ICT-56-2020 "Next Generation Internet of Things"

Grant Agreement number: 957218

IntelIIoT

Deliverable D5.6
Validation & Evaluation (final version)

Deliverable release date	
Authors	1. Siemens: Andreas Zirkler, Andreas Ziller, Arne Broering 2. EURECOM: Jérôme Härri 3. AAU: Beatriz Soret, Lam Nguyen 4. UOULU: Sumudu Samarakoon 5. TTC: Martijn Rooker, Gerald Fritz 6. TUC/TSI: Andreas Brokalakis, Charalampos Savvakos, Thomas Kyriakakis, Fotini Karetsi 7. Philips: Thomas Bastiaansen, Bas Flaton 8. SANL: Konstantinos Fysarakis, Antonios Paragioudakis, Sotirios Michail 9. HSG: Simon Mayer, Jérémy Lemée 10. HOLO: Carina Pamminger, Nour Fendri 11. AVL: Holger Burkhardt, Wolfgang Hollerweger 12. PAGNI: Maria Marketou, Ioannis Anastasiou 13. VIDAVO: Alexis Fourlis 14. STARTUPC: Maren Lesche
Editors	Andreas Brokalakis (TSI), Charalampos Savvakos (TSI)
Reviewers	Sumudu Samarakoon (UOULU), Konstantinos Fysarakis (SANL)
Approved by	PTC Members: (Vivek Kulkarni, Konstantinos Fysarakis, Sumudu Samarakoon, Beatriz Soret, Arne Bröring, Maren Lesche) PCC Members: (Vivek Kulkarni, Jérôme Härri, Beatriz Soret, Mehdi Bennis, Martijn Rooker, Sotiris Ioannidis, Anca Bucur, Georgios Spanoudakis, Simon Mayer, Carina Pamminger, Holger Burkhardt, Maren Lesche, Georgios Kochiadakis)
Status of the Document	Final
Version	1.0
Dissemination level	Public

Table of Contents

Acronyms and Definitions

Acronym	Definition
5G NR	5G New Radio
AAA	Authentication, Authorization, and Accounting
AAMAS	Autonomous Agents and MultiAgent Systems
AR	Augmented Reality
CAGR	Compound Annual Growth Rate
D2D	Device to Device
DLT	Distributed Ledger Technology
DRL	Deep Reinforcement Learning
ECG	Electrocardiogram
ETSI	European Telecommunications Standards Institute
FL	Federated Learning
GUI	Graphical User Interface
HIL	Human In the Loop
HMD	Head-Mounted Display
IAKM	Infrastructure Assisted Knowledge Management
ICT	Information and Communication Technology
IDE	Integrated Development Environment
IDS	Intrusion Detection System
IPFS	InterPlanetary File System
KPI	Key Performance Indicator
LL-MEC	Low Latency Multi-access Edge Computing
MAS	Multi-Agent System
MEC	Multi-access Edge Computing
ML	Machine Learning
mMTC	massive Machine-Type Communication
MTD	Moving Target Defence
NB-IoT	Narrow Band IoT
NFV	Network Function Virtualization
NG IoT	Next Generation Internet of Things
OEM	Original Equipment Manufacturer
RAN	Radio Access Network
RL	Reinforcement Learning
RPC	Remote Procedure Call
RPM	Remote Patient Monitoring
SGD	Stochastic Gradient Decent
TSN	Time Sensitive Networking
URLLC	Ultra Reliable Low Latency Communication
WoT	Web of Things

1. INTRODUCTION

Deliverable D5.6 is the final report of the validation and evaluation procedures, reflecting the final outcome of the development and integration processes that led to the final version of the IntellIoT framework and its deployment in the project's three use-cases.

As such, the deliverable reports on the results of the final activities of Task 5.3, that aimed to provide the validation and evaluation procedures carried within the project, and to verify whether the goals set – e.g., scenarios, requirements and KPIs, as defined in Deliverables D2.5 – "Technology Analysis & Requirements Specification (final version)" and D2.6 – "High level architecture (final version)" – have been met, and at which level of success. Task 5.3 naturally follows the work carried out within Task 5.1 (Integration and Implementation) and Task 5.2 (Deployment, Testing and Demonstration) that build on the technical developments of Work Packages (WPs) 3 and 4.

To provide the above, Section 2 outlines the validation and evaluation methodology. It highlights the experimentation processes and testbeds and defines the different parts of the validation process concluding a validation and evaluation template that summarises all this information. This template is then used in Sections 3 to 7 to report on the validation and evaluation procedures and outcome of each functional, non-functional, and technical requirement as well as KPIs related to the IntellIoT framework, use-cases and overall project ambitions.

It should be noted that all requirements and KPIs are traceable to their specifications in Deliverables D2.5 and D2.6, elaborating & refining upon the original ones (as defined within IntellIoT's Description of Action). Compared to the initial version of this deliverable (Deliverable D5.3), several requirements have been removed or consolidated as a result of the development activities (Section 2.2 provides a detailed list of the requirements that have been removed or consolidated along with a reasoning behind this action). Therefore, these requirements have been removed from this document.

2. VALIDATION & EVALUATION METHODOLOGY

The validation and evaluation methodology that is followed within IntelIIoT focuses on establishing a clear baseline for the assessment of the different components and the overall framework that is developed within the project. The three use-cases implemented in the project provide a real-world testing environment to validate and assess the use and successful applicability of each component and the framework as a whole. Beyond that, the results are also evaluated against the overall KPI targets set by the consortium.

In this section, the methodology for carrying out the validation and evaluation process is going to be outlined. The description of this methodology will set the scene for the assessment report in the next sections of the document.

It should be noted that since the first version of this document (D5.3) was released, IntelIIoT has carried out two cycles of Open Calls reaching out to external entities with the aim to extend and/or use the different components developed within the project. As a result, the validation and evaluation procedures reported within this document also include feedback and results that have emerged beyond the project's three main use cases, via the inclusion of the work carried out within said Open Calls.

The validation and evaluation methodology is composed of a few distinguished steps. First, the experimentation testbeds and methodologies that have been used to test and validate IntelIIoT components (each one alone or their integrated versions with other components) are going to be presented. Then the validation and evaluation strategy is going to be described, as well as the evaluation scope. To help structure the process, an evaluation template is derived and is used throughout the next sections to provide the validation and evaluation report for each requirement and KPI. The overall process is depicted in Figure 1.

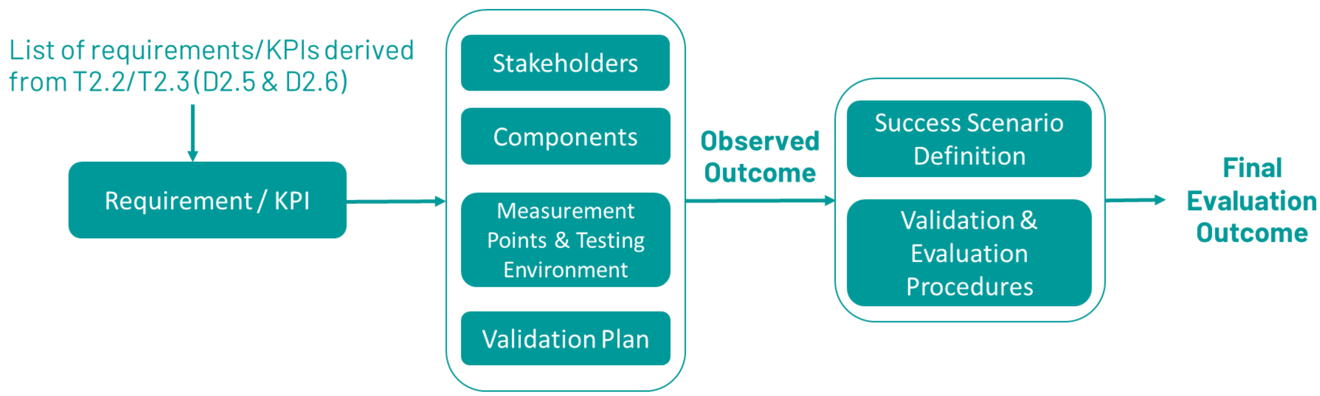


Figure 1. Validation and Evaluation Procedure for each Requirement or KPI.

2.1 Experimentation Methodology

2.1.1 REQUIREMENTS

To validate and evaluate the overall IntelIIoT framework and its successful application in the three use-cases chosen, three main sets of requirements have been set: functional, non-functional and technical (the first two documented in deliverables D2.5, and the latter documented in D2.6). On top of this primary categorization, the requirements have been associated with either the general IntelIIoT framework (i.e., those requirements that are applicable and common for all use cases and constitute basic framework functionality) or the three use cases (i.e., requirement that are specifically derived from - and pertain to - the particularities of each use case).

2.1.2 RESPONSIBLE STAKEHOLDERS (OWNERS)

For each requirement that must be validated and evaluated, the involved stakeholders have been defined. These include the primary developers of the associated components and in some cases the main users of the services that these components provide.

2.1.3 EVALUATION SCOPE

For each requirement or KPI that has to be evaluated, the scope provides an insight on where this requirement originated from. In the context of IntellIoT, the different evaluation scopes are:

1. **IntellIoT use cases**: the requirement or KPI arises from the specifics of the different use cases (Agriculture, Healthcare, Manufacturing) that are considered within IntellIoT.
2. **IntellIoT framework**: the requirement or KPI arises from the development of the overall IntellIoT framework and generally to all (or most) use cases or is fundamental for the functioning of a coherent framework.
3. **IntellIoT project**: the requirement or KPI arises from the general goals and ambitions of the project.

2.1.4 VALIDATION AND EVALUATION STRATEGY

Two main approaches are to be followed for validation & evaluation: (i) the first one is ***demonstration-based***, and aims to validate and evaluate requirements and KPIs through setting up a proper testing procedure and validating functionality or performance in a simulated or actual testbed; the second is ***review-based***, and focuses on validating and evaluating requirements and KPIs against well-defined documentation and scientific publications. The subsections below provide more details on each of these two approaches.

2.1.4.1 DEMONSTRATION-BASED VALIDATION

When applying demonstration-based validation to a requirement or KPI a series of (one or multiple) test cases must be specified. For each test case, the following need to be provided:

1. **Test case workflow specification:** this describes the actors involved in the test case and the procedures and practices followed as well as possible inputs and data sets used.
2. **Test acceptance criteria:** these determine when test results indicate that a certain test has been passed (fully or partially).
3. **Test results:** these are the actual results upon which a successful/failed outcome and evaluation can be based upon.

The testing infrastructure, both hardware and software -wise, is described in Deliverable D5.5.

2.1.4.2 REVIEW-BASED ANALYSIS

A review-based analysis is used to validate a requirement or KPI by providing reference to IntellIoT's documentation (e.g., project's deliverables, component specifications, scientific publications). This review-based analysis approach can also be referred to as document-based analysis. When using this approach, explanations must be provided why the given references are adequate to satisfy the specific requirement or KPI.

