equipment with remote support from an expert, facilitated by an augmented reality (AR) maintenance application. The application captures and transmits the surrounding three-dimensional environment, allowing the expert to perceive the technician's context in real time. Through the integration of 3D avatars and augmented visual cues, the expert's presence is effectively projected into the technician's environment, enhancing situational awareness and enabling more precise and effective guidance during maintenance operations.

AS-IS Architecture (Current Operational State)

The existing training architecture is predominantly analogue and paper-based. It relies on printed manuals and direct expert assistance provided on the shop floor. This structure lacks digital integration, immersive technologies, or adaptive mechanisms. As a result, trainees receive static instruction without feedback loops, validation mechanisms, or the capacity to tailor training content to their individual needs or performance.

The maintenance architecture is similarly fragmented and relies on manual and paper-driven processes. When a problem occurs, a paper-based work order is generated and forms the backbone of the workflow. Maintenance personnel attempt to address the issue using their personal expertise, supported by static consultation of printed manuals. In cases where additional expertise is required, off-site experts are contacted through conventional communication channels, with their recommendations based solely on problem descriptions provided remotely. Where remote assistance proves insufficient, the architecture necessitates the physical deployment of the expert to the site. The current ("AS-IS") training architectures are illustrated in Figure 5.

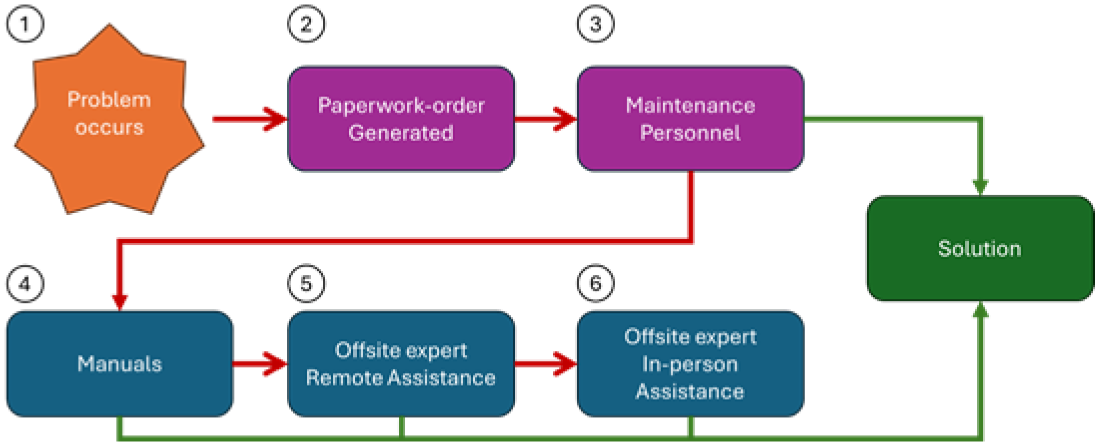

Figure 5: *Use Case 3* current AS-IS architecture

TO-BE Architecture (TALON-Orchestrated Coordination)

The TO-BE architecture focuses on two scenarios: VR training for personnel in training, see Figure 6, and on-site AR-assisted maintenance, see Figure 7. In the first scenario, trainees undergo VR-based immersive training, including exposure to hazardous environments, with the environment simulated using a digital twin. Trainees follow assigned goals while real-time reports are transmitted to remote experts for interactive guidance. In the second scenario, on-site maintenance personnel use AR headsets connected to edge nodes, transmitting real-time point-of-view feeds to remote experts. TALON's AI capabilities analyse and semantically enrich the data, providing actionable information on hazardous materials, access points, and equipment directly on the AR headsets to guide staff during servicing. Automatic reports are generated throughout both scenarios to document activities and support ongoing evaluation.

TALON provides a common orchestration layer to manage resources and data flows more effectively. For the VR training application, this means a clearer path towards adapting training materials to individual needs and ensuring more efficient use of available computational resources. For the AR maintenance application, TALON helps coordinate real-time communication and improve consistency when scaling the solution across different sites. More broadly, such orchestration has a potential to support incremental steps towards more efficient and sustainable operations.

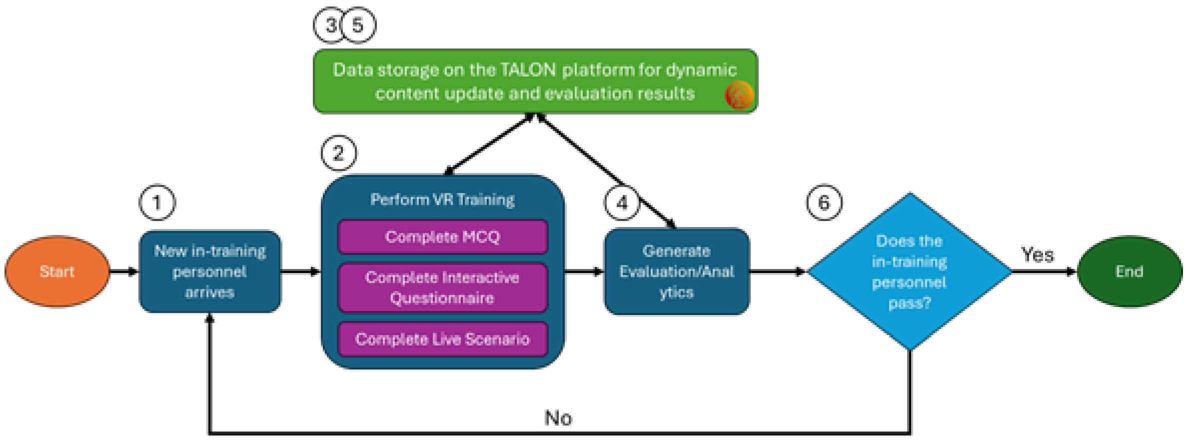

Figure 6: Use Case 3 TO-BE architecture of VR training application.

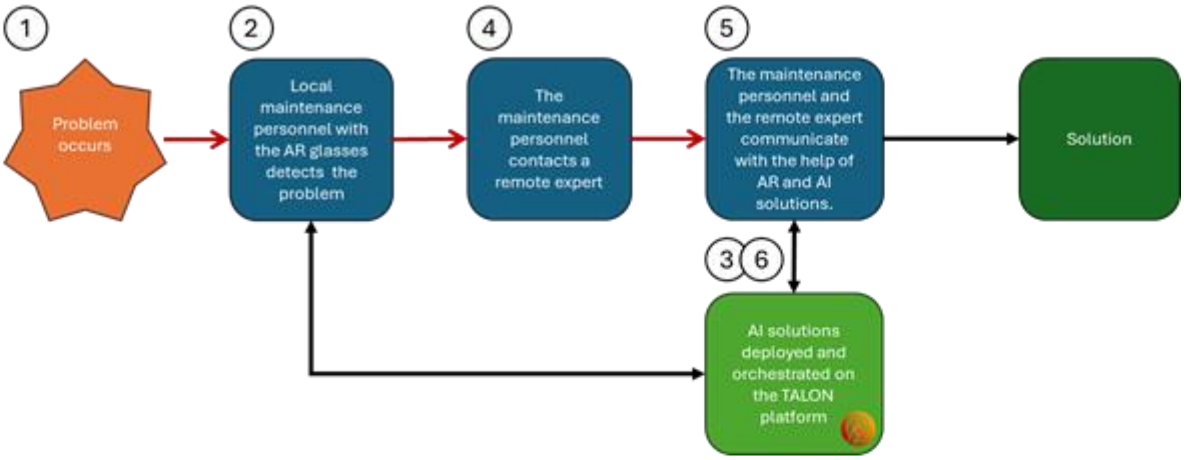

Figure 7: Use Case 3 TO-BE architecture of AR maintenance application.

Technical Evaluation

System Components:

- In-training personnel wearing a VR headset immersed in a simulated environment to safely practice workplace tasks and procedures.
- On-site technician wearing an AR headset provided real-time visual guidance and remote expert support during equipment maintenance.
- Dashboard enables experts to monitor, communicate, and guide on-site personnel through interactive AR interfaces.
- TALON Dashboard for energy and emission monitoring displays real-time metrics on energy consumption and emissions to support sustainable industrial operations.
- AI based life-long learning object detection algorithms to continuously improve the recognition of equipment and components to enhance AR maintenance procedures over time.

Data Format:

- Input
  - Image data captured by the AR headset during maintenance operations.
  - Depth data representing the spatial environment surrounding the AR headset user.
  - JSON annotations supporting life-long learning for object and context recognition.
- Output
  - JSON containing classified objects, such as personal protective equipment (PPE), fire hazards, and various factory equipment.
  - JSON containing real-time energy and emission metrics.
  - 3D visualisations delivered as a byte-stream for display on the remote expert dashboard.

End-User Evaluation

<u>Benefits to operators:</u>

- Safer skill acquisition: VR training allows personnel to practice hazardous procedures in a controlled, risk-free environment.
- Human-in-the-loop support: Operators retain control over tasks while benefiting from augmented guidance and expert input.
- Enhanced situational awareness: AR visualisation of equipment and 3D environment helps operators navigate complex maintenance tasks.
- Low cognitive load: AI-enhanced AR guidance helps technicians correctly identify and handle equipment components. Step-by-step AR instructions reduce the mental burden on technicians, allowing focus on critical decisions.
- Real-time feedback: VR and AR applications provide immediate performance insights and training adaptation to individual needs.
- Sustainability: Efficient, guided maintenance reduces unnecessary equipment handling and travel for experts, lowering energy use and operational waste.

End-User Feedback:

The system was described as highly usable and intuitive, providing a positive overall experience.

Interactive questions felt realistic and closely aligned with everyday tasks, supporting knowledge transfer.

Live scenarios were seen as both challenging and engaging, reinforcing learn-by-doing as a practical training approach.

Manufacturing personnel noted a clear reduction in workload, with less effort, physical demand, and frustration during maintenance tasks.

Real-time voice and video communication was highlighted as improving coordination and efficiency.

Action cards were appreciated for offering useful guidance without restricting hands-on activity.

Recommendations / Notes by KU

- Incremental Deployment: AR maintenance solutions should be introduced gradually, beginning with non-critical systems, to refine responsiveness, usability, and integration with existing practices.
- User-Centred Design: Future development should prioritise ergonomics, accessibility, and ease of use to ensure adoption by maintenance personnel and trainees with varying levels of digital literacy.

4.3.1 Key Performance Indicators' Final Evaluation

KPI_01: Latency reduction

Objective: Reduce AR application response time by lowering network latency below 20 ms.

Data Collection: To assess network performance, evaluations were conducted across all available communication mediums, including Wi-Fi 4, Wi-Fi 5, Wi-Fi 6, and Ethernet. The average round-trip latency recorded across these mediums was 7.47 ms. This latency value was combined with the previously calculated average transmission time of 7.58 ms, resulting in a new total response time of: 7.58 ms (transmission)+7.47 ms (latency)=15.05 ms. This total response time was derived from monitored data communication between the AR maintenance application and the backend solution over a defined period. For each request, the response time was calculated by summing the transmission time and the round-trip latency.

Conclusion: Latency was significantly reduced from the baseline of 50 ms to a new total of 15.05 ms, which is well below the 20 ms target, demonstrating a successful optimisation of network performance across multiple mediums.