A review-based analysis should consist of the following:

1. References to technical documentation or peer-reviewed scientific publications that explain some technical feature or certain characteristics;
2. evaluation, i.e., mapping of the provided documentation to the requirement or KPI in scope, and;
3. justification why the documentation provides proof for the validation of the requirement or KPI

2.1.5 ASSESSMENT

Using the results of either the review-based or demonstration-based validation and evaluation procedure, the assessment report summarises the outcome of the process. For requirements that are related to functionality issues, possible outcomes may be in the form of ***Pass/Fail***, although in some cases ***Partial*** success may be claimed if: (i) the outcome only manages to cover a certain part of the requirement (e.g., not all packets are registered in a network monitoring scenario), or; (ii) the requirements/KPI has not been thoroughly tested in adequate conditions and therefore functional testing and validation is of limited scope.

2.1.6 VALIDATION AND EVALUATION TEMPLATE

A common template has been devised for evaluating the requirements and KPIs of the IntellIoT project. It is presented in the following table along with a proper description for each field.

Table 1. Validation and Evaluation Template

Requirement/KPI ID	*Identifier for the specific requirement or KPI under evaluation. This can be traced back directly to deliverables D2.5 and D2.6 where the requirement and KPIs are specified*
Owner	*Stakeholders of the specific requirement or KPI*
Evaluation Scope	*Evaluation Scope (IntellIoT use cases, IntellIoT framework, IntellIoT project, Open Calls)*
Components(s)	*A list with the components relevant to this requirement*
Measurement Point(s)	*This may be a list of components, servers, etc where measurements to validate the specific requirement are expected to be taken. For example, in the case of measuring the number of packets that the IDS component is able to handle, this has to be carried out at the IDS client in the devices that are hosted*
Goal (Success Criteria)	*A small description of what is expected as a success scenario*
Validation Strategy	*Demonstration / Review*
Validation Plan	*A description of the process that is going to be followed depending on the validation strategy to validate the requirement / KPI.*
Validation Assessment	*[Pass/Fail/Partial]* *Upon executing the validation and evaluation plan, the assessment discusses on the results and outcomes of the process. For purely functional requirements, a pass/fail/partial success assessment is expected. The same can be expected for non-functional and technical requirement, however more details may need to be provided.*

2.2 Changes with respect to Deliverable 5.3

As mentioned in the introductory section, this document reflects the changes that have been done during the second (and final) cycle of the project. As part of the activities of the second cycle, several initial requirements defined during Cycle 1, have either been removed or consolidated with other requirements. As such, these initial requirements will not be further reported in the present deliverable, and subsection 2.2.1 provides the reasoning behind each of these changes.

In the same context, the continued development activities, framework and use case refinements, have also led to the introduction of some new requirements. These have been introduced in the updated material of Deliverables D2.5 and D2.6, while a list of these requirements is also reported in subsection 2.2.2 below.

2.2.1 REQUIREMENTS REMOVED DURING CYCLE 2

Table 2. Requirements removed during Cycle 2 of the project along with the justifications for their removal.

Requirement ID	Justification
GFR.2	Redundant with GFR.1.
GFR.8	After rework of the architecture (see TR.15), HyperMAS polls the state of machines and robots. This simplifies the architecture because we do not need REST client and server on each device.
GFR.9	After rework of the architecture (see TR.15), HyperMAS polls the state of machines and robots. This eases the architecture because we do not need REST client and server on each device.
GFR.13	Feature requirement removed as it was not needed.
GNFR.9	Redundant, as the authentication has already taken place to be able to access that part of the SAP platform.
GNFR.14	Vague description, covered by other more specific requirements GNFR.13, FR.UC1#7, FR.UC3#7, UCNFR.6.
TR.12	Redundant with TR.24.
TR.14	This is covered by the adherence to W3C WoT TD in GNFR.18.
TR.16	After rework of the architecture (see TR.15), HyperMAS polls the state of machines and robots. This eases the architecture because we do not need REST client and server on each device.
TR.17	After rework of the architecture (see TR.15), HyperMAS polls the state of machines and robots. This eases the architecture because we do not need REST client and server on each device.
TR.18	After rework of the architecture (see TR.15), HyperMAS polls the state of machines and robots. This eases the architecture because we do not need REST client and server on each device.
TR.22	No responsibility on data collection and labelling and thus, training new models is not possible.
TR.23	Merged with TR.10 and TR.11, where the existence and correct operation of DLT and smart contract has been done in the local testbed during cycle 1, and pending to be validated in the final demos of the framework in cycle 2.
TR.29	After optimization of the architecture towards minimum reaction time, HIL Application directly controls the robot, without intermediation of the robot controller.
TR.31	Covered by TR.28.
TR.33	Covered by TR.28.
TR.35	Only ASCII input. Vector Graphics is computed internally.
TR.37	Covered by GNFR.18.
TR.39	The placing of the workpiece is now pre-programmed.
TR.45	Covered by TR.3.
TR.47	After rework of the architecture (see TR.15), HyperMAS polls the state of machines and robots. This eases the architecture because we do not need REST client and server on each device.

TR.48	Not needed anymore. First tests have shown that the camera at the robot is not sufficient to get overview pictures, because the focus of this camera cannot be configured so that both gripper and viewpoint images are sharp. Therefore, we now have separate cameras for overview pictures.
TR.73	The specific requirement is completely covered by TR.72 and is deemed redundant.
TR.74	For service level guarantees health state based on delay and jitter is sufficient. No app interface is needed. Unclear how further insights via in-app diagnostics further help.
TR.94	Only text is engraved on the workpiece.
FR.UC1#4	Tractor is returning its status (ETCS.state = error in case of obstacle). What is done with that information is not part of tractor interface.
FR.UC1#6	New interface: tractor returns error state, any component can listen to tractor status and request control over the tractor.
FR.UC1#9	Merged with FR.UC1#8 based on the changes made on UC1.2.1
FR.UC1#13	Requirement removed as it was not useful for the user.
FR.UC1#15	All passive interaction with tractors was considered "not useful" for the operator. By utilising the MQTT message broker, the AI could trigger the request automatically. Having a tool that could allow connection to tractor by the user, would have negatively impacted the test case.
FR.UC1#16	Redundant because of FR.UC1#13.
FR.UC1#17	No drone provider found during Open Calls, therefore the requirement was removed.
FR.UC1#19	This was never used. Kept only to not break the numbering.
FR.UC1#20	The indirect control of the tractor is beyond the scope of the project.
FR.UC1#21	The indirect control of the tractor is beyond the scope of the project.
FR.UC2#9	There are general requirements covering the existence of DLT clients and registration of transactions from the devices, i.e., GFR.5, TR.10, and they apply to the whole IntelIIoT framework and not only UC2.
FR.UC3#3	This FR is not required anymore since the engraver does not need the image but only ASCII of the to-be-engraved text. The corresponding vector graphics are rendered internally.
FR.UC3#4	Already covered in TR.30
UCNFR.1	This non-functional requirement is removed, as it is already covered by GNFR.15.
UCNFR.8	After rework of the architecture, HyperMAS polls the state of machines and robots. Thus, periodic state information updates do not exist anymore and TR.9 thus don't need to be acknowledged.

2.2.2 REQUIREMENTS INTRODUCED DURING CYCLE 2

During the second cycle of the project, a number of technical requirements (TR#) have been added to the previously identified ones. Furthermore, an additional Non-Functional Requirement specific to a Use Case has been introduced. The following list provides the identification codes of the newly introduced requirements.

TR.40a	TR.80	TR.84	TR.88	TR.92	TR.96
TR.40b	TR.81	TR.85	TR.89	TR.93	TR.97
TR.44a	TR.82	TR.86	TR.90	TR.94	UCNFR.9
TR.79	TR.83	TR.87	TR.91	TR.95	

3. VALIDATION AND EVALUATION OF GENERAL REQUIREMENTS

This section provides the validation and evaluation report for the requirements that are associated with the IntellIoT framework (general functional and non-functional requirements, within subsections 0 and 0, respectively) as specified in Deliverable D2.5. Furthermore, the technical requirements specified in Deliverable D2.6 are also assessed in this Section (subsection 0).

3.1 Validation and Evaluation of General Functional Requirements

Requirement	GFR.1
Owner	HSG
Evaluation Scope	IntellIoT Framework
Components(s)	Agent IDE, HyperMAS infrastructure, Organization IDE
Measurement Point(s)	HyperMAS Infrastructure
Goal (Success Criteria)	Agents and their organizations can be created by the Agent IDE and run on the HyperMAS Infrastructure
Validation Strategy	Demonstration
Validation Plan	1. Create agent procedural knowledge with the Agent IDE. 2. Deploy the agent on the HyperMAS Infrastructure. 3. Create multi-agent organization with the Organization IDE. 4. Deploy the multi-agent organization to the HyperMAS Infrastructure.
Validation Assessment	**Partial**. The integration of the generation of the procedural knowledge and deployment fully works on a former version of the HyperMAS Infrastructure but needs to be adjusted for the current version. The technical video for UC3 shows how to create an agent with the Agent IDE. The creation of the multi-agent organization with the Organization IDE works but the deployment to the HyperMAS Infrastructure remains to be done.

Requirement	GFR.3
Owner	HSG
Evaluation Scope	IntellIoT Framework
Components(s)	HyperMAS Infrastructure, Deployed Agents, Goal Specification Interface
Measurement Point(s)	Goal Specification Interface
Goal (Success Criteria)	The user receives an acknowledgment that a goal has been received.
Validation Strategy	Demonstration
Validation Plan	1. User defines a goal in the Goal Specification Interface 2. User sends this goal to the HyperMAS infrastructure 3. Agents in the HyperMAS infrastructure starts working towards the goal 4. The user receives an acknowledgment via the Goal Specification Interface
Validation Assessment	**Pass**. The process described in the validation plan has been successfully completed as part of the testing procedures and working prototype demonstrations during the integration phase.

Requirement	GFR.4
Owner	HSG
Evaluation Scope	IntellIoT Framework
Components(s)	HyperMAS Infrastructure
Measurement Point(s)	Goal Specification Interface
Goal (Success Criteria)	The user receives an acknowledgment that a goal has been reached.
Validation Strategy	Demonstration
Validation Plan	1. User defines a goal in the Goal Specification Interface 2. User sends this goal to the HyperMAS infrastructure 3. Agents in the HyperMAS infrastructure starts working towards the goal 4. Agents in the HyperMAS infrastructure reach the goal 5. The user receives an acknowledgment via the Goal Specification Interface
Validation Assessment	**Pass.** The process described in the validation plan has been successfully completed as part of the testing procedures and working prototype demonstrations during the integration phase.

Requirement	GFR.5
Owner	AAU
Evaluation Scope	IntellIoT Framework
Components(s)	DLT client
Measurement Point(s)	DLT client
Goal (Success Criteria)	The DLT client receives data and transactions to be recorded from sensors or other gathering devices
Validation Strategy	Demonstration
Validation Plan	1. Devices send data/transactions to the DLT client 2. The data/transactions are recorded in the ledger 3. The new data/transactions are available and can be seen in the demo
Validation Assessment	**Pass**. The creation and generation of new transactions by the DLT client has been integrated and demonstrated in the three UCs, the content of the DLT at the end of the demos can be fetched and a visualization tool is also available in the case of UC3.

Requirement	GFR.6
Owner	AAU
Evaluation Scope	IntellIoT Framework
Components(s)	DLT clients and manager
Measurement Point(s)	DLT ledger
Goal (Success Criteria)	Check behaviour of the smart contract
Validation Strategy	Demonstration

Validation Plan	1. DLT record operational data 2. DLT operational data is compared to smart contract terms to check if there is a mismatch
Validation Assessment	**Pass**. The DLT and smart contract framework has been tested and validated in the three UCs.

Requirement	GFR.7
Owner	AAU
Evaluation Scope	IntellIoT Framework
Components(s)	DLT clients and manager
Measurement Point(s)	DLT client
Goal (Success Criteria)	Inform human operator about a mismatch data/smart contract
Validation Strategy	Demonstration
Validation Plan	1. DLT record operational data 2. DLT operational data is compared to smart contract terms 3. Show potential mismatch in the human operator DLT client
Validation Assessment	**Pass**. This information is available in the blockchain and can be fetched in a text file and checked by the human operator.

Requirement	GFR.10
Owner	HSG
Evaluation Scope	IntellIoT Framework
Components(s)	HyperMAS Infrastructure, Deployed Agents, Goal Specification Interface
Measurement Point(s)	Goal Specification Interface
Goal (Success Criteria)	A human user can specify, review, and rephrase goals for the HyperMAS
Validation Strategy	Demonstration
Validation Plan	1. User defines a goal in the Goal Specification Interface 2. User sends this goal to the HyperMAS infrastructure 3. Agents in the HyperMAS infrastructure starts working towards the goal 4. User reviews the sent goal(s) using the Goal Specification Interface 5. User changes goal(s) using the Goal Specification Interface 6. Agents in the HyperMAS Infrastructure drop the old goal and work on the changed goal.
Validation Assessment	**Partial**. The user can change the goal if a notification that the sent goal is not valid is received. However, a valid goal sent to the HyperMAS will necessarily be executed.

Requirement	GFR.11
Owner	UOULU, Philips
Evaluation Scope	UC1 (Scenario2-2.1), UC2, UC3 (Scenario-2)

Components(s)	Agriculture AI, Healthcare AI, Manufacturing AI
Measurement Point(s)	Inferred decisions at AI components
Goal (Success Criteria)	To define a rule that clarifies the likelihood of failures at AI inference
Validation Strategy	Demonstration
Validation Plan	Artificially create a scenario outside of training scenario
Validation Assessment	**Pass** for UC1, the validation is assessed with a laboratory demo setting as well as integrated in the tractor controller. The model indicates both its ability to detect and bypass the obstacle and to notify when inference fails. **Not applicable** for UC2, the collected data was not sufficient to get a proper model performance on a validation dataset (see Appendix A: Healthcare Use Case – Data analysis and model development). Therefore, no attempt was made to implement any mechanisms that measure model confidence. The model selected for UC2 was a regression model, which makes it less evident what model confidence would mean (contrary to classifiers). In our case, we would have kept track of predictions vs actual measurements over time, as that is easy to do given the timeseries input data. **Pass** for UC3, the AI model is integrated and tested in the demo setup. The confidence of the AI inference is computed based on the likely of detecting the usable area and feasible grab point coordinates.

Requirement	GFR.12
Owner	SANL, TSI
Evaluation Scope	UC3 (Key scene: UC.3.3.9 and related scenes)
Components(s)	SAP, MTD, TSN
Measurement Point(s)	SAP for alerting & action triggering. Malicious actor endpoint to verify isolation.
Goal (Success Criteria)	Upon triggering of isolation by operator, to verify that the malicious actor no longer has access.
Validation Strategy	Demonstration
Validation Plan	1. Trigger malicious activity through "compromised" client device 2. Operator observes alert & requests lockout 3. SAP triggers the MTD 4. MTD interacts with TSN to isolate malicious host 5. Confirm that malicious entity is isolated
Validation Assessment	**Pass**. Demonstration of full sequence in UC3.

Requirement	GFR.14
Owner	HOLO, Siemens
Evaluation Scope	UC3
Components(s)	HIL-AR-Application, HIL Service, HoloLens 2
Measurement Point(s)	Pop-up on the HL2 client application to operators offering help.
Goal (Success Criteria)	HIL Service informing operators offering help that the task is no longer existent/being taken care of.

Validation Strategy	Demonstration
Validation Plan	1. Operator offers help to HIL Service 2. HIL Service responds that help is no longer needed 3. Operator receives a pop-up containing the response of the HIL Service
Validation Assessment	**Pass**. A popup is shown to all connected operators. An operator can take over by clicking the "OK" button. If another user already accepted the task, other users will see a pop-up message informing them that help is no longer needed.

Requirement	GFR.15
Owner	HOLO
Evaluation Scope	UC3
Components(s)	HIL-AR-Application, Robot Controller, HoloLens 2
Measurement Point(s)	Pop-up event messages on the HL2
Goal (Success Criteria)	Correct events messages are being displayed on to the Operator running the application on the HL2
Validation Strategy	Demonstration
Validation Plan	1. Operator controls the robot 2. Controller sends back event state messages 3. Application on the HL2 displays these events to the Operator
Validation Assessment	**Pass**. Event messages are being displayed to the operator.

Requirement	GFR.16
Owner	HOLO
Evaluation Scope	UC3
Components(s)	HIL-AR-Application, UR5 Robot
Measurement Point(s)	HIL-AR-Application, UR5 Robot
Goal (Success Criteria)	Receiving and recording information from the machine
Validation Strategy	Demonstration
Validation Plan	1. Operator in control of the Robot 2. Robot sends relevant state information 3. HIL-AR-Application receives and records information
Validation Assessment	**Pass.** Information is received and can be recorded through screen recording tools, e.g. OBS.

Requirement	GFR.17
Owner	HOLO, Siemens
Evaluation Scope	UC3
Components(s)	HIL-AR-Application, Robot System

Measurement Point(s)	WebRTC Connection successfully established Stream latency
Goal (Success Criteria)	Frames received by the HIL-AR-Application
Validation Strategy	Demonstration
Validation Plan	1. HIL-AR-Application and robot system establish a P2P WebRTC connection. 2. System sends encoded frames through the interface to the HIL-AR-Application 3. HIL-AR-Application receives frames
Validation Assessment	**Pass.** HIL Application successfully receives and displays frames in-app. Robot movement and video streams from the two cameras on the robot and above the workpiece table can be seen in the HoloLens. Video streams appear synchronized and with low delay.

Requirement	GFR.18
Owner	HOLO
Evaluation Scope	UC3
Components(s)	HIL-AR-Application, HoloLens 2
Measurement Point(s)	Stream latency
Goal (Success Criteria)	Display frames on the HL2 to Operator
Validation Strategy	Demonstration
Validation Plan	1. HIL-AR-Application decodes frames and displays them in the AR environment 2. Operator sees the environment through the HL2
Validation Assessment	**Pass.** Frames are visualised in the AR glasses, while the app is running on a laptop. Video streams and robot movement appear synchronized and with low delay.

Requirement	GFR.19
Owner	HOLO
Evaluation Scope	UC3
Components(s)	HIL-AR-Application, HIL Service, AI
Measurement Point(s)	AI, HIL-AR-Application
Goal (Success Criteria)	Control is returned to the AI
Validation Strategy	Demonstration
Validation Plan	1. Operator sends a request to the HIL Service to end the session 2. HIL Service informs AI that the Operator is finished and to resume control of the robot
Validation Assessment	**Pass.** The operator can return control to the AI through menu interaction. This also closes the connection between the robot and the HIL Application.

Requirement	GFR.20
Owner	HOLO

Evaluation Scope	UC1 & UC3
Components(s)	HIL Application (user interface + VR/AR glasses), Interface that allows for machine (robot or tractor) control
Measurement Point(s)	HIL Application, Machine (robot or tractor)
Goal (Success Criteria)	To support the (semi-)autonomous system by remotely controlling the machine that doesn't know how to handle the current situation
Validation Strategy	Demo
Validation Plan	1. HIL Application connects to the interface that allows for sending machine controls 2. Operator sends control commands using HMIs (Stylus (UC3) or VR controllers (UC1)) to the machine through said interface in an attempt to solve the problem
Validation Assessment	**Pass.** The operator can control the machines (tractor and robot) through their respective interfaces.

Requirement	GFR.21
Owner	UOULU, Philips
Evaluation Scope	IntellIoT Framework, All use cases.
Components(s)	Local and Global AI components
Measurement Point(s)	Retraining of AI components
Goal (Success Criteria)	To continuously improve AI models
Validation Strategy	Review
Validation Plan	For UC1 and UC3: 1. Creating several artificial situations where AI escalates to human operator 2. Retraining is triggered by multiple escalations to human Similar procedure as mentioned here will be utilized. For UC2: Start with a poorly trained model and two phones that run the local AI with a good amount of data. Create a training configuration on the global AI. Start the apps on the two phones and validate that after a few minutes, new weights have been uploaded for aggregation. After aggregation, evaluation showed that the model accuracy has improved.
Validation Assessment	**UC2 pass:** tested with known well-training model on fashion MNIST use case in on-desk setup and with simple model on actual demo setup. **UC1 and UC3 pass**: In simulated settings, local AI raining is carried out with a portion of datasets allowing them to result in inaccurate inference. Followed by several hundreds of inference failures, retraining of local AI is triggered to result in a model with improved inference accuracy. This experiment is repeated several iterations to illustrate the diminishing need of human interventions. The results are disseminated in D3.6.

Requirement	GFR.22
Owner	SANL

Evaluation Scope	UC2 & UC3
Components(s)	SAP
Measurement Point(s)	SAP Incident Response component
Goal (Success Criteria)	To provide the means to specify parameters regarding response of security modules to detected events / incidents.
Validation Strategy	Demonstration
Validation Plan	1. Verify that user interface provides the needed means to specify the response of the security modules. 2. Verify that specified behaviour is followed by affected security modules.
Validation Assessment	**Pass**. Demonstration of above plan steps in UC2 & UC3.

3.2 Validation and Evaluation of General Non-Functional Requirements

Requirement	GNFR.1
Owner	SANL
Evaluation Scope	Event Captors
Components(s)	Event Captors (secondary: Trust Broker, SAP, IDS, MTD)
Measurement Point(s)	Event Captors, Trust Broker
Goal (Success Criteria)	Ensure that identified relevant trust-related events are captured by corresponding Event Captors
Validation Strategy	Demonstration
Validation Plan	1. Trigger each identified relevant trust-related event 2. Verify developed event captors accurately capture event 3. Verify event evidence is relayed to SAP through Trust Broker
Validation Assessment	**Pass**. Demonstration of above steps in UC1, UC2 & UC3.