Baseline: 50 ms

Final Result: 15.05 ms

KPI 13: AI-human collaboration effectiveness

Objective: Improve the effectiveness of AI-human collaboration in AR-supported maintenance and training tasks.

Data Collection: Surveys (5-point Likert scale), structured interviews, and user interaction logs were used. Surveys measured satisfaction and usability; logs measured task accuracy and completion time. Results were analysed statistically and visualised with bar charts and radar charts.

Conclusion: Collaboration effectiveness improved beyond the target, showing increased satisfaction and efficiency.

Baseline: 2.0 (Likert average)

Final Result: 3.727 (Likert average)

KPI 14: AR-to-node AR POV latency reduction

Objective: Decrease AR-to-node POV transmission latency by more than 90%.

Data Collection:

Monitoring tools recorded timestamps for AR frame capture and node reception before and after system optimisation.

Conclusion: Transmission latency was reduced by >90%, meeting the KPI target.

Baseline: 2 s

Final Result: 198 ms

KPI 15: Increased reusability (training attendance)

Objective: Increase reusability through training attendance rates above 95%.

Data Collection: Attendance logs were digitally recorded and compared to the number of invited participants. Attendance percentages were analysed session by session.

Conclusion: Training attendance rates increased significantly and met the ≥95% target.

Baseline: 1/20 participants (5%)

Final Result: 19/20 participants (95%)

KPI 16.1: Increased environment recognition

Objective: Improve AR system recognition accuracy for environments and gestures above 90%.

Data Collection: The system's recognition outputs were compared with a labelled dataset. Accuracy was calculated as (correct recognitions ÷ total cases) × 100.

Conclusion: Recognition accuracy improved from 70% at baseline to 90.03%, just meeting the 90% target and confirming system improvement.

Baseline: 70% accuracy

Final Result: 90.03% accuracy

KPI 16.2: Preservation of anonymisation (newly introduced and measured KPI)

Objective: Guarantee 100% anonymisation of detected sensitive data (faces, numbers, text).

Data Collection: Anonymised data were reprocessed using the same detection algorithms initially applied to identify sensitive elements. Success was defined as no detectable sensitive information in anonymised outputs.

Conclusion: Anonymisation was fully preserved, ensuring compliance with privacy requirements.

Baseline: N/A

Final Result: 100%

Key performance indicators							
ID	**Name**	**Reference to mentioned use case objectives**	**How will be measured / assessed?**	**Baseline Value**	**Target Value**	**Early Results**	**Final Results**
PI_01	Latency reduction	Latency time reduction	The response time will be measured and assessed by utilising performance monitoring tools to track network latency and response times in real-time, implementing logging within the AR application to record precise timestamp at request initiation and response reception, and analysing this data to ensure	50ms	20ms	21.58ms	15.05ms

			the response time reaches below 20ms, measured by ($\|T2-T1\|$)					
KPI_13	AI-human collaboration effectiveness		AI-human collaboration effectiveness	The AI-human collaboration effectiveness will be measured through a combination of surveys and feedback forms to gauge human satisfaction and user experience during the AR-maintenance scenario and training activities. Additionally, performance metrics will be collected to evaluate the efficiency and accuracy of tasks completed through AI-human collaboration, aiming for over 50% effectiveness. User interaction data and qualitative assessments will further validate improvements in user experience and interaction, measured by a 5-point Likert system	2.0 average on 5-point Likert system	3.0+ average on 5-point Likert system	N/A	3.852 average on 5-point Likert system
KPI_14	AR-to-node AR POV latency reduction		Decrease in AR-to-node AR point-of-view (POV) transmission latency	The >90% decrease in AR-to-node POV transmission latency will be measured by comparing the latency metrics	2s	200ms	200ms	200ms

				before and after implementing the new system. Continuous monitoring tools will record the time taken for data to transmit from the AR device to the node. Baseline latency values will be established prior to the new system's deployment, and subsequent latency measurements will be taken to ensure they reflect a decrease of over 90%. Regular performance reports and real-time analytics will be used to validate the latency reduction such as in the case of data required for the training with and without few-shot learning, measured by ($\lvert T2 - T1 \rvert$)				
KPI_15	Increased reusability		Increased reusability and training attendance rate	The training attendance rate of >95% will be measured by tracking participant attendance records for each training session. Attendance logs will be maintained and reviewed to calculate the percentage of attendees relative to the	1/20 participants	19/20 participants	N/A	19/20

				total number of invited participants. The assessment will be conducted regularly to ensure the attendance rate consistently meets or exceeds 95%, and reports will be generated to document compliance with this metric.				
KPI_16.1	Increased environment recognition		Increased gesture / environment recognition	The >90% accuracy in environment recognition will be measured by comparing the system's recognition outputs against a predefined set of known environments. The system's recognition accuracy will be evaluated through a series of tests, where the correct identification rate is calculated by dividing the number of correctly recognised objects by the total number of test objects. Regular performance evaluations and benchmarking against these known environments will ensure the accuracy consistently meets or exceeds the	70% accuracy	>90% accuracy	*90.03%*	*90.03%*

				90% threshold, measured by $(\frac{V2-V1}{\lvert V1 \rvert}X100)$				
KPI_16.2	PII preservation		Preserve workers' sensitive and personal identifiable information	The 100% PII preservation through facial, numerical, and textual data anonymisation will be measured by conducting audits on the anonymised data aiming to reidentify the related removed information. These evaluations will involve cross-referencing the anonymised data against the detection algorithms that were used initially to ensure no personally identifiable information (PII) can be retrieved, measured by $(\frac{V2-V1}{\lvert V1 \rvert}X100)$	n/a	100%	100%	100%

4.4 Human Robot Collaboration

Objective

The objective of UC4 is to apply the TALON system in a factory setting for the intelligent, efficient, and timely identification of workers' compliance with personal protective equipment (P.P.E.) requirements, such as helmets, vests, and other safety gear. Traditionally, this process relied exclusively on empirical methods, with safety officers manually supervising the premises, which introduced the risk of human oversight. In the proposed approach, a drone, integrated with the TALON system, will autonomously monitor the site and detect compliance violations in real time. This transition enables data-driven, automated supervision while reducing manual workload.

Usage Scenario

In a factory environment, a field engineer and a drone (robot) collaborate to monitor worker safety and compliance. The drone collects visual data (images/videos) of the workspace, while TALON's AI pipeline processes the data to detect objects such as helmets, vests, and protective equipment. If safety violations are detected, TALON triggers alerts, enabling the field engineer to take corrective actions before accidents occur.

Coordination is achieved via:

- Drone Operator (Field Engineer): Oversees the collaborative process and takes corrective action if needed.
- AI Preprocessing Pipeline: Performs multi-object detection and safety compliance checks.
- TALON Dashboard: Provides real-time visualization of safe/unsafe working conditions.

AS-IS Architecture (Current Operational State)

Main Characteristics:

- Safety monitoring is performed manually by the field engineer/safety officer.
- Compliance checks rely solely on human observation and judgment.
- No automated detection of protective equipment (helmets, vests, etc.).
- High risk of human error or oversight in safety monitoring.
- Accident prevention depends on direct human-to-human communication.

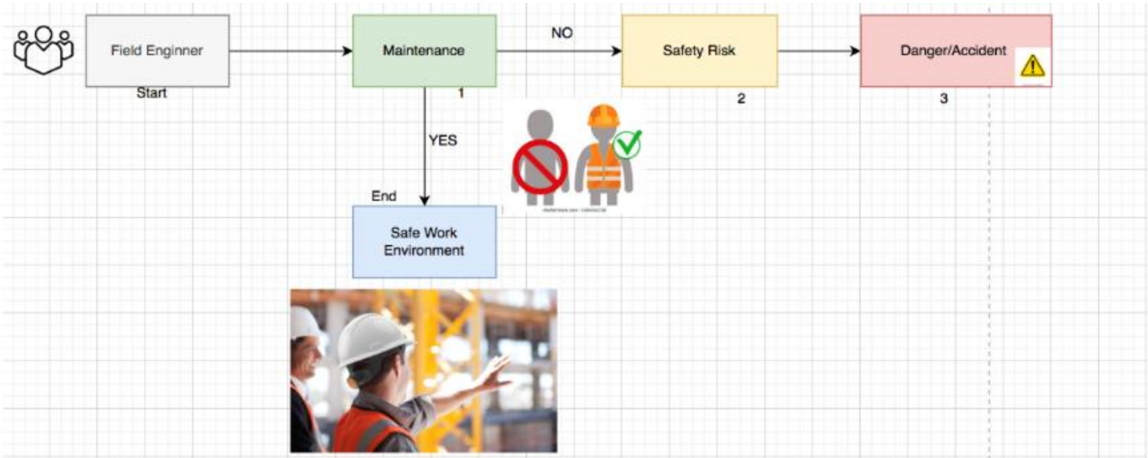

Figure 8: Use Case 4 AS-IS Diagram

TO-BE Architecture (TALON-Orchestrated Coordination)

Enhanced Features:

- Drone-assisted monitoring of working conditions through video/image capture.
- AI-based preprocessing pipeline for real-time detection of safety equipment.
- Task distribution is dynamically managed based on the assessed complexity of each task.
- Automatic alerts for unsafe conditions, reducing response time.
- User-friendly TALON dashboard for visualization of compliance and alerts.

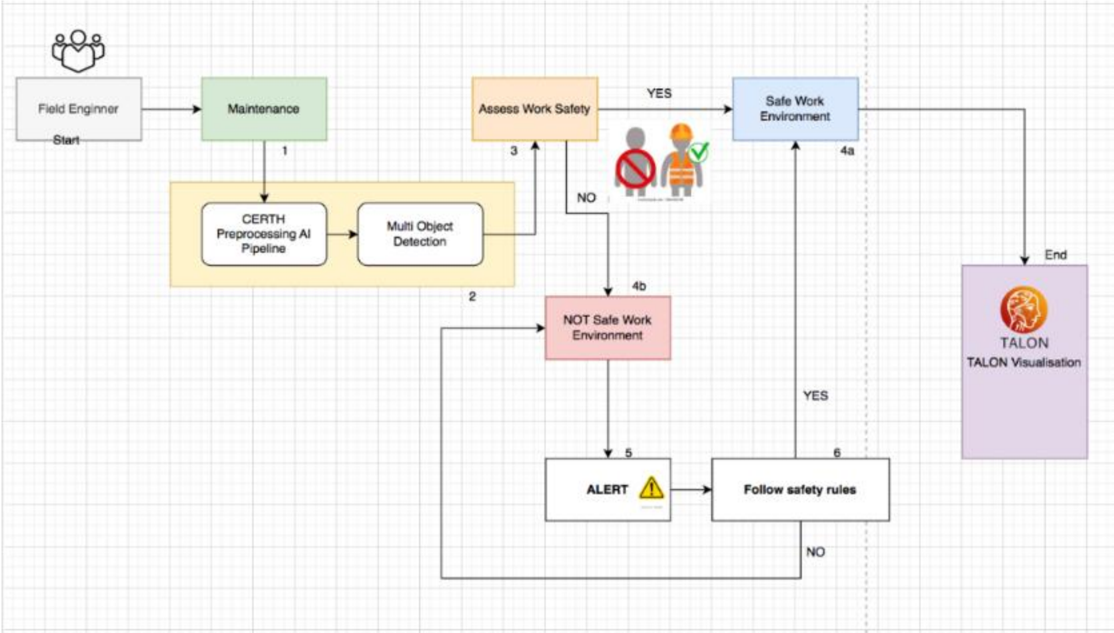

Figure 9: Use Case 4 TO-BE Diagram

Technical Evaluation

<u>System Components</u>

- Drone (Robot): Captures live videos/images in the workspace.
- Field Engineer (Human): Supervises safety and ensures compliance when alerts are triggered.
- AI Pipeline: Multi-object detection (helmets, vests, PPE) for safety assessment.
- TALON Dashboard: Provides real-time visualization and notifications of safe/unsafe conditions.