Requirement	GNFR.2
Owner	SANL
Evaluation Scope	SAP Monitor
Components(s)	SAP
Measurement Point(s)	SAP Monitor, Event Captors
Goal (Success Criteria)	Ensure that the raw events received from Event Captors are ingested by SAP and the appropriate monitoring rules are triggered.
Validation Strategy	Demonstration
Validation Plan	1. Follow validation process defined in GNFR.1 2. For each event, verify corresponding monitoring rule is triggered 3. Verify that satisfaction / violation of rule is correctly reasoned based on raw evidence received
Validation Assessment	**Pass**. Demonstration of above steps in UC1 & UC2.

Requirement	GNFR.3
Owner	SANL
Evaluation Scope	Trust Broker
Components(s)	Trust Broker
Measurement Point(s)	SAP, IDS, MTD
Goal (Success Criteria)	Verify that QoS parameters are configurable through Trust Broker, and that these are enforced/respected in message exchanges of Trust components interacting via Broker
Validation Strategy	Demonstration
Validation Plan	For each QoS mode available in Trust Broker: 1. Set QoS mode 2. Trigger test messages 3. Verify test messages respect QoS parameters for each of the involved endpoints.
Validation Assessment	**Pass**. Demonstration of above steps in UC3.

Requirement	GNFR.4
Owner	TUC/TSI, SANL
Evaluation Scope	Event Captors, Trust IDS
Components(s)	Event Captors, Trust IDS (secondary: SAP)
Measurement Point(s)	Event Captors, SAP Monitor, Trust IDS
Goal (Success Criteria)	Verify that all types of malicious activity covered in the 3 different Trust scenarios (one scenario per UC) are captured by the Event Captors and/or the Trust IDS system.
Validation Strategy	Demonstration
Validation Plan	For each UC: 1. Execute all Trust Scenario key scenes 2. Verify malicious activity is captured by Event Captor and/or the Trust IDS (depending on type of activity) 3. Verify that SAP's monitoring component receives, ingests & reasons upon malicious activity evidence, based on corresponding rule
Validation Assessment	**Pass**. Successful demonstration for all types of malicious activity covered in Cycle 1 & 2 demos across UCs.

Requirement	GNFR.5
Owner	TSI, SANL
Evaluation Scope	Security Components
Components(s)	MTDs, (Secondary: Trust IDS, Event Captors, SAP)
Measurement Point(s)	MTDs, Trust IDS, SAP, Malicious hosts
Goal (Success Criteria)	Verify that misbehaving nodes are detected and automatically excluded or isolated.
Validation Strategy	Demonstration

Validation Plan	Use Case 1: Use nping to generate excessive traffic. Trust IDS should detect that activity and issue a warning towards the MTD Server. MTD Server isolates the offending node by changing the network configuration for the rest nodes. It also shares this information with SAP. Use Case 2: Detection of malicious botnet activity on patient systems. Detection of malicious ransomware activity on clinician systems. Mitigation of both. Use Case 3: Detection & mitigation of malicious remote robot operator (through the proper API call to the TSN controller to isolate the offending node).
Validation Assessment	**Pass.** Demonstrated the detection and exclusion of misbehaving nodes on each Use Case, as described in the validation plan above. For each Use Case, depending on the specific components deployed, different ways have been used to achieve the desired outcome.

Requirement	GNFR.6
Owner	SANL
Evaluation Scope	SAP Front end
Components(s)	SAP
Measurement Point(s)	SAP
Goal (Success Criteria)	Ensure that the SAP Graphical User Interface (GUI) has the capacity to provide a real time view of the security posture of the protected deployment
Validation Strategy	Review
Validation Plan	Design and specification review to ensure the SAP front-end integrates the specified capabilities.
Validation Assessment	**Pass**. Design of SAP user interface provides the needed capabilities to provide real time view of the security posture in terms of GUI features (also demonstrated across UCs).

Requirement	GNFR.7
Owner	SANL
Evaluation Scope	SAP
Components(s)	SAP (Secondary: Event Captors, Trust IDS)
Measurement Point(s)	SAP front end (GUI)
Goal (Success Criteria)	Ensure that the SAP GUI
Validation Strategy	Demonstration
Validation Plan	For each UC: 1. Execute Trust scenario and associated key scenes 2. For each security-pertinent incident monitored, verify that the appropriate monitoring rule is defined.

	3. Verify that satisfactions and violations of rules are tracked and visualised in real-time for the operator to view.
Validation Assessment	**Pass**. Successful validation for Cycle 1 & 2 key scenes in relevant demonstrations, across Use Cases.

Requirement	GNFR.8
Owner	SANL
Evaluation Scope	SAP
Components(s)	SAP (Secondary: MTDs)
Measurement Point(s)	SAP (Incident Response module)
Goal (Success Criteria)	Verify that the operator can define responses to incidents and that these are enforced through the trust components.
Validation Strategy	Demonstration
Validation Plan	1. Define a set of defence / incident response actions through IR Playbooks, within the corresponding SAP capability 2. Deploy the defined IR flow 3. Emulate malicious activity to trigger Playbook 4. Verify defined response action is relayed to and executed through corresponding trust component (e.g., MTD).
Validation Assessment	**Pass.** Demonstrated in UC1 (notification only), UC2 (automated mitigation) & UC3 (mitigation with human-in-the-loop).

Requirement	GNFR.10
Owner	SANL
Evaluation Scope	SAP
Components(s)	SAP (Secondary: Event Captors)
Measurement Point(s)	SAP Monitor
Goal (Success Criteria)	The SAP Monitor should provide the means to monitor uptime.
Validation Strategy	Review
Validation Plan	Design and documentation review to ensure the monitoring component of SAP provides the means to monitor satisfaction and violation of uptime requirements.
Validation Assessment	**Pass**. Capability verified as reported in Deliverable D4.8.

Requirement	GNFR.11
Owner	SANL, TUC/TSI
Evaluation Scope	All trust components
Components(s)	All trust components
Measurement Point(s)	All trust components
Goal (Success Criteria)	Ensure the secure configuration and bootstrapping of all trust components

Validation Strategy	Review
Validation Plan	Review of trust components' deployment and configuration practices to verify that properties hold according to best known practices.
Validation Assessment	**Pass**. Manual review & verification across UC testbeds. Trust components use a secure broker for their communication. All their traffic is encrypted using TLS. The relevant keys and certificates are distributed offline (before system bootstrapping) for each component during deployment, and they are loaded via their configuration files. See also TR.21.

Requirement	GNFR.12
Owner	HOLO, AVL
Evaluation Scope	UC1
Components(s)	Tractor, 5G, HIL
Measurement Point(s)	VR glasses
Goal (Success Criteria)	Image resolution is good enough to be visible in the VR space but not too high that causes delay and FPS issues.
Validation Strategy	Demonstration
Validation Plan	Check image at VR glasses composed from tractor cameras
Validation Assessment	**Pass.** During the integration meetings, it was confirmed that the resolution of the video feed is high enough to fully understand the environment and the situation, but not too high to cause noticeable delays with the transmission. A smooth interaction with the tractor was confirmed.

Requirement	GNFR.13
Owner	Siemens, TTC, AVL
Evaluation Scope	UC1 (Key Scene 2.1); UC3 (general)
Components(s)	Tractor, Robot
Measurement Point(s)	Tractor, Robot
Goal (Success Criteria)	The component (i.e., tractor or robot) doesn't move anymore after it is in a situation which it doesn't know how to handle, or which is configured to be unsafe and waits for human input.
Validation Strategy	Demonstration
Validation Plan	Command the tractor or robot to move to an unsafe position, observe that it stops without damage.
Validation Assessment	**Pass** **UC1**: The tractor utilizes the way points to follow its defined trajectory until it encounters an obstacle which it doesn't know how to handle. **UC3**: Robot controller only accepts grab spots in a defined area which is known to be safe. Remote movement via stylus is limited in force and speed, so that the robot changes to protective stop mode before it damages anything.

Requirement	GNFR.15

Owner	TTC, EURECOM, Siemens, AAU
Evaluation Scope	UC1 (Key Scene 2.4); UC3 (Key Scene 2.1)
Components(s)	Tractor, Robot, VR/AR Infrastructure, 5G Infrastructure
Measurement Point(s)	Tractor, Robot
Goal (Success Criteria)	To have the component (i.e., tractor or robot) to react in less than 30ms E2E (10ms over wireless) to the control commands that the human operator is providing to the component.
Validation Strategy	Demonstration
Validation Plan	1. Human operator is controlling the robot via the VR system 2. Speed of tractor is (potentially) reduced 3. Measurement of delay between command given by the human and the tractor actual performing the command.
Validation Assessment	**Partial**. Test labs showed 5G can support 10ms over wireless URLL, but no end-2-end evaluation yet.

Requirement	GNFR.16
Owner	HSG
Evaluation Scope	IntellIoT Framework
Components(s)	Agent IDE, HyperMAS Infrastructure
Measurement Point(s)	Agent IDE
Goal (Success Criteria)	Domain experts (DE) can successfully program and deploy Agents using the Agent IDE
Validation Strategy	Review (User Study)
Validation Plan	Extended user study based on this paper[1] 1. DE is given realistic scenario 2. DE successfully creates agent procedural knowledge with the Agent IDE 3. DE successfully creates multi-agent organization with the Agent IDE 4. DE successfully deploys both agent procedural knowledge and multi-agent organization to the HyperMAS Infrastructure Steps are performed with enough users to yield significant results.
Validation Assessment	**Partial**. The implementation for all steps has been performed but step 3 has not been evaluated with a user study.

Requirement	GNFR.17
Owner	HSG

[1] S. Burattini, A. Ricci, S. Mayer, D. Vachtsevanou, J. Lemee, A. Ciortea and A. Croatti, "Agent-Oriented Visual Programming for the Web of Things," in *Proceedings of the International Workshop on Engineering Multi-agent Systems*, 2022

Evaluation Scope	IntellIoT Framework
Components(s)	HyperMAS Infrastructure, Agent IDE
Measurement Point(s)	HyperMAS Infrastructure
Goal (Success Criteria)	It is possible to extend the system with additional services and changed Agent procedural knowledge at run time.
Validation Strategy	Demonstration
Validation Plan	1. User monitors Agents running in the HyperMAS Infrastructure using the Agent IDE 2. Additional service(s) are displayed by the Agent IDE 3. User modifies Agent procedural knowledge to use an additional service and redeploys
Validation Assessment	**Partial**. Steps 1 and 2 work. The extension of the agent procedural knowledge at run time is work in progress.

Requirement	GNFR.18
Owner	HSG
Evaluation Scope	IntellIoT Framework
Components(s)	Agent IDE, HyperMAS Infrastructure
Measurement Point(s)	HyperMAS Infrastructure and Agent IDE
Goal (Success Criteria)	Agent IDE and HyperMAS Infrastructure are compatible with any service providing a TD and input/output schema specification.
Validation Strategy	Review (Standards-based Argument)
Validation Plan	1. Each service provides its TD 2. The TDs are exposed on the HyperMAS Infrastructure 3. The user can access the TDs from the Agent IDE and can create agents based on the TDs.
Validation Assessment	**Pass**. Since all services adhere to the W3C WoT TD standard and since the HyperMAS Infrastructure and Agent IDE use the same standard.

3.3 Validation and Evaluation of Technical Requirements

Requirement	TR.1
Owner	SIEMENS, HSG
Evaluation Scope	IntellIoT Framework
Components(s)	All services implementing an API for any machine (tractor, robot, etc) that are managed by the Edge Orchestrator
Measurement Point(s)	HyperMAS, Edge Orchestrator
Goal (Success Criteria)	Successful registration on HyperMAS
Validation Strategy	Demonstration

Validation Plan	1. The Edge Orchestrator initiates registration of an unbound service API to the HyperMAS Infrastructure by providing the needed W3C WoT Thing Description Template (TDT) 2. HyperMAS has knowledge of the unbound service's interface (but not the form binding)
Validation Assessment	**Pass.** All relevant services were successfully initiated and used in the UC3 demo setup.

Requirement	TR.2
Owner	All service owners
Evaluation Scope	IntellIoT Framework
Components(s)	All services implementing an API for any machine (tractor, robot, etc) that are not managed by the Edge Orchestrator
Measurement Point(s)	HyperMAS
Goal (Success Criteria)	Successful registration on HyperMAS
Validation Strategy	Demonstration
Validation Plan	1. A service initiates registration to HyperMAS by providing the needed W3C WoT Thing Description (TD) 2. HyperMAS can use the API described in the TD by using the specified protocol bindings and security definition
Validation Assessment	**Pass.** This has been demonstrated completely for UC3 during integration meetings at Siemens in Munich and for UC1 during integration meetings at AVL in Steyr.

Requirement	TR.3
Owner	HSG, Siemens
Evaluation Scope	UC3
Components(s)	Hypermedia MAS, Edge Orchestrator
Measurement Point(s)	Hypermedia MAS, Edge Orchestrator
Goal (Success Criteria)	When an agent intends to use the service of an edge app that is represented through a W3C WoT TDT, upon a create orchestration request, the Edge Orchestrator instantiates the service and registers a protocol bound W3C WoT TD at the HyperMAS.
Validation Strategy	Demonstration
Validation Plan	1. Service stub is registered to Hypermedia MAS only through a W3C WoT TDT 2. Domain Expert programs agent behaviour using the Agent IDE. This behaviour includes usage of the service stub. One Agent that implements this behaviour is deployed to the Hypermedia MAS 3. The Agent attempts to task the service. Since the TDT does not provide a form binding, the Agent instead contacts the Edge Orchestrator and requests the TD 4. The Edge Orchestrator instantiates the service and returns a TD that specifies the service interface in a form binding 5. The Agent successfully uses the newly bound service
Validation Assessment	**Partial**. Features to support using TDTs have been implemented in the HyperMAS Infrastructure and the Agent IDE. However, the communication between the agent and the

	Edge Orchestrator has been conceptually but not practically evaluated due to priority on UC3 demonstrator stabilization.

Requirement	TR.4
Owner	All service owners
Evaluation Scope	IntellIoT Framework
Components(s)	All services implementing an API for any machine (tractor, robot, etc)
Measurement Point(s)	HyperMAS and all services that integrate with it
Goal (Success Criteria)	Services can be polled for their health status
Validation Strategy	Demonstration
Validation Plan	1. HyperMAS issues a HTTP GET request towards registered services 2. Healthy services respond with HTTP 200 OK
Validation Assessment	**Pass.**

Requirement	TR.5
Owner	HSG
Evaluation Scope	IntellIoT Framework
Components(s)	HyperMAS
Measurement Point(s)	HyperMAS Infrastructure
Goal (Success Criteria)	Delegating a goal to the agent once it is received by the HyperMAS Infrastructure
Validation Strategy	Demonstration
Validation Plan	1. A goal is received by a manager agent in the HyperMAS 2. The manager agent selects the agent that can work on the goal 3. The manager agent sends the goal to this agent
Validation Assessment	**Partial**. The mechanism to support this has been implemented. However, during the integration meetings, only single agent deployment (where the manager agent and the executor agent are the same agent) have been considered for easiness of use.

Requirement	TR.6
Owner	HSG
Evaluation Scope	IntellIoT Framework
Components(s)	HyperMAS Infrastructure
Measurement Point(s)	HyperMAS Infrastructure
Goal (Success Criteria)	Agents send requests to services based on their TDs
Validation Strategy	Demonstration
Validation Plan	1. A service exposes its TD on the HyperMAS 2. An agent is programmed (with the Agent IDE) to use the service based on this TD 3. The agent runs on the HyperMAS and uses the service

Validation Assessment	**Pass.** This workflow has been fully demonstrated within our lab. During integration meetings, the process has been simplified so that the agents can be programmed beforehand (based on the known TDs).

Requirement	TR.7
Owner	All service owners
Evaluation Scope	IntellIoT Framework
Components(s)	All services implementing an API for any machine (tractor, robot, etc)
Measurement Point(s)	HyperMAS
Goal (Success Criteria)	Services acknowledge they have received tasks from Agents in the HyperMAS
Validation Strategy	Demonstration
Validation Plan	1. HyperMAS agent assigns a task to a registered service 2. Upon receival, the service acknowledges it received the task 3. The service initiates the assigned task
Validation Assessment	**Pass**. This has been tested during integration meeting in AVL Steyr for UC1 and Siemens Munich for UC3.

Requirement	TR.8
Owner	EURECOM
Evaluation Scope	UC1, UC3
Components(s)	IAKM
Measurement Point(s)	IAKM edge server; IAKM cloud server
Goal (Success Criteria)	MQTT message sent with the correct AI semantic to a IAKM server (cloud)
Validation Strategy	Demonstration
Validation Plan	1. IAKM edge server receives a request for AI from an IAKM client 2. IAKM edge queries its knowledge base for AI 3. If not available, IAKM edge server sends a request (MQTT) to the global IAKM cloud server (one single entity)
Validation Assessment	**Pass.** During the demonstration activities for UC1 and UC3, the IAKM cloud server successfully received the request, found the model and sent it to the IAKM edge server.

Requirement	TR.9
Owner	UOULU
Evaluation Scope	UC1.1.4
Components(s)	AI for Obstacle Bypassing
Measurement Point(s)	Communication from AI component to Vehicle Control
Goal (Success Criteria)	Issue control suggestions from the AI component to the Vehicle Control to manoeuvre around the obstacle
Validation Strategy	Demonstration

Validation Plan	1. AI component on the tractor receives sensing information 2. AI component infers control suggestions and sends to the Vehicle Control
Validation Assessment	**Pass.** Demonstrated in the lab setting and tractor controller and AI model communication is validated with integrating AI as a service on the tractor controller.

Requirement	TR.10
Owner	AAU
Evaluation Scope	IntellIoT framework
Components(s)	DLT
Measurement Point(s)	DLT clients
Goal (Success Criteria)	Make sure a DLT client exists and is operational
Validation Strategy	Demonstration
Validation Plan	Check a DLT client exists and is operational
Validation Assessment	**Pass**. During the demo activities, it was checked that the DLT clients and corresponding smart contracts were up and running.

Requirement	TR.11
Owner	AAU
Evaluation Scope	IntellIoT framework
Components(s)	DLT
Measurement Point(s)	DLT clients
Goal (Success Criteria)	Make sure a smart contract exists and is operational
Validation Strategy	Demo
Validation Plan	Check a smart contract exists and is operational
Validation Assessment	**Pass**. During the demo activities, it was checked that the DLT clients and corresponding smart contracts were up and running

Requirement	TR.13
Owner	AVL, TTC, HOLO
Evaluation Scope	Key Scene 2.3
Components(s)	Camera System on the Tractor
Measurement Point(s)	Tractor
Goal (Success Criteria)	The human operator receives the video stream from multiple cameras
Validation Strategy	Demonstration
Validation Plan	• Human operator connects to tractor • HIL application receives camera streams from the tractor • Multiple camera streams are composed to a single stream • Composed stream is displayed in VR glasses

Validation Assessment	**Pass.** The HIL Application receives the video stream of 4 different cameras placed around the tractor and visualizes them in the VR head set for the human operator.

Requirement	TR.15
Owner	Siemens
Evaluation Scope	UC3
Components(s)	• HyperMAS • Edge applications
Measurement Point(s)	HyperMAS
Goal (Success Criteria)	Update state information in HyperMAS
Validation Strategy	Demonstration
Validation Plan	Observe state information in HyperMAS
Validation Assessment	**Pass.** Replaces TR.16, TR.17 and TR.18. After rework of the architecture, HyperMAS polls the state of machines and robots instead of receiving regular updates. HyperMAS successfully polls the states of robot, milling machine, and laser engraver and uses the information for the control of the demo process in UC3.

Requirement	TR.19
Owner	AAU
Evaluation Scope	UC3
Components(s)	DLT client/manager
Measurement Point(s)	Smart contract
Goal (Success Criteria)	Transactions recorded correctly
Validation Strategy	Demo
Validation Plan	1. Receive transactions to be registered at smart contract application from the robot or other machines 2. Register transactions in the DLT
Validation Assessment	**Pass**. The creation and registration of transactions has been tested in the 3 UCs demos

Requirement	TR.20
Owner	AAU
Evaluation Scope	UC3
Components(s)	DLT client/manager
Measurement Point(s)	Smart contract
Goal (Success Criteria)	Smart contract registers to the Hypermedia MAS when started
Validation Strategy	Demo
Validation Plan	1. Run smart contract 2. Register to Hypermedia MAS

Validation Assessment	**Pass**.

Requirement	TR.21
Owner	TUC/TSI
Evaluation Scope	Use Cases 1, 3
Components(s)	• Trust-based IDS • MTD Client / Server • Broker • Security Assurance Platform
Measurement Point(s)	Broker
Goal (Success Criteria)	Establish secure communication between all security components.
Validation Strategy	Demonstration
Validation Plan	Initialise all security components, at all nodes of the test site as well as the central nodes that instantiate the Trust Broker and the SAP and configure the broker to orchestrate all communications through secure channels (TLS). If all the required network ports are properly configured, the communication between the components will be established, indicating additionally that the TLS keys are properly set and exchanged between the different nodes.
Validation Assessment	**Pass.** As part of the integration and demonstration activities (reported in Deliverables D5.4 and D5.5), the different security tools have been successfully connected with each other. The demonstrator scenes engaging the security tools that have been carried out within the UC1 and UC3 activities prove that the tools are properly initiated and configured and can collectively execute tasks that require a communication between them. Communication is handled using TLS, with self-signed certificates using the TLS-gen tool.