<u>Data Formats</u>

- Input: Video/image streams from drone, transferred to preprocessing pipeline.
- Output: Structured compliance reports, visual alerts, and safety logs.

End-User Evaluation

<u>Benefits to Operators</u>

- Low Cognitive Load: Drone + AI reduces the need for constant manual inspections.
- Real-Time Awareness: Immediate alerts for safety violations via TALON dashboard.
- Resilience: Hybrid collaboration allows fallback to manual supervision.
- Safety Assurance: Automated detection minimizes risks from human oversight.
- Efficiency: Faster detection of non-compliance ensures safer, more productive operations.

<u>End-User Feedback</u>

- Positive reception of automated safety checks compared to purely manual monitoring.

- Notable reduction in time required to identify and correct unsafe conditions.
- Improved trust in safety processes due to AI-driven detection and explainable results.

Recommendations / Notes by CERTH

- Extend the system from single-drone to multi-drone monitoring for large factories.
- Enhance AI pipeline to support multimodal input (e.g., thermal + RGB for low-light safety checks).
- Integrate predictive safety analytics using historical data to prevent recurring risks.
- Explore integration of wearable sensors for more precise worker safety monitoring.

4.4.1 Key Performance Indicators' Final Evaluation

The objective of this scenario-UC4 is to develop an intelligent, autonomous control system that ensures compliance with safety regulations in a factory environment, ultimately enhancing worker safety. Specifically, the system is designed to recognize adherence to safety protocols by detecting the use of personal protective equipment (PPE), such as vests and helmets, within the workplace. The final implementation leverages the TALON system and its suite of advanced tools, including an AI orchestrator, blockchain technology, distributed federated learning, data anonymization capabilities and other integrated technologies that support secure, scalable, and intelligent system operation.

The UC4 scenario demonstrates a transformative shift in occupational safety monitoring within industrial environments. Previously, compliance with personal protective equipment (PPE) requirements relied entirely on human oversight by safety officers. This method was limited in coverage, subject to human error, and lacked real-time responsiveness.

With the integration of the TALON system and aerial drones, safety monitoring is now automated, intelligent, and data-driven. TALON enables real-time identification of workers not complying with PPE regulations using computer vision technology. This shift not only improves accuracy and efficiency but also supports scalable operations and the measurement of system performance through key performance indicators (KPIs). The result is a more proactive and reliable approach to worker safety.

Next, we present the methodology behind the KPI calculations for this scenario:

KPI_16 - Increased object / environment recognition

- This KPI focuses on enhancing the detection of humans and commonly encountered objects in industrial sites, including Personal Protective Equipment (PPE) such as helmets and vests, as well as identifying their absence. This, in turn, contributes to improving worker safety in industrial environments.
- Specifically, the following metrics are used: Recall (%), Precision (%), and mean Average Precision (mAP %).
- A YOLOv8n model was trained on the training set of the Worker-Safety dataset and evaluated on its test set, achieving over 90% across all examined classes and metrics, as shown in the following table.

- **Conclusion:** Fine-tuning a pretrained object detection model, such as YOLOv8n, demonstrates the capability to accurately detect workers and PPE in industrial environments, with performance surpassing 90% accuracy.

Class	Precision(%)	Recall(%)	mAP(%)
person	97.9	91.7	99.3
helmet	98.7	98.0	99.3
no-helmet	95.8	100.0	99.5
vest	97.8	94.5	97.9
no-vest	99.5	96.0	99.2
all	97.9	96.0	99.0

KPI_17 - Preservation of anonymisation

- This KPI does **not** evaluate the **face detection rate**, as no existing technique currently guarantees 100% accuracy in face detection. Instead, the KPI measures the effectiveness of anonymization (i.e., obfuscation) on the faces that have been detected, where the module achieves 100% success.

- How is this validated? Each anonymized image is reprocessed using the **same face detection algorithm** initially used to identify faces. The KPI is considered successful when **no faces are detected** in the anonymized image, proving that all previously detected faces have been effectively anonymized.
- **Conclusion:** The Image Anonymization Module guarantees full anonymization (**100% effectiveness**) of all faces it successfully detects, ensuring compliance with privacy requirements in all processed cases.

KPI_18 - AI-to-AI communication

- The KPI **"Reduction in AI-to-AI communication latency"** was evaluated by measuring the time elapsed between two consecutive events:
  - **T1**: when an AI module (e.g., perception/vision AI) generates and transmits a data request.
  - **T2**: when the receiving AI module successfully receives and acknowledges that request.
- Latency was defined as the absolute difference |T2 – T1|. Measurements were carried out under realistic pilot conditions, with multiple repetitions to capture variability in network load, computational resources, and message sizes. Both the baseline system (AS-IS) and the optimized TALON-enabled system (TO-BE) were tested using the same experimental setup to ensure fair comparability.

In the CERTH pilot (Use Case 4), the baseline latency was consistently below 1 second. After applying TALON's benchmarking tools and optimization framework, latency was reduced to below 0.25 ms, as measured through direct time-stamping and logging within the system's edge-cloud communication pipeline.

KPI_19 - Human vs Robot Inspection time

- According to inspection records, the average time for a security officer to patrol a site is **20 minutes**, which serves as the **baseline value** for KPI 19.
- When using a drone to supervise the same area, time was measured using a stopwatch, yielding an average of **4 minutes** per site.
- **Conclusion**: Drone implementation leads to an **80% reduction** in inspection time per site.

KPI_20 - Robust AI models

The robustness of the computer vision models was assessed through a comparative evaluation between the baseline (AS-IS) and the TALON-optimized models.

Datasets Used:

- A combination of benchmark datasets (e.g., industrial object detection datasets) and real pilot data collected from Pilot 4.
- The dataset was divided into training, validation, and test subsets to ensure reproducibility and prevent data leakage.

Evaluation Metrics:

- Accuracy (%) was selected as the primary metric, computed as the proportion of correct predictions (true positives + true negatives) over the total number of samples.
- In addition, confusion matrices and error analysis were carried out to ensure balanced performance across different object classes.

- The initial AI models trained before TALON optimisations achieved 60% accuracy on the test set.
- Using TALON's enhancements (e.g., advanced training pipelines, data augmentation, hyperparameter optimisation, and improved model architectures), the optimised models were retrained and tested on the same test dataset for comparability.
- The final measured accuracy was 80%, representing a +20% absolute improvement.

Validation Procedure:

- To confirm reliability, cross-validation and multiple training runs were conducted.
- Performance stability was also tested under pilot-specific conditions

The final accuracy value (80%) was measured by evaluating the TALON-optimised computer vision models on benchmark datasets and real pilot data from Use Case 4. The same test dataset was used for both baseline (AS-IS) and optimised (TO-BE) models, ensuring comparability. Accuracy was computed as the proportion of correct predictions, with additional validation through cross-validation and repeated training runs to confirm robustness under pilot-specific conditions.

KPI_21 - Protective equipment readiness

- In a manual patrol scenario, an alert is triggered when a security officer identifies a worker without the required protective equipment.
- Since detection depends on when the worker is encountered, the time varies. However, complete area coverage requires a **20-minute patrol**, which is used as the average time to raise an alert.

- With a drone, area supervision takes about **4 minutes**, and initiating the alert process through the **TALON system** takes an additional **1 minute**.
- **Conclusion**: Drone-assisted supervision reduces the average alert time to 5 minutes, resulting in a **75% reduction** compared to manual patrolling.

KPI_22 - Personnel safety

- Adopting a similar methodology to that presented in KPI_16, this KPI evaluates the detection accuracy of pretrained object detectors fine-tuned to recognize PPE, including helmets and vests, in realistic industrial settings, with the aim of improving worker safety in such environments.
- In this KPI, a dataset comprising drone-acquired images of workers in real-world industrial environments was employed to fine-tune and evaluate the YOLOv8n object detection model.
- Applying the same evaluation metrics as in KPI_16, the fine-tuned YOLOv8n model achieved a mAP of 70.5% for helmet detection and 93.7% for vest detection. The complete results are presented in the following table.
- **Conclusion:** Fine-tuning lightweight pretrained object detection models, such as YOLOv8n, on drone-acquired imagery for PPE detection in real-world industrial settings can yield high detection accuracy, contributing substantially to improved worker safety in such environments.

Class	Precision(%)	Recall(%)	mAP(%)
helmet	80.5	71.6	70.5
vest	96.5	88.0	93.7
all	88.5	79.8	82.1

Key performance indicators							
ID	**Name**	**Reference to mentioned use case objectives**	**How will be measured / assessed?**	**Baseline Value**	**Target Value**	**Early Results**	**Final Results**
KPI_16	Increased object / environment recognition	Increased object recognition	Experiments for accuracy between previous and optimised AI Models, measured as Delta function to calculate the improvement $\|Acc2 - Acc1\|$	85%	>90%	90%	99% overall mAP averaged across all classes

KPI_17	Preservation of anonymisation	Preserve workers' sensitive and personal identifiable information	The anonymized image is re-processed through the same face detection model used. The absence of detected faces confirms the preservation of the anonymization.	AS-IS no anonymisation is applicable	100% (numerical, textual and imagery)	-	100%
KPI_18	AI-to-AI communication	Reduction in AI-to-AI communication latency	Latency measured before and after TALON's AI benchmarking and optimisation, measured as $(\|T2 - T1\|)$	AS-IS time for AI-to-AI communication is <1sec	TO-BE time for AI-to-AI communication will be <0.25ms	-	< 0.25ms
KPI_19	Human vs Robot Inspection time	Robot production efficiency	Timer, measured as $(\|\mathrm{T2} - \mathrm{T1}\|)$	20 min (person walking)	4 min (drone flying)	For 4 places Person time: 20-25 min Drone time: 4-6 min	80% decrease in inspection time
KPI_20	Robust AI models	Robust computer vision models	Experiments for accuracy between previous and optimised AI Models, measured as Delta function to calculate the improvement $\|Acc2 - Acc1\|$	60%	>80%	75%	80%
KPI_21	Protective equipment readiness	Robust computer vision models	Timer, measured as $\|\mathrm{T2} - \mathrm{T1}\|$	2o min (time to raise an alert)	5 min	For 4 places Person time: 180 min	75% decrease in time to raise an alert

						Drone time: 120 min	
KPI_22	Personnel safety	Multimodal object detection	Measured by means of AS-IS and TO-BE difference, expressed by AI Model accuracy as $\|Acc2 - Acc1\|$	AS-IS: no AI / ML method is currently used; only visual personnel inspection	>70% personnel safety increase, expressed by a robust object classifier / detector (i.e., accuracy, f1, precision, recall)	Accepted values > 70%	82.1% overall mAP averaged across all classes

5 End User Evaluation Results

In order to examine the overall satisfaction of the Talon platform integration, usage and tools that have been made available to the users, developers and everyone associated with it, a survey has been conducted and disseminated in the pilot sites. The survey has had 43 responses.