Requirement	TR.24
Owner	HSG
Evaluation Scope	IntellIoT Framework
Components(s)	HyperMAS Infrastructure
Measurement Point(s)	HyperMAS Infrastructure, DLT Service
Goal (Success Criteria)	Agents in the Hypermedia MAS notify the DLT Service when they generate a new task for a service/machine/artifact.
Validation Strategy	Demonstration
Validation Plan	1. Agent in the HyperMAS Infrastructure generates a new task for a service/machine/artifact 2. HyperMAS Infrastructure notifies DLT Service about this new task.
Validation Assessment	**Pass.** This has been demonstrated for UC3 during integration meetings at Siemens in Munich.

Requirement	TR.25

Owner	UOULU
Evaluation Scope	UC3 (Scenario 1 – Key scene 1.1)
Components(s)	Manufacturing AI
Measurement Point(s)	Manufacturing AI – Engrave area detection
Goal (Success Criteria)	To identify the engraving area and inform the HyperMAS
Validation Strategy	Simulation and Demo
Validation Plan	1. Place workpieces on the storage 2. HyperMAS request workpiece information 3. AI compute the engraving area and share the requested information In simulations, step 3 is validated. In demo, all three steps are validated.
Validation Assessment	**Pass**. Successfully sharing the workpiece information in the demo setting. For each request from the HyperMAS agent, the AI agent returns the usable information or grab spot information either as a JSON object or as a visual image depending on the request. Exemplary results are disseminated in D3.6.

Requirement	TR.26
Owner	EURECOM, AAU
Evaluation Scope	UC1; UC3
Components(s)	FlexRIC
Measurement Point(s)	FlexRIC, 5G Communication Manager (5GCM)
Goal (Success Criteria)	• FlexRIC provides ORAN-compliant Python APIs for the 5GCM to send FlexRIC instructions to the 5G RAN component • The 5GCM sends ORAN requests to FlexRIC be notified of a new required 5G Slice (i.e., new required resources) • 5G Communication Manager sends FlexRIC a request for a new 5G Slice • FlexRIC notifies the 5GCM of the request • Edge Controller increases local available resources for the new 5G slice
Validation Strategy	Demonstration
Validation Plan	The Edge orchestrator via the 5GMC needs to get the 5G RAN quality to know if a new 5G RAN needs to be deployed or more resources are required to be allocated. The 5GCM contacts the FlexRIC and if the 5G RAN KPI quality is lower than a threshold, the Edge orchestrator increases allocated resources.
Validation Assessment	**Pass**. Interfaces defined and tested with the 5GCM for DL slices.

Requirement	TR.27
Owner	EURECOM, AAU
Evaluation Scope	UC1, UC3
Components(s)	FlexRIC; 5G Communication Manager (5GCM)
Measurement Point(s)	FlexRIC; 5G Communication Manager (5GCM)
Goal (Success Criteria)	• FlexRIC provides ORAN-compliant Python APIs for the Edge Controller to send FlexRIC instructions to the 5G RAN component

	• The 5GCM sends ORAN requests to the FlexRIC (e.g., 5G RAN KPI quality) • FlexRIC periodically sends 5G-RAN data to Edge Controller
Validation Strategy	Demonstration
Validation Plan	5GCM requires to know the 5G KPI CQI (channel Quality Indicator); FlexRIC provides it for the 5G slice
Validation Assessment	**Pass**. Interfaces defined; Tested with the 5GCM.

Requirement	TR.28
Owner	HSG, Siemens
Evaluation Scope	UC3
Components(s)	HyperMAS, Edge Orchestrator
Measurement Point(s)	Edge Orchestrator
Goal (Success Criteria)	The edge app becomes instantiated in the edge infrastructure
Validation Strategy	Demonstration
Validation Plan	Trigger the agent to request a service
Validation Assessment	**Pass.** Agents can successfully request to use services during our integration meetings at Siemens in Munich.

Requirement	TR.30
Owner	HOLO
Evaluation Scope	UC3
Components(s)	Robot Controller, Interface, HIL-AR-Application
Measurement Point(s)	Robot Controller
Goal (Success Criteria)	Valid information about digital twin sent to the Robot Controller
Validation Strategy	Demonstration
Validation Plan	1. HIL-AR-Application has direct connection to the Robot Controller through the interface 2. HIL-AR-Application sends digital twin related information to the Robot Controller
Validation Assessment	**Pass.** Digital twin interaction input sent to real robot and controls its movements.

Requirement	TR.32
Owner	All service owners
Evaluation Scope	Use Case Integration
Components(s)	All edge applications
Measurement Point(s)	HyperMAS
Goal (Success Criteria)	Edge Applications can communicate with HyperMAS using their provided TD.
Validation Strategy	Demonstration

Validation Plan	- An Agent intends to task a service through that service's provided TD - The Agent can successfully task that service.
Validation Assessment	**Pass,** communication between agents and services is demonstrated in UC 1 and 3

Requirement	TR.34
Owner	UOULU
Evaluation Scope	UC3-Scenario 1: Key Scenes 1.1 & 1.2, UC3-Scenario 2: Key Scenes 2.1
Components(s)	Manufacturing AI
Measurement Point(s)	Manufacturing AI – Grab spot detection
Goal (Success Criteria)	Provide confidence of the AI grabbing decision
Validation Strategy	Demonstration
Validation Plan	In both simulation and demo: 1. Capture the image of the work piece 2. Compute grabbing spot and confidence
Validation Assessment	**Pass.** The confidence of the grabbing spot is calculated and returned as a percentage, which is to be used to determine human intervention. This is validated with the integrated AI model in the demos setting. The results are disseminated in D3.6 and D5.5.

Requirement	TR.36
Owner	HSG
Evaluation Scope	IntellIoT Framework
Components(s)	HyperMAS Infrastructure
Measurement Point(s)	Agent IDE Back End
Goal (Success Criteria)	The back end of the Agent IDE represents the interface to the HyperMAS
Validation Strategy	Demonstration
Validation Plan	1. When the back end of the Agent IDE receives a request from the Agent IDE to create an agent, it sends the request to the HyperMAS Infrastructure 2. The HyperMAS Infrastructure instantiates the agent
Validation Assessment	**Pass.** This has been tested within our lab. For integration meetings, a script was used instead to initialize the agents because it is more convenient for the project partners.

Requirement	TR.38
Owner	UOULU
Evaluation Scope	UC3 (Scenario 1: Key scene 1.1)
Components(s)	Manufacturing AI
Measurement Point(s)	Manufacturing AI: grabbing task
Goal (Success Criteria)	To successfully grab workpieces

Validation Strategy	Demonstration
Validation Plan	1. Position a workpiece on storage 2. AI to infer the grab spot with the confidence
Validation Assessment	**Pass**. AI infers the grabbing point information and the robot control receiving the grabbing coordinates. This is validated in the final demonstration and related results are shared in D5.5.

Requirement	TR.40
Owner	HOLO
Evaluation Scope	UC3
Components(s)	HIL-AR-Application, Robot Controller, Interface
Measurement Point(s)	Robot Controller
Goal (Success Criteria)	UR5 format-compatible instructions are received by the Robot Controller
Validation Strategy	Demonstration
Validation Plan	1. Operator generates movement commands 2. HIL-AR-Application transforms commands into UR5 compatible instructions 3. Instructions are sent through the interface to the Robot Controller 4. Robot Controller understands the received instructions
Validation Assessment	**Pass.** Robot understands received instructions and moves accordingly.

Requirement	TR.40a
Owner	HOLO, Siemens
Evaluation Scope	UC3
Components(s)	Robot Controller, HIL-AR-Application, Interface
Measurement Point(s)	Robot Controller, Robot
Goal (Success Criteria)	When AI requests robot controller to move the robot to a grab spot using x, y, alpha coordinates, robot controller shall translate the coordinates to TCP coordinates a UR5 compatible format.
Validation Strategy	Demonstration
Validation Plan	Send movement commands with x, y, alpha coordinates to the robot controller, check if the robot moves to the specified position.
Validation Assessment	**Pass**. The robot has shown to be able to reach all specified positions, i.e. to be able to grab workpieces from any location on the workpiece table.

Requirement	TR.40b
Owner	HOLO, Siemens
Evaluation Scope	UC3

Components(s)	HIL-AR-Application, Robot Controller, Interface
Measurement Point(s)	Robot Controller, Robot
Goal (Success Criteria)	When HyperMAS requests robot controller to move the robot to a named position, robot controller shall translate the position to TCP coordinates a UR5 compatible format.
Validation Strategy	Demonstration
Validation Plan	Send a sequence of movement commands with named positions to the robot controller, check if the robot moves to the specified position.
Validation Assessment	**Pass**. The robot has shown to be able to reach all named positions and to move between them in arbitrary order.

Requirement	TR.41
Owner	Siemens, EURECOM
Evaluation Scope	UC3
Components(s)	Edge orchestrator, TSN controller, communication resource manager
Measurement Point(s)	Edge orchestrator
Goal (Success Criteria)	Edge orchestrator can request communication services from TSN controller and communication resource manager
Validation Strategy	Demonstration
Validation Plan	Trigger the edge orchestrator to request for communication services from TSN controller and communication resource manager, e.g., by setting up a HIL action. Observe the communication reservation.
Validation Assessment	**Partial**. Reservation of network resources via TSN controller and 5G communication resource manager have been implemented and tested. Measurements have shown that both are working and improving the QoS of the reserved communication services. The reservation mechanism has not yet been integrated in the final demo setup because the 5G system gets unstable some minutes after a slice is configured.

Requirement	TR.42
Owner	Siemens
Evaluation Scope	UC 3
Components(s)	• HIL service • Edge orchestrator
Measurement Point(s)	Edge orchestrator
Goal (Success Criteria)	The edge orchestrator shall be able to instantiate the HIL service on any edge device. However, it is the preferred replication strategy, that exactly one instance is instantiate on one selected edge device.
Validation Strategy	Demonstration
Validation Plan	Trigger the edge orchestrator to instantiate a HIL service, observe proper instantiation.

Validation Assessment	**Pass.** The orchestration template for the HIL service includes the setting that the HIL service is instantiated exactly once on a defined edge device. The correct execution was verified.

Requirement	TR.43
Owner	Siemens, EURECOM
Evaluation Scope	UC3
Components(s)	• Edge Orchestrator • TSN Controller • Communication Resource Manager
Measurement Point(s)	Edge Orchestrator
Goal (Success Criteria)	Request will be responded with a success response
Validation Strategy	Demonstration of effect
Validation Plan	Trigger the Edge Orchestrator for a resource cancellation and observe the reservation.
Validation Assessment	**Partial**. Reservation of network resources via TSN controller and 5G communication resource manager have been implemented and tested. Measurements have shown that both are working and improving the QoS of the reserved communication services. The reservation mechanism has not been integrated in the final demo setup because the 5G system gets unstable some minutes after a slice is configured.

Requirement	TR.44
Owner	Siemens
Evaluation Scope	UC3
Components(s)	• HIL Service • HIL Application • MQTT Broker
Measurement Point(s)	HIL Application
Goal (Success Criteria)	Inform all HIL applications of a new service request
Validation Strategy	Demonstration
Validation Plan	Trigger a HIL service request at HIL Service, observe the request being received by operators.
Validation Assessment	**Pass.** We have demonstrated in UC3 that any HIL service request, triggered by the AI, is shown in the HoloLens of any connected operator.

Requirement	TR.44a
Owner	HOLO
Evaluation Scope	Use Case 3
Components(s)	• HIL Service • HIL Application

	• MQTT Broker
Measurement Point(s)	MQTT Message Broker application
Goal (Success Criteria)	When an operator wants to take over a help request, it shall send an operator takeover event containing its own IP address to HIL service.
Validation Strategy	Update to new notification scheme for HIL workers
Validation Plan	Does the user receive a message, e.g. a help request, when the AI runs into an issue, which results in the takeover of the robot arm? Possible features: • Accept help request and take over control; • Reject help request and inform HyperMAS that another request has to be sent; • Completed request, once the issue has been solved;
Validation Assessment	**Pass**. Through the MQTT Message broker application, the user will receive a popup window with the help request. Once confirmed the robot control will be given to the user. If rejected, the information is sent back to the HyperMAS and another user will be requested to help. Once an issue was fixed, the control is returned to the robot arm, by selecting "Completed".

Requirement	TR.46
Owner	AAU
Evaluation Scope	UC3
Components(s)	Smart contract
Measurement Point(s)	Smart contract
Goal (Success Criteria)	Smart contract application is aware of a new task to be performed
Validation Strategy	Demonstration
Validation Plan	1. New task created 2. Smart contract application is aware about this new task and resources to be used
Validation Assessment	**Pass.** The smart contract and transaction registration has been demonstrated in the integrated UC demos.

Requirement	TR.49
Owner	Siemens
Evaluation Scope	UC3
Components(s)	• HyperMAS • Robot Controller • UR5 Robot
Measurement Point(s)	Robot
Goal (Success Criteria)	Abstract the complex coordinate interface of UR5 and allow a simple, REST-based control of the robot using named poses.
Validation Strategy	Demonstration

Validation Plan	Request the robot to move to a named pose through robot controller's REST interface. Observe robot movement.
Validation Assessment	**Pass.** The robot is controlled via this REST interface in the UC3 demo.

Requirement	TR.50
Owner	HOLO
Evaluation Scope	UC3
Components(s)	Robot Controller, Interface, HIL-AR-Application
Measurement Point(s)	Robot Controller
Goal (Success Criteria)	Commands are successfully sent through the interface
Validation Strategy	Demonstration
Validation Plan	Ability to send commands through the TCP endpoint connecting the HIL-AR-Application and Robot Controller
Validation Assessment	**Pass.** Commands are sent through the TCP endpoint connection.

Requirement	TR.51
Owner	EURECOM
Evaluation Scope	UC1; UC3
Components(s)	IAKM edge server; MEC component
Measurement Point(s)	MEC component
Goal (Success Criteria)	• An IAKM edge server is created • The IAKM edge server sends an authentication message over HTTPS to the MEC component • The MEC validate the IAKM edge server credential
Validation Strategy	Demonstration
Validation Plan	The Edge controller creates an IAKM edge server with preloaded credentials; The IAKM authenticate to the MEC components; the IAKM successfully operates
Validation Assessment	**Partial**. Interfaces defined on LL-MEC, but will not be implemented. Following the deprecation of the 4G+ LL-MEC architecture on OAI, the replaced 5G FlexCN (during the project) did not have the same level of maturity and the security function have not required in the integration framework.

Requirement	TR.52
Owner	EURECOM
Evaluation Scope	UC1; UC3
Components(s)	IAKM edge server, MEC Component
Measurement Point(s)	IAKM edge server
Goal (Success Criteria)	• An IAKM edge server is created

	• The IAKM edge server sends an authentication message over HTTPS to the MEC component as for TR.51 • Upon validation of IAKM credential, the MEC component in turns sends its credential to the IAKM edge server • The IAKM edge server validates the MEC credentials
Validation Strategy	Demonstration
Validation Plan	Following TR.51, MEC sends its credential to the IAKM edge server, which checks its credentials.
Validation Assessment	**Partial**. Interfaces defined on LL-MEC, but will not be implemented. Following the deprecation of the 4G+ LL-MEC architecture on OAI, the replaced 5G FlexCN (during the project) did not have the same level of maturity and the security function have not required in the integration framework.

Requirement	TR.53
Owner	UOULU
Evaluation Scope	Local AI
Components(s)	Global and Local AI components
Measurement Point(s)	Retraining of local AI components
Goal (Success Criteria)	To exchange the status of the AI models with one another
Validation Strategy	Simulation
Validation Plan	Initiate re-train requests at local AI component
Validation Assessment	**Pass.** For UC1 and UC3, the validation is carried out in a simulated setting due to the need of hundreds of new interventions. This is emulated by partitioning the dataset into small portions and training the initial models to inherit low accuracy with a smaller fraction of data. After a predefined number of inference failures (100-200), the retraining is triggered. The results are disseminated in D3.6.

Requirement	TR.54
Owner	TUC/TSI, SANL
Evaluation Scope	IAKM
Components(s)	IAKM (Secondary: components interacting with IAKM)
Measurement Point(s)	IAKM
Goal (Success Criteria)	Ensure secure communication is in place
Validation Strategy	Review
Validation Plan	Review IAKM deployment and configuration to ensure TLS-grade security is applied across its communication channels with other components.
Validation Assessment	**Partial**. Interfaces defined on LL-MEC, but will not be implemented. Following the deprecation of the 4G+ LL-MEC architecture on OAI, the replaced 5G FlexCN (during the project) did not have the same level of maturity and the security function have not required in the integration framework.

Requirement	TR.55
Owner	EURECOM
Evaluation Scope	UC1, UC3
Components(s)	MEC (Secondary: any component interacting with MEC)
Measurement Point(s)	MEC
Goal (Success Criteria)	• A component is created at the MEC • It sends its credentials over HTTPS to the MEC • MEC validates the component credentials
Validation Strategy	Demonstration
Validation Plan	Edge controller creates a component at the MEC; The component is authenticated by the MEC
Validation Assessment	**Partial**. Interfaces defined on LL-MEC, but will not be implemented. Following the deprecation of the 4G+ LL-MEC architecture on OAI, the replaced 5G FlexCN (during the project) did not have the same level of maturity and the security function have not required in the integration framework.

Requirement	TR.56
Owner	EURECOM
Evaluation Scope	UC1, UC3
Components(s)	MEC (Secondary: any component interacting with MEC)
Measurement Point(s)	MEC
Goal (Success Criteria)	• A new component is created at the MEC • It subscribes to MEC services • The MEC enter the component into the subscription database
Validation Strategy	Demonstration
Validation Plan	Edge controller creates a component at the MEC; it subscribes to the MEC; subscription is rejected if authentication fails (TR.55), otherwise succeeds. MEC can monitor the list of subscribed entities in its subscription database.
Validation Assessment	**Partial**. Interfaces defined on LL-MEC but will not be implemented. Following the deprecation of the 4G+ LL-MEC architecture on OAI, the replaced 5G FlexCN (during the project) did not have the same level of maturity and the security function have not required in the integration framework.

Requirement	TR.57
Owner	SANL
Evaluation Scope	Event Captors
Components(s)	Event Captors (Secondary: Edge devices)
Measurement Point(s)	Event Captors, Edge Devices
Goal (Success Criteria)	Ensure Event Captors can capture telemetry from edge devices involved in UCs
Validation Strategy	Review

Validation Plan	Review design, implementation details and documentation to ensure that Event Captors can capture the telemetry provided by the edge devices present in the UC environments (superset of edge devices monitored in the context of the trust scenarios and associated key scenes).
Validation Assessment	**Pass**. Successful verification of Cycle 1 & Cycle 2 event captors, as demonstrated in UC1, UC2 & UC3 demos. Review of core design principles in D4.8.

Requirement	TR.58
Owner	SANL
Evaluation Scope	Event Captors
Components(s)	Event Captors (Secondary: Trust Broker, SAP)
Measurement Point(s)	Event Captors, Trust Broker, SAP Monitor component
Goal (Success Criteria)	Ensure Event Captor alerts are transmitted and ingested by SAP
Validation Strategy	Demonstration
Validation Plan	1. Trigger all Event Captors developed within each UC 2. Verify alert is relayed to Assurance Platform via Trust Broker
Validation Assessment	**Pass**. Demonstration across UCs.

Requirement	TR.59
Owner	UOULU
Evaluation Scope	Manufacturing AI
Components(s)	Manufacturing AI: Engrave area detection & Grab spot detection
Measurement Point(s)	Pre-training of AI at the plant edge
Goal (Success Criteria)	Obtain pre-trained models for engrave area detection and grab spot detection. Given a new workpiece, infer the engraving area and grabbing spot similar to the labels available in the training data.
Validation Strategy	Demonstration
Validation Plan	1. Offline training over the labelled dataset 2. Offline validation over a labelled testing dataset 3. Online validation over the demo setup
Validation Assessment	**Pass**: A single model is trained for both tasks using augmented data. The functionality of the model has been validated in offline data as well as after the integration in the demo setting. The results are shared in D3.6 and D5.5.

Requirement	TR.60
Owner	UOULU
Evaluation Scope	Manufacturing AI
Components(s)	Manufacturing AI: Engrave area detection & Grab spot detection
Measurement Point(s)	Pre-training of AI at the plant edge

Goal (Success Criteria)	Obtain dataset for pretraining the Manufacturing AI models. 1. Generate training and testing datasets 2. AI inference accuracy over testing data is close to the accuracy over training data
Validation Strategy	Demonstration
Validation Plan	Option 1: Manual data collection 1. Mark engraving area on workpieces and define grab spot by grabbing with the robot arm 2. Place the workpieces on the storage area/machines 3. Collect images and produce the dataset Option 2: Data augmentation 1. Collect few images of workpieces on storage area/machines 2. Define engraving area and grab spot using image processing tools 3. Augment new data (images) by rotating, scaling, and skewing the processed images
Validation Assessment	**Pass**: Inference accuracy is validated over augment data (Option 2 is successful) as well as in experiments carried out after the integration (Option 1 is successful). The results are shared in D3.6 and D5.5.

Requirement	TR.61
Owner	SANL
Evaluation Scope	SAP
Components(s)	SAP
Measurement Point(s)	SAP, Event Captors
Goal (Success Criteria)	Verify SAP aggregates, ingests, and stores raw evidence received from Event Captors
Validation Strategy	Demonstration
Validation Plan	Following execution of trust scenarios defined within each UC, cross-check to verify all raw evidence relayed by event captors have been received, ingested, and triggered appropriate monitoring rules' reasoning. The corresponding events should be listed within the relevant monitoring assessment page.
Validation Assessment	**Pass**. Successful verification for all Cycle 1 & Cycle 2 scenarios, across UCs.

Requirement	TR.62
Owner	SANL
Evaluation Scope	Trust Broker
Components(s)	Trust Broker
Measurement Point(s)	Trust Broker
Goal (Success Criteria)	Verify deployment and correct configuration and integration of Trust Broker
Validation Strategy	Review, Demonstration
Validation Plan	1. Review of design and deployment to verify presence of Trust Broker as a core integration point between Trust Enablers.