The questionnaire was designed to capture both quantitative and qualitative feedback across different aspects of the TALON platform. The first set of questions assessed satisfaction with key functionalities such as the dashboard, ease of use of the tools, data security and privacy, user interface, data visualisations, and authentication process. Responses were measured on a five-point Likert scale, ranging from Strongly Disagree to Strongly Agree, allowing us to measure the overall acceptance and usability trends.

The second and third part of the survey focused on more detailed ratings, using 1–5 and 1–10 scales, to evaluate dimensions such as dashboard navigation, the intuitiveness of explainable AI outputs, the efficiency of AIOps procedures, the usefulness of security and privacy mechanisms, and the effectiveness of the Zero-Touch AI Orchestrator. Respondents were also asked to rate the overall design, behaviour and capabilities of the TALON platform, as well as its potential usefulness in the future. Finally, the survey concluded with a Net Promoter Score (NPS) style question, asking participants how likely they would be to recommend the TALON platform to colleagues. This structure ensures that both numerical measures and qualitative insights are gathered, providing a holistic view of user satisfaction, system reliability and future potential.

Overall, the survey results reflect a very positive reception of the TALON platform among the participants. Most respondents agreed or strongly agreed that the dashboard is user-friendly, the tools are easy to use, and the general usage of the platform was straightforward. In particular, over two-thirds of the answers for usability-related questions (e.g. dashboard navigation, authentication, visualisations) were in the Agree or Strongly Agree categories, while only a very small number of respondents reported disagreement.

Taken together, the survey confirms that TALON succeeded in providing a usable and accessible platform that meets user expectations across pilots.

The results of the questions asked are presented in the following figures as percentages of each answer given:

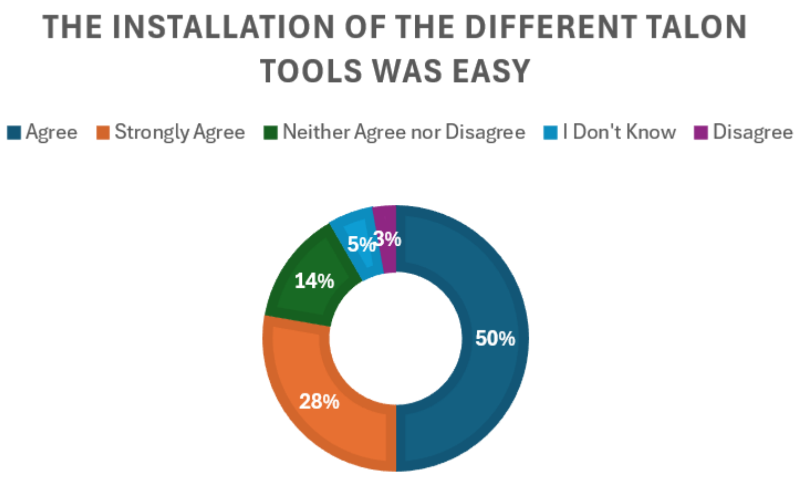

Figure 10: Level of satisfaction with the following aspect of the application (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree): The installation of the different TALON tools was easy.

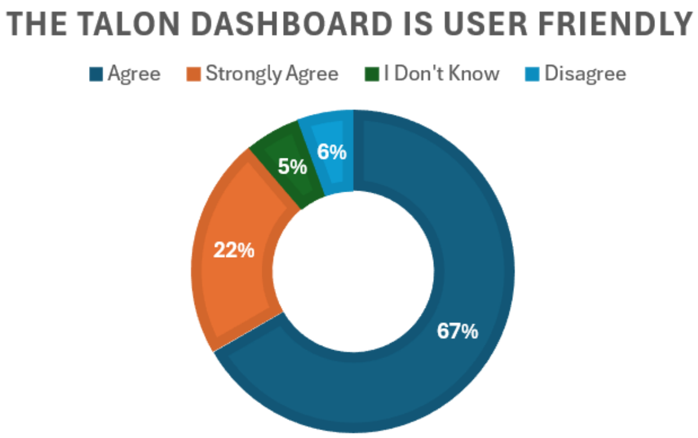

Figure 11: Level of satisfaction with the following aspect of the application (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree): The TALON Dashboard is user-friendly.

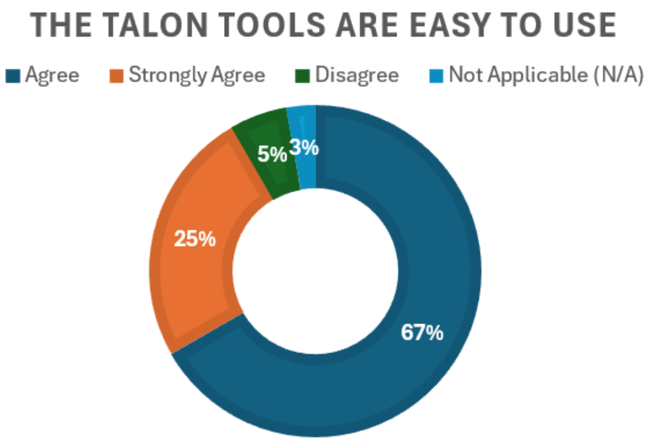

Figure 12: Level of satisfaction with the following aspects of the application (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree): The TALON tools are easy to use.

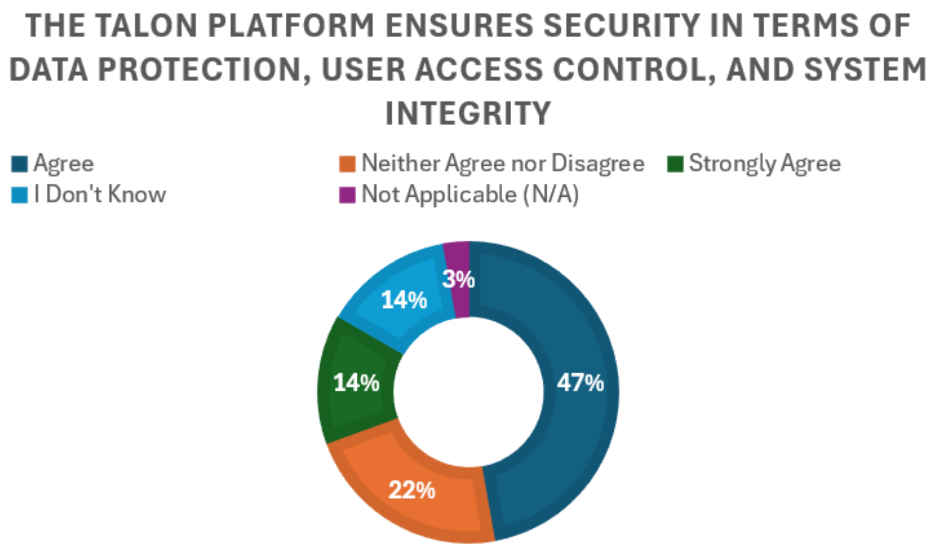

Figure 13: Level of satisfaction with the following aspect of the application (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree): The TALON Platform ensures security in terms of data protection, user access control, and system integrity.

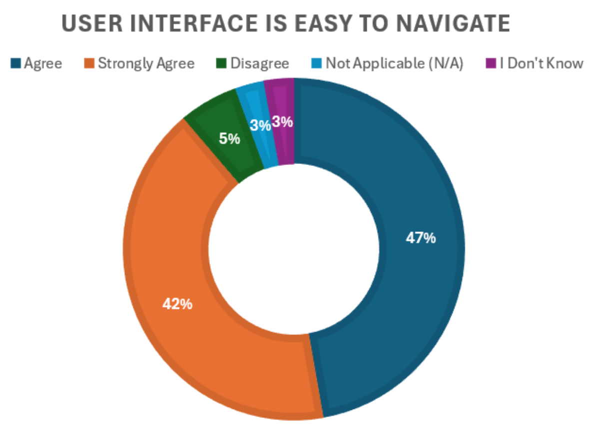

Figure 14: Level of satisfaction with the following aspect of the application (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree): User interface is easy to navigate.

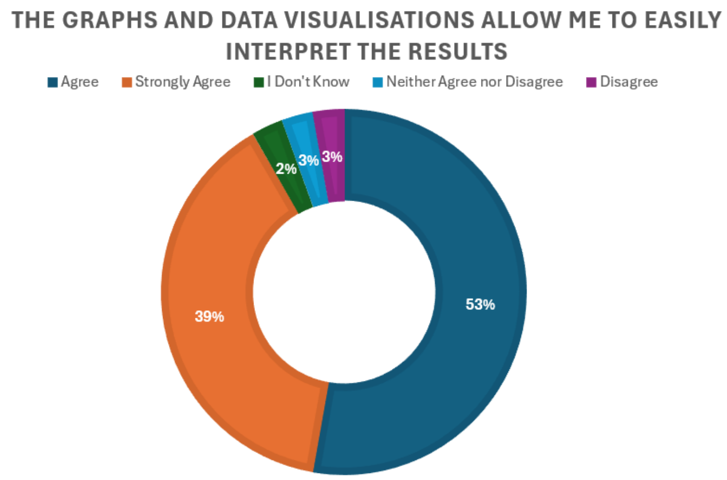

Figure 15: Level of satisfaction with the following aspect of the application (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree): The graphs and data visualisations allow me to easily interpret the results.

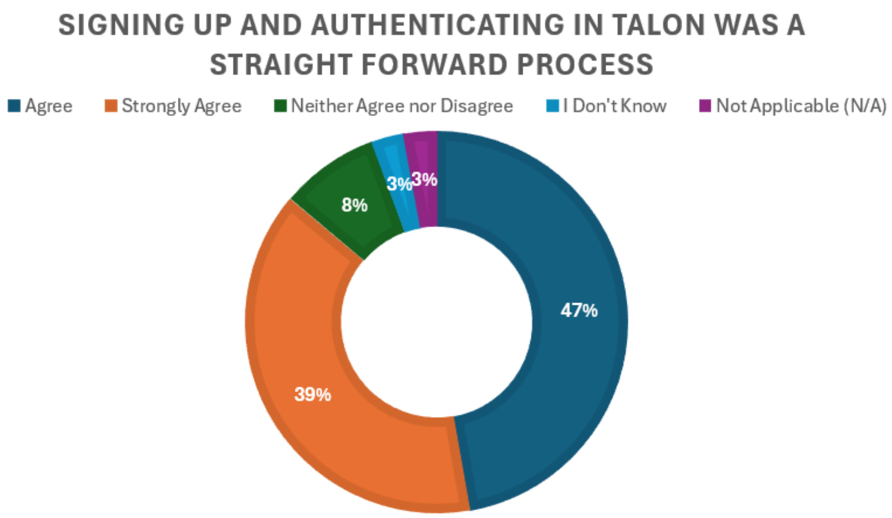

Figure 16: Level of satisfaction with the following aspect of the application (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree): Signing up and authenticating in TALON was a straightforward process.