	2. Execution of UC demonstrators (Trust scenarios) to verify proper functioning of deployed Trust Broker, as a core integration point between Trust Enablers.
Validation Assessment	**Pass**. Design review (D4.8). Demonstration across UCs.

Requirement	TR.63
Owner	SANL
Evaluation Scope	Trust Broker
Components(s)	Trust Broker
Measurement Point(s)	Trust Broker
Goal (Success Criteria)	Verify correct configuration of demonstrator environments to allow functioning of Trust Broker.
Validation Strategy	Review
Validation Plan	Review of configuration / setup of all 3 UC environments, to verify port 5671 is not blocked.
Validation Assessment	**Pass**. Successful verification for Cycle 1 & Cycle 2 across UC environments.

Requirement	TR.64
Owner	Siemens, TTC
Evaluation Scope	UC1 (Scenario 3 – Trustworthiness)
Components(s)	Tractor Controller, 5G MEC, Edge Node
Measurement Point(s)	Tractor, Edge Node, 5G MEC
Goal (Success Criteria)	Installation and operation of trust enablers
Validation Strategy	Demonstration
Validation Plan	1. Bring the trust enablers containers up 2. If proper access rights are given (e.g., network traffic monitoring) all trust enablers operate as expected
Validation Assessment	**Pass:** The solutions for the trustworthiness have been installed on the tractor controller and the 5G MEC. Remote tests have been performed on the tractor controller (clients) and the 5G MEC (server) and they provide correct operation and functionality.

Requirement	TR.65
Owner	TUC/TSI
Evaluation Scope	UC1, Open Call 2 (DotSoft, WasteLocker)
Components(s)	MTD Client
Measurement Point(s)	Edge Nodes
Goal (Success Criteria)	An MTD Client is installed and working on every edge device
Validation Strategy	Demonstration
Validation Plan	Install the MTD Client on each related Use Case node. Deploy using docker image for both arm64 and amd64 architectures. Test end to end communication with MTD Server.

Validation Assessment	**Pass.** All Edge Nodes used on UC1 including the tractor controller and the Raspberry Pi boards have an instance of MTD Client installed and running. Each MTD Client is connected securely through the Trust Broker to the MTD Server. The MTD Clients have successfully completed all test scenarios related to UC1 (reported in Deliverable D5.5) proving their correct operation and functionality. Additionally, the MTD Client components have been installed successfully on edge nodes of the Open Call 2 partners and have proven their availability and correct deployment and functionality in their respected environments and test cases. Notice compared to Cycle 1: in previous iteration of this document (D5.3), it was reported that this requirement would also be validated/evaluated in the context of UC3. However, in UC3, the MTD Clients could affect the TSN network and therefore it was decided that they would not be deployed. Since, the MTD Clients have been successfully validated in both UC1 and the external environments of the Open Call partners, the validation of this requirement can be considered completed.

Requirement	TR.66
Owner	TUC/TSI
Evaluation Scope	UC1, Open Call 2 (DotSoft, WasteLocker)
Components(s)	MTD Client, MTD Server
Measurement Point(s)	MTD Server
Goal (Success Criteria)	Edge device information is sent to the MTD Server upon registration
Validation Strategy	Demonstration
Validation Plan	1. MTD Server connects to Trust Broker and waits for MTD Clients registration 2. MTD Client connects to Trust Broker 3. MTD Client sends registration message 4. MTD Server receives the registration message 5. All needed information is available 6. MTD Client receives network configuration and is successfully connected
Validation Assessment	**Pass.** MTD Server receives via secure channel all the basic information (Edge IP, MAC address, Public Key) of the MTD Client that wants to register in the system, which is verified by the MTD Server logs. This is validated on UC1 testbed and those of the Open Call 2 partners.

Requirement	TR.67
Owner	TUC/TSI, SANL
Evaluation Scope	UC1, UC3
Components(s)	MTD Server Event Captor IDS SAP
Measurement Point(s)	MTD Server
Goal (Success Criteria)	MTD Server receives a warning containing relevant information.
Validation Strategy	Demonstration

Validation Plan	For each use case: 1. A detection mechanism sends a warning 2. MTD Server receives the warning 3. The warning contains all relevant data
Validation Assessment	**Pass.** For UC1, the component responsible to trigger the MTD Server about a misbehaving node is TrustIDS. TrustIDS, upon detection, sends via the secure broker a JSON message to the MTD Server containing all needed information namely the node IP, the action to take (block), the TrustIDS public key and the reason of issuing a warning. This can easily be assessed by checking the TrustIDS and MTD Server logs where the sending and receiving of the warning is shown (reported in Deliverable D5.5). For UC3, the component that triggers the MTD Server is the Security Assurance Platform. It sends again a similar JSON message to the MTD Server via the secure broker, using a dedicated topic (mtd.alert). The difference between the two Use Cases lies on the mitigation action. This can easily be assessed by checking the MTD Server logs where the receiving of the warning is shown (reported in Deliverable D5.5).

Requirement	TR.68
Owner	TUC/TSI
Evaluation Scope	UC1, Open Call 2 (DotSoft, WasteLocker)
Components(s)	MTD Client MTD Server
Measurement Point(s)	MTD Server
Goal (Success Criteria)	MTD Client graceful shutdown
Validation Strategy	Demonstration
Validation Plan	1. Assuming an MTD Client registration has taken place as described in TR.65 2. MTD Client sends a de-registration message before shutdown 3. MTD Server receives the de-registration message 4. MTD Server removes MTD Client from its known clients list 5. MTD Server sends an updated configuration to the remaining clients
Validation Assessment	**Pass.** MTD Clients, upon shutdown, deregister from the MTD Server so that the network configuration can be updated and MTD Server has a correct current state. The MTD Server generates the new configuration and sends it to the remaining clients, excluding the MTD Client that requested deregistration from the clients list. This can be verified by the MTD Server/Client logs. This is considered completed as it has been validated on UC1 testbed and on those of the Open Call 2 partners.

Requirement	TR.69
Owner	TUC/TSI
Evaluation Scope	UC1, Open Call 2 (DotSoft, WasteLocker)
Components(s)	MTD Client MTD Server
Measurement Point(s)	MTD Server

Goal (Success Criteria)	Ensure that when an MTD Client does not answer in a Keep-Alive request in a timely manner, the MTD Server will consider it deregistered.
Validation Strategy	Demonstration
Validation Plan	1. Assuming an MTD Client registration has taken place as described in TR.65 2. MTD Server sends Keep-Alive requests to the registered MTD Client 3. MTD Client does not respond to Keep-Alive requests in a timely manner 4. MTD Server removes MTD Client from its known clients list, which is equivalent to a deregistration as described in TR.68 5. MTD Server sends an updated configuration to the remaining clients
Validation Assessment	**Pass.** All unresponsive MTD Clients, regardless of the underlying reason, are automatically deregistered from the MTD Server. The MTD Server removes the MTD Client from the clients list and sends an updated configuration to the remaining clients. This can be verified by the MTD Server/Client logs. This is considered completed as it has been validated on UC1 testbed and on those of the Open Call 2 partners.

Requirement	TR.70
Owner	TUC/TSI
Evaluation Scope	UC1, Open Call 2 (DotSoft, WasteLocker)
Components(s)	MTD Client MTD Server
Measurement Point(s)	MTD Client
Goal (Success Criteria)	Network configuration relevant information is included in the configuration message.
Validation Strategy	Demonstration
Validation Plan	1. MTD Server, either periodically or because of a triggering event, sends a new configuration to each registered MTD Client 2. MTD Client receives the configuration 3. The configuration includes all network data to keep it operational
Validation Assessment	**Pass.** The MTD Server sends the new configuration that includes all the relevant information (Server Public Key, Routing Table, Routes to Add, Routes to Remove) to each trusted MTD Client. This can be verified by the MTD Server/Client logs. This is considered completed as it has been validated on UC1 testbed and on those of the Open Call 2 partners.

Requirement	TR.71
Owner	TUC/TSI
Evaluation Scope	UC1
Components(s)	IDS
Measurement Point(s)	Edge nodes
Goal (Success Criteria)	IDS is successfully deployed on each edge node.
Validation Strategy	Demonstration
Validation Plan	Install Trust-IDS on each related Use Case node. Deploy using docker image for both arm64 and amd64 architectures. Test end to end communication with MTD Server.

Validation Assessment	**Pass.** All Edge Nodes used on UC1 including the tractor controller and the Raspberry Pi boards have an instance of Trust-IDS connected securely through the Trust Broker to the MTD Server. Notice compared to Cycle 1: in previous iteration of this document (D5.3), it was reported that this requirement would also be validated/evaluated in the context of UC3. There we demonstrated the integration with SAP being responsible for sending warnings about offending nodes, instead of the TrustIDS component.

Requirement	TR.72
Owner	TUC/TSI
Evaluation Scope	UC3
Components(s)	Interoperability Box
Measurement Point(s)	Edge node connected with limited-resource devices
Goal (Success Criteria)	The Interoperability Box has a proper description of the properties and available actions of each limited resource device connected.
Validation Strategy	Demonstration
Validation Plan	• An IoT application that wants to use information or services from the limited resource devices communicates with the Interoperability Box installed on an edge node • The Interoperability Box has a proper description of the limited resource device properties and available actions and is therefore capable to act as a middleman between the IoT application and the device. • All requested actions (or information exchange) between the IoT application and the limited resource device are correctly carried out.
Validation Assessment	**Pass.** Interoperability Box is installed on Use Case 3 testbed where it interfaces with MiroCard which offers communication only through BLE. MiroCard regularly broadcasts temperature readings which are received and stored on the Interoperability Box. An HTTP REST API is offered for all external services that need to poll those temperature readings. More details and REST API invocation is available on Deliverable D5.5.

Requirement	TR.75
Owner	TUC/TSI
Evaluation Scope	UC3
Components(s)	AAA
Measurement Point(s)	AAA
Goal (Success Criteria)	AAA provides access tokens to legitimate applications
Validation Strategy	Demonstration
Validation Plan	1. Legitimate applications have pre-shared client credentials from AAA 2. An application uses these credentials to access the AAA API 3. AAA responds with a valid JWT token
Validation Assessment	**Pass.** A reverse proxy (nginx/lua-resty-openidc) is used between external services that issue requests towards Edge applications. External services are registered as trusted clients to the Keycloak and can acquire JWT tokens using a pre-shared client id/secret.

	The same applies for the reverse proxy, which can subsequently validate the external services JWT token with the Keycloak server.

Requirement	TR.76
Owner	TSI
Evaluation Scope	UC3
Components(s)	AAA Reverse Proxy
Measurement Point(s)	Reverse Proxy
Goal (Success Criteria)	Reverse Proxy validates incoming requests before routing to the internal services
Validation Strategy	Demonstration
Validation Plan	1. An application sends a request with a valid token to an internal service 2. Reverse Proxy acts as a mediator and validates the token through AAA 3. Reverse