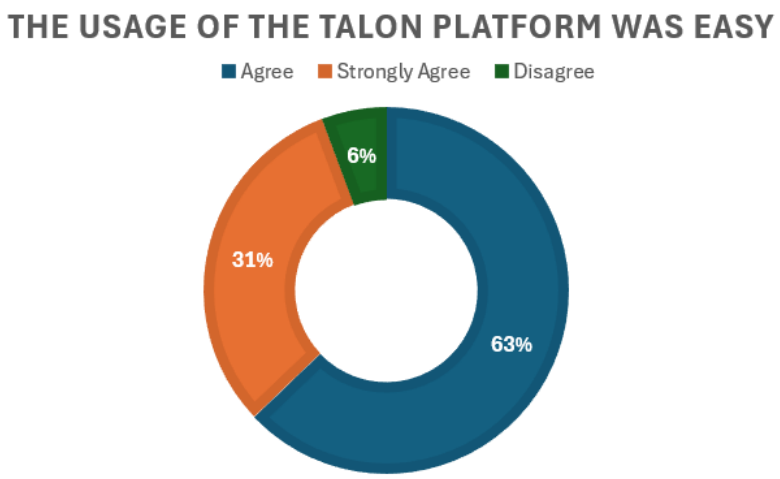

Figure 17: Level of satisfaction with the following aspect of the application (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree): The usage of the TALON Platform was easy.

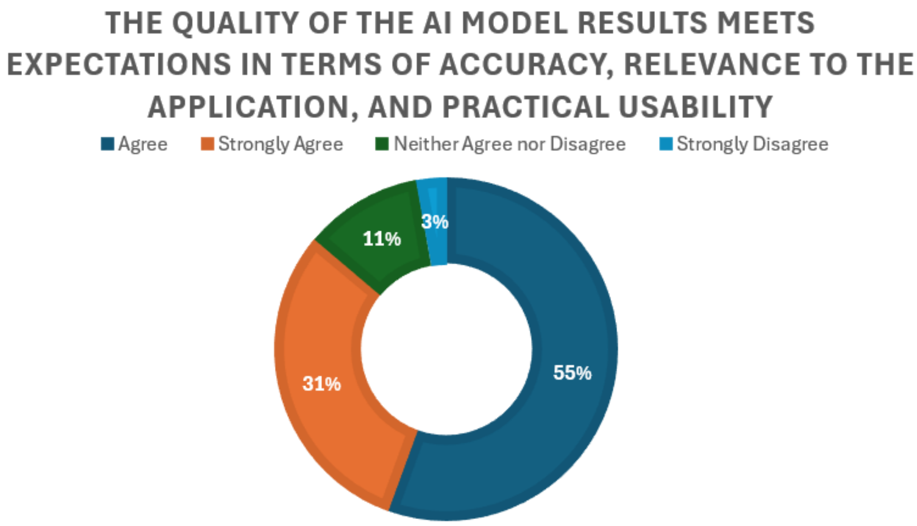

Figure 18: In terms of performance and interest, give your opinion on these aspects (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree): The quality of the AI model results meets expectations in terms of accuracy, relevance to the application, and practical usability.

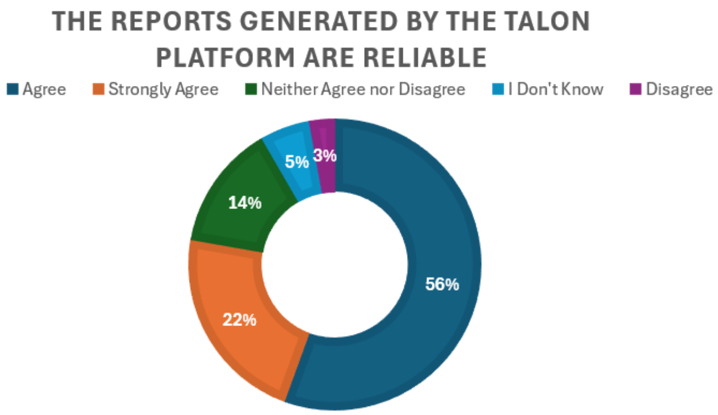

Figure 19: In terms of performance and interest, give your opinion on these aspects (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree): The reports generated by the TALON Platform are reliable.

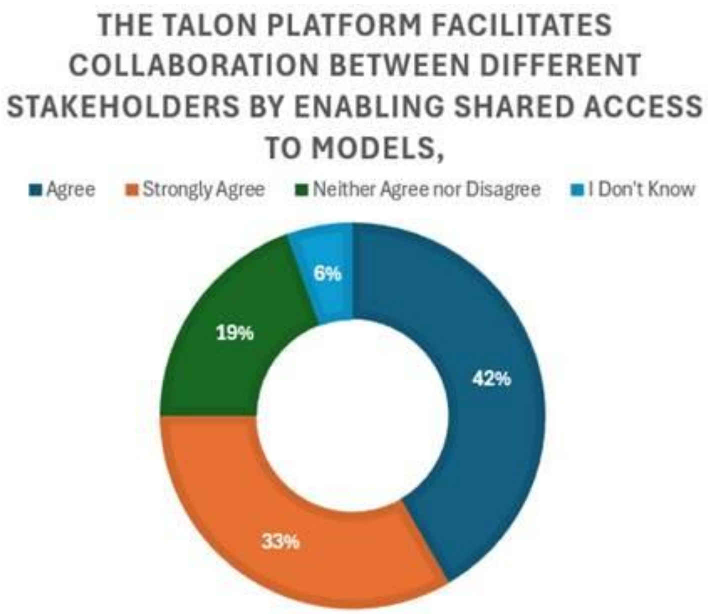

Figure 20: In terms of performance and interest, give your opinion on these aspects (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree): The TALON Platform facilitates collaboration between different stakeholders by enabling shared access to models.

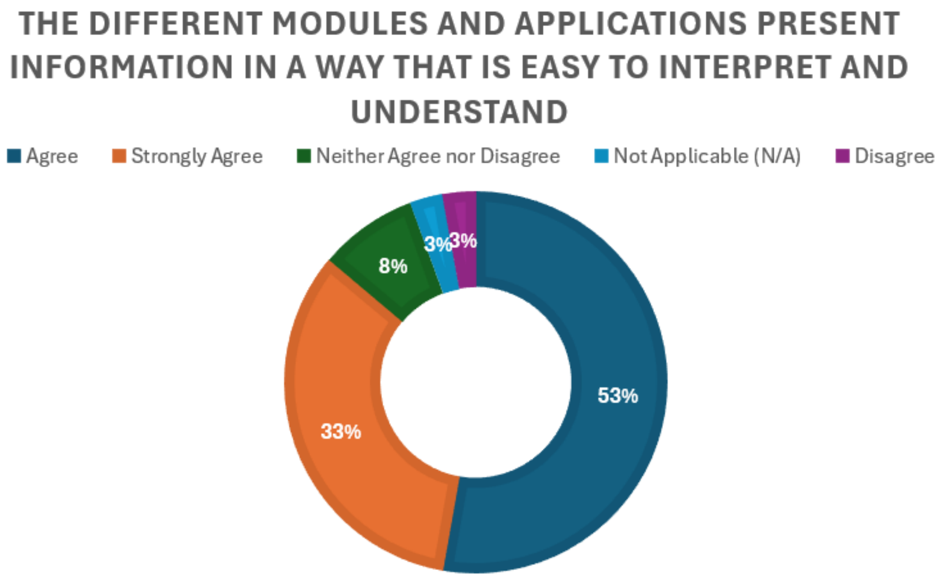

Figure 21: In terms of performance and interest, give your opinion on these aspects (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree): The different modules and applications present information in a way that is easy to interpret and understand.

In the next set of questions, the users have been asked to provide a rating as an answer with a scale ranging from 1-5, as well as N/A (Not Applicable) and I Don't Know, in order to respondents to make quick judgments without overthinking. The answers are presented in the graphs below:

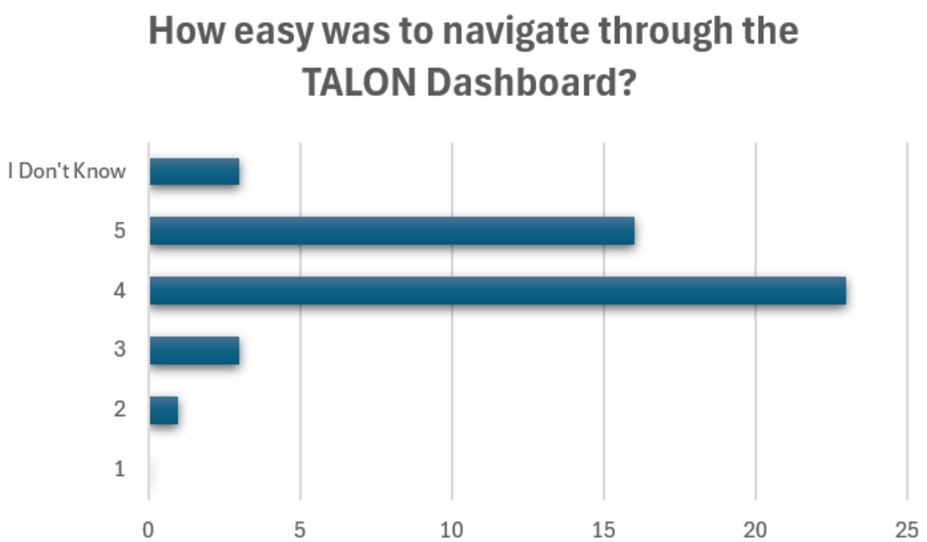

Figure 22: How easy was to navigate through the TALON Dashboard? (1=Not at all, 5=Extremely)

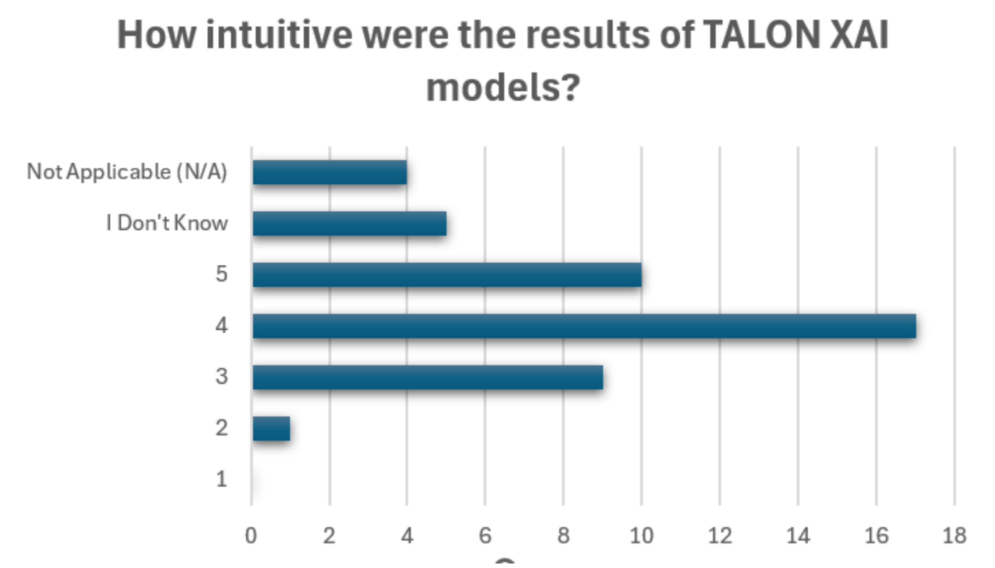

Figure 23: How intuitive were the results of TALON XAI models? (1=Not at all, 5=Extremely)

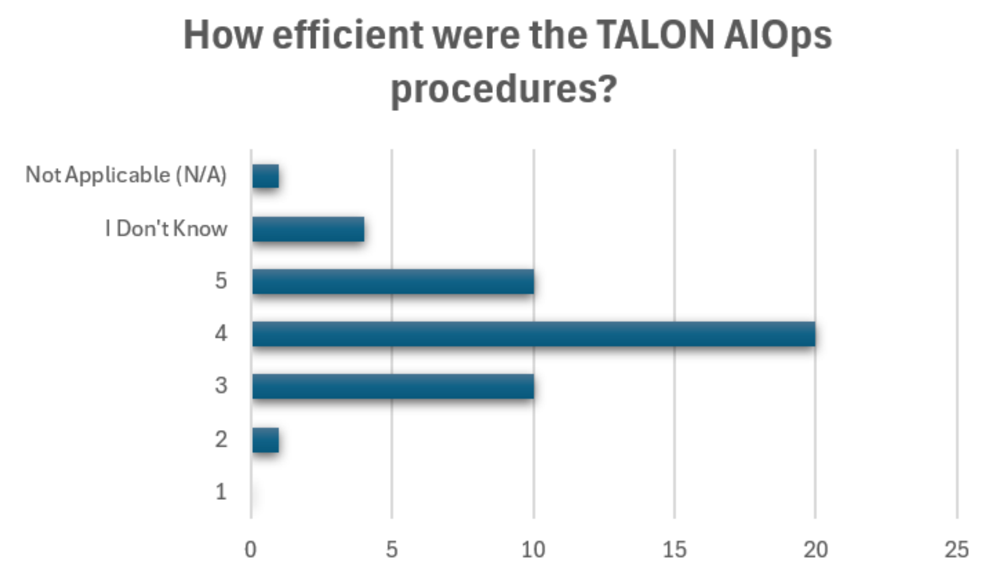

Figure 24: How efficient were the TALON AIOps procedures? (1=Not at all, 5=Extremely)

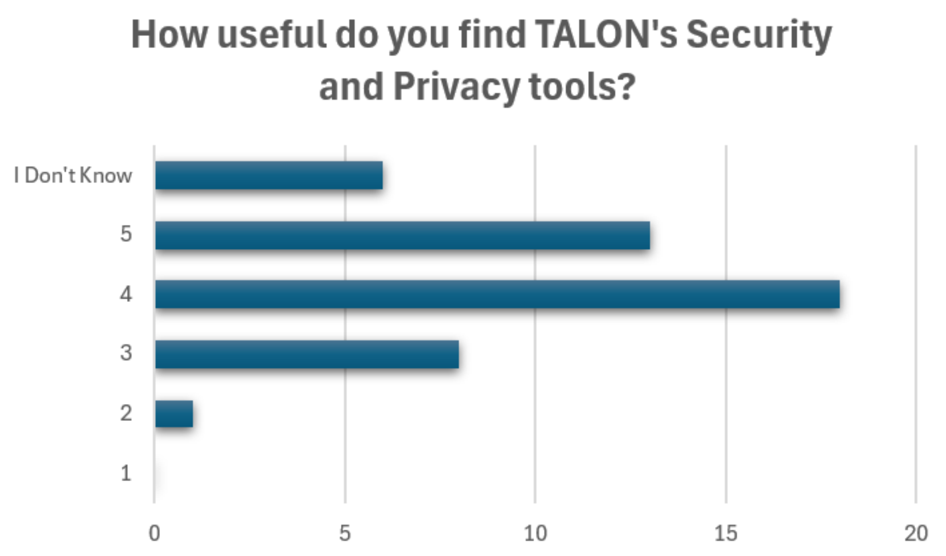

Figure 25: How useful do you find TALON's Security and Privacy tools? (1=Not at all, 5=Extremely)

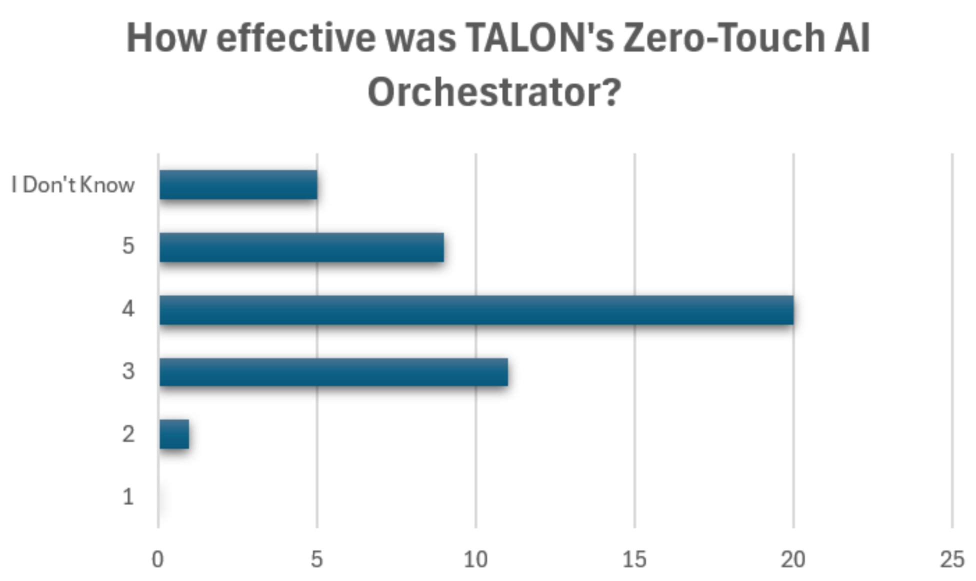

Figure 26: How effective was TALON's Zero-Touch AI Orchestrator? (1=Not at all, 5=Extremely)

In the last set of questions, the users have been asked to provide a rating as an answer with a scale ranging from 1-10, as well as N/A (Not Applicable) and I Don't Know, in order for respondents to provide more fine-grained distinctions on their answers for the survey. The answers are presented in the graphs below:

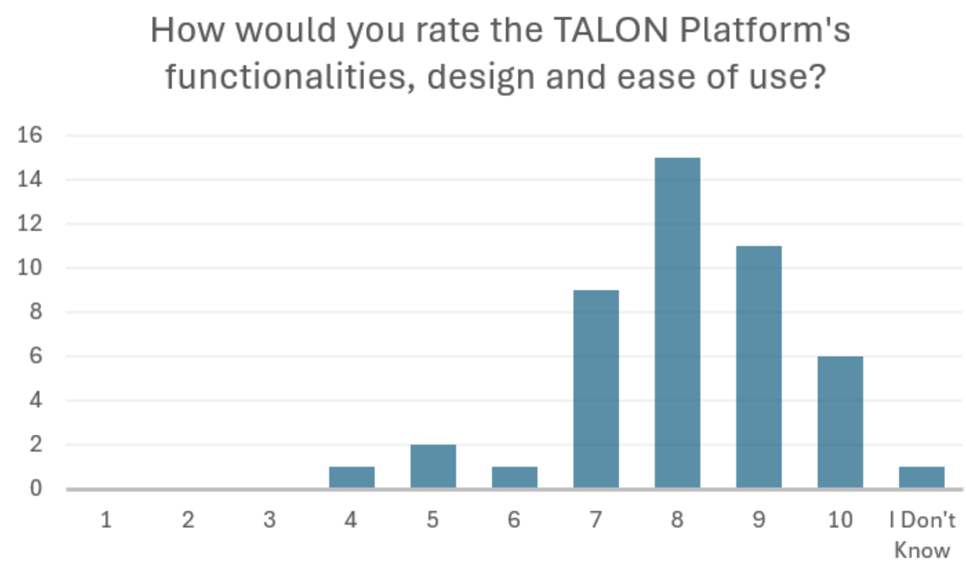

Figure 27: How would you rate the TALON Platform's functionalities, design and ease of use? (1=Extremely poor. 10=Excellent)

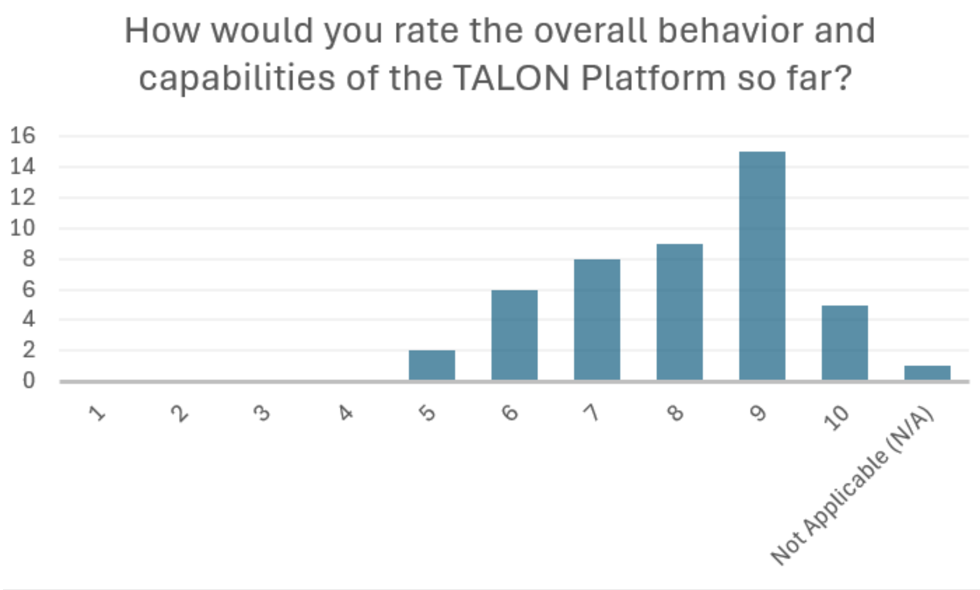

Figure 28: How would you rate the overall behaviour and capabilities of the TALON Platform so far? (1=Extremely poor. 10=Excellent)

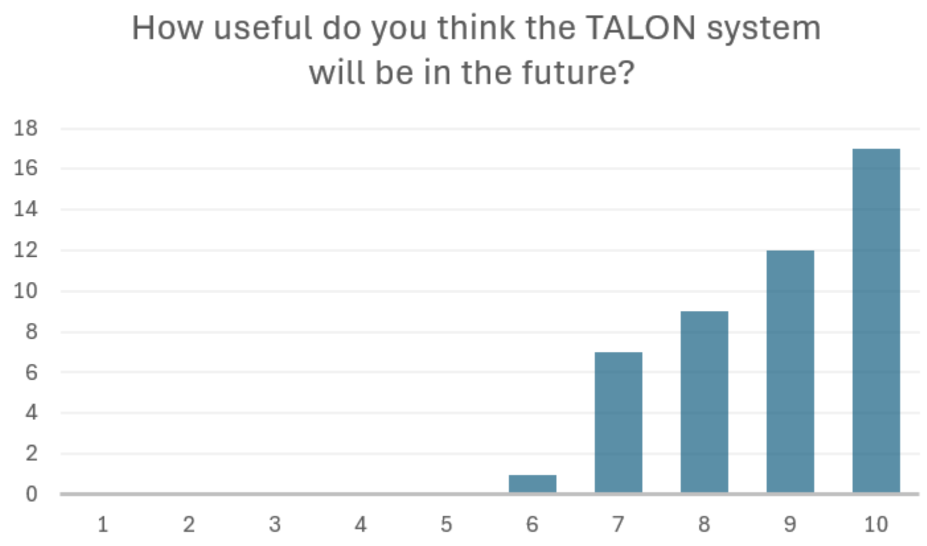

Figure 29: How useful do you think the TALON system will be in the future? (1=Extremely poor. 10=Excellent)

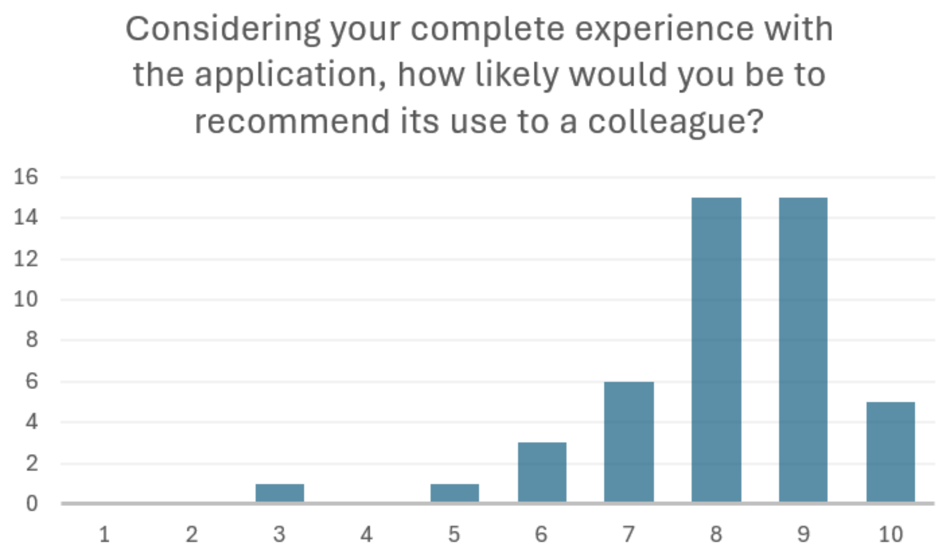

Figure 30: Considering your complete experience with the application, how likely would you be to recommend its use to a colleague? (1=Extremely poor. 10=Excellent)

In order to provide a more detailed analysis on the survey answers, the same questions and answers have been grouped and analysed per industry. The corresponding pie charts per industry are presented in the following sections.

5.1 Pilot: UATV's coordination

UATV's coordination pilot counts 14 responses, the highest response rate for the overall survey. Results are strongly positive, with high levels of agreement and confidence across most questions. The corresponding graphs are presented below.

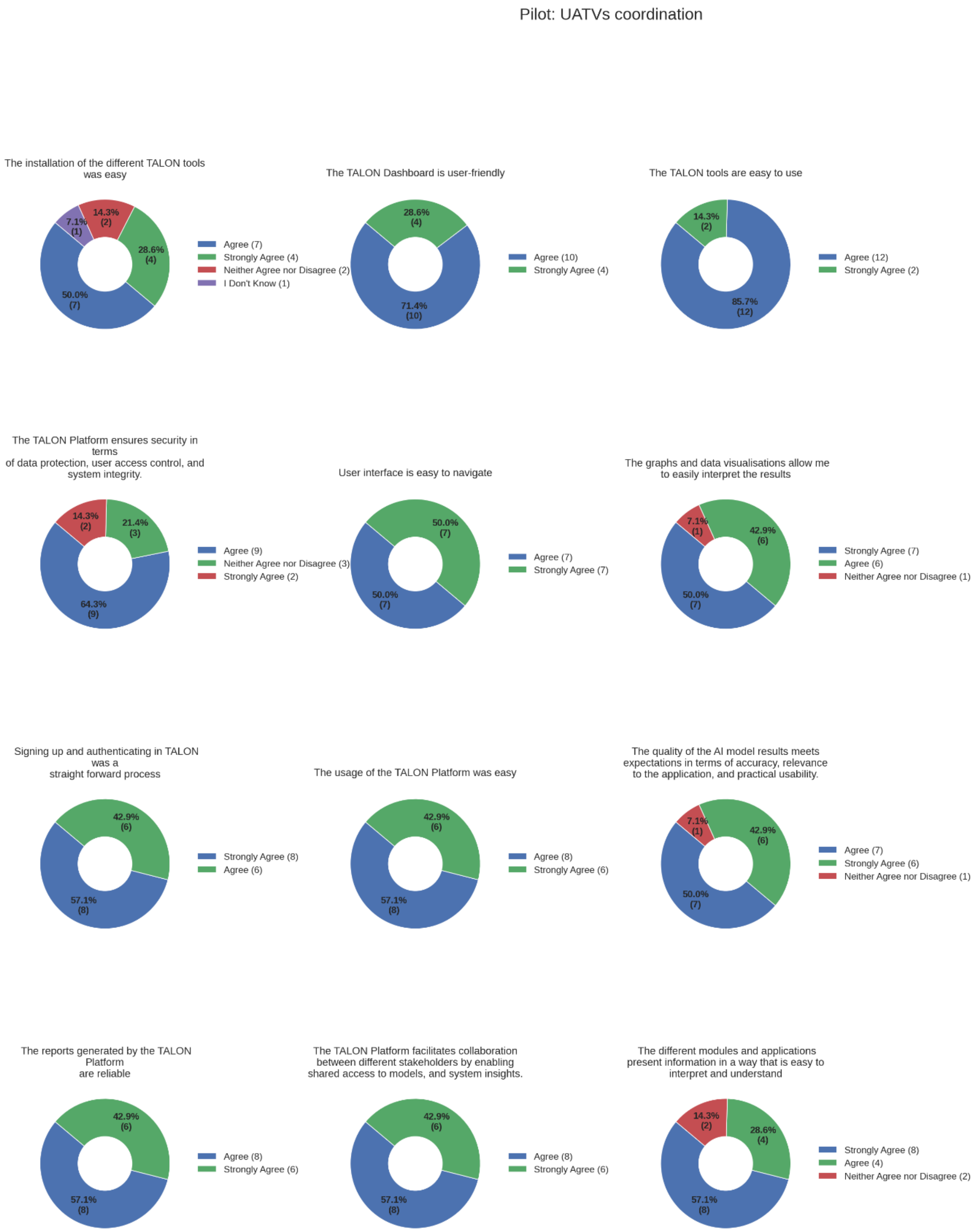

Figure 31: UATVs coordination survey results (1/2)

Figure 32: UATVs coordination survey results (2/2)

5.2 Pilot: Human Robot Collaboration

Human Robot Collaboration pilot has been answered by 11 participants and their respective answers as graphs are shown below:

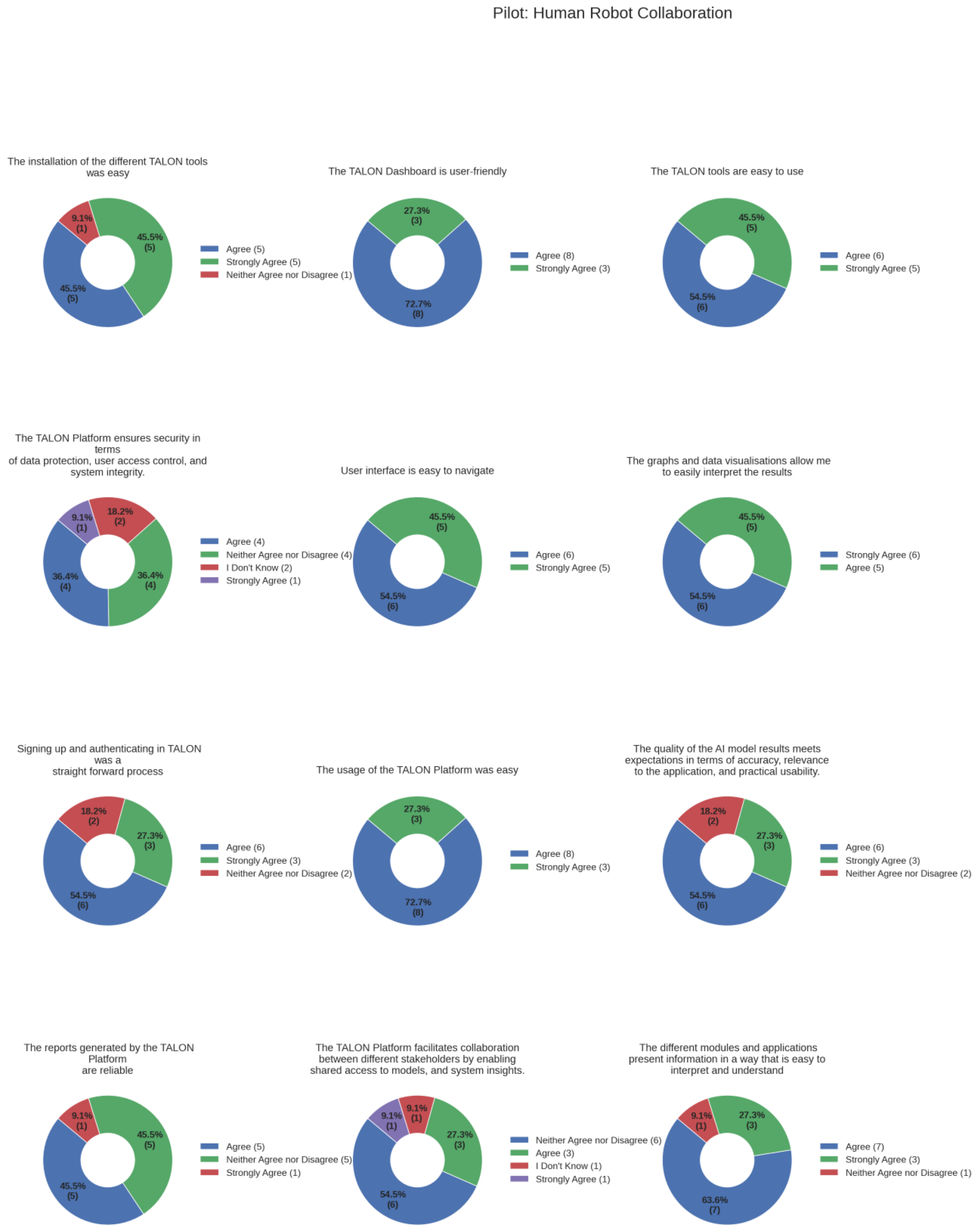

Figure 33: Human Robot Collaboration survey results (1/2)

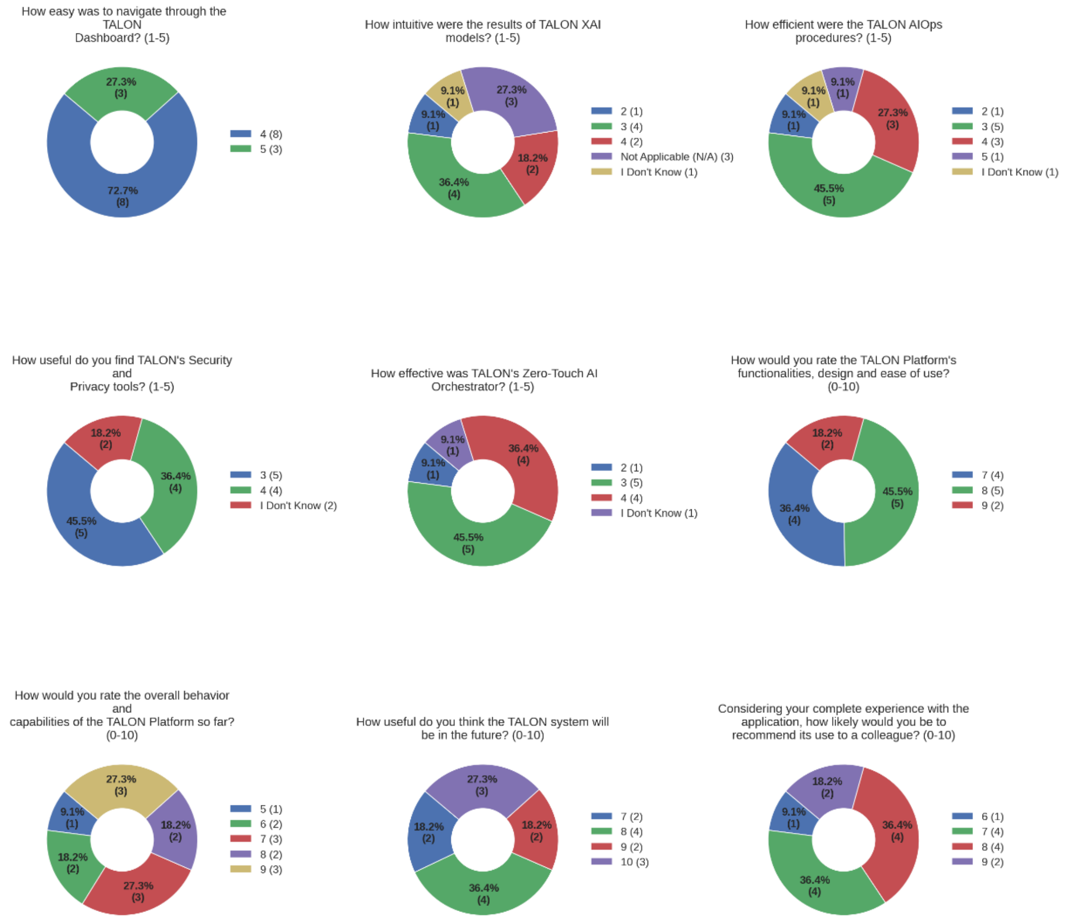

Figure 34: Human Robot Collaboration survey results (2/2)

5.3 Pilot: AR/VR for Training & Maintenance

AR/VR for training & Maintenance pilot has had 10 respondents, and the results to the survey are presented as percentages on the figures below:

Figure 35: AR/VR for training & maintenance survey results (1/2)

Figure 36: AR/VR for training & maintenance survey results (2/2)

5.4 Pilot: I5.0 Automation and Planning

Finally, I5.0 Automation and Planning pilot was represented by 8 participants. The corresponding results as percentages can be found in the below figures:

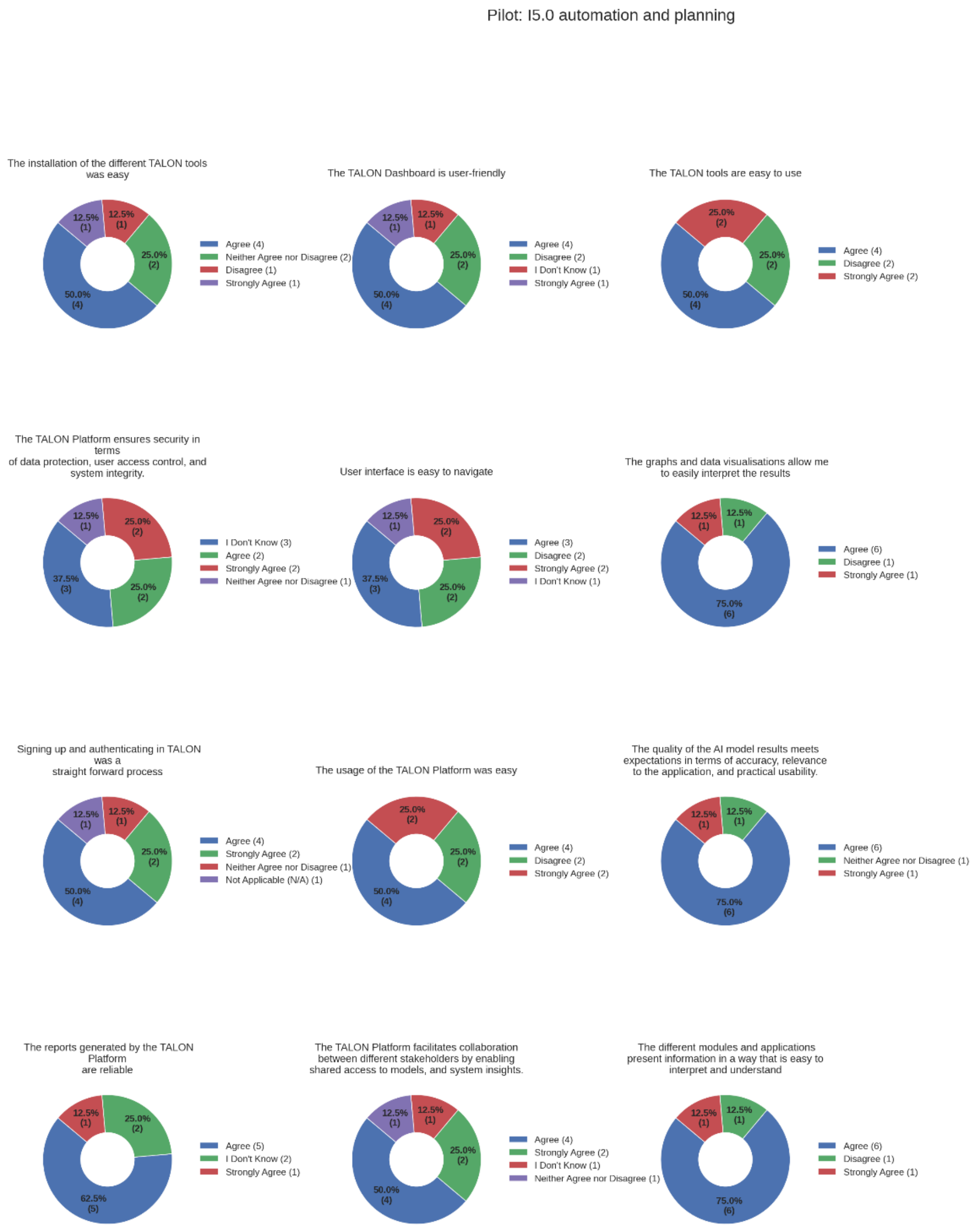

Figure 37: I5.0 automation and planning survey results (1/2)

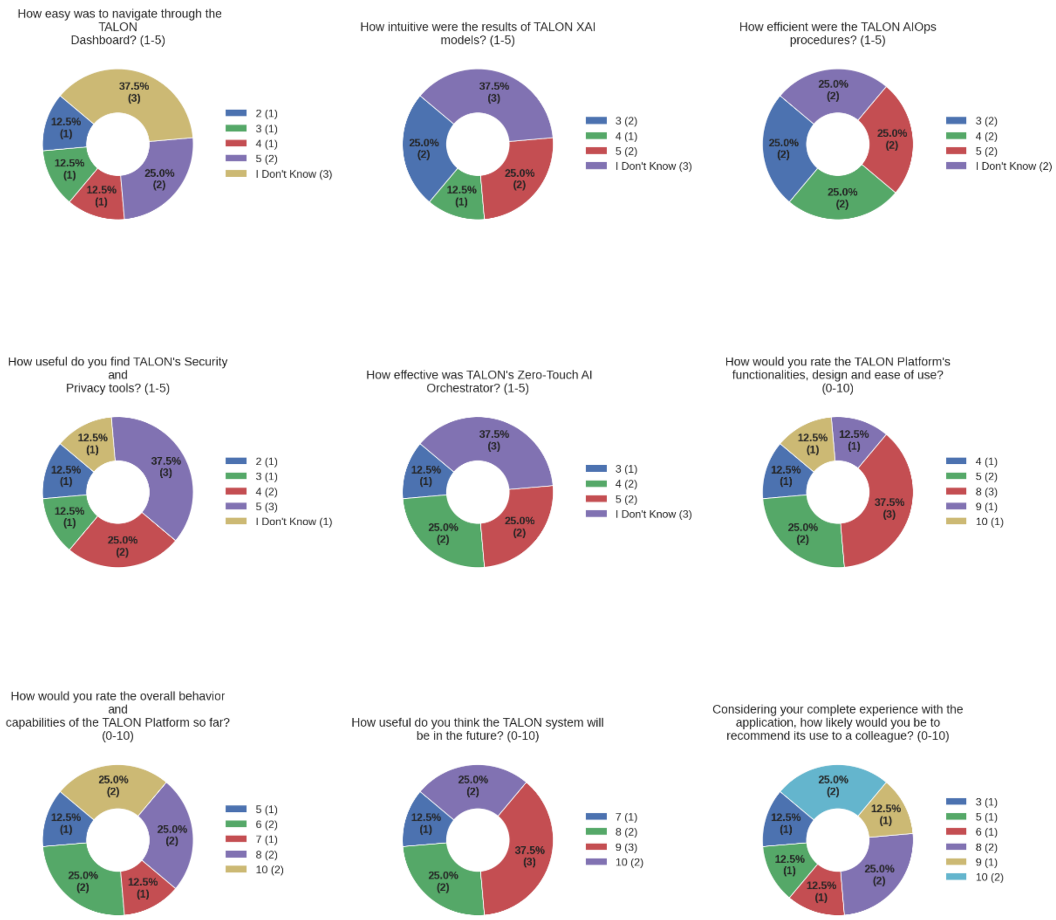

Figure 38: I5.0 automation and planning survey results (2/2)

6 Conclusion

Deliverable D5.4 - *Final TALON Platform Setup, Operation, Continuous Integration & Maintenance Report* documents the work carried out in the TALON Pilot sites, undertaken under Task 5.3: Automatic UATV Coordination, Task 5.4: I5.0 Automation & Planning, Task 5.5: AR/VR for Training & Maintenance and Task 5.6: Human Robot Collaboration.

This document presents the final results of the TALON evaluation activities conducted across all demonstrators, together with a comprehensive assessment of the technical components comprising the TALON Integrated Platform.

The consolidated evaluation results demonstrate that:

- The technical outputs of the project are of high quality, have been extensively validated, and are characterised by a high degree of security and interoperability.
- The final evaluation of the Key Performance Indicators (KPIs) confirms the effectiveness of the TALON Platform across the various demonstration scenarios.
- The end-user evaluation highlights that the project results are not only technically robust but also enjoy a high level of acceptance among the intended user community.

D5.4 is the final deliverable of a series of four deliverables related to the "Work Package 5: Integration, Validation & Demonstration" of TALON, which marks the successful completion of Milestone 8 (MS8), related to the successful final release of TALON Integrated platform in M36 concluding demonstration and validation activities undertaken in TALON project

Funded by
the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101070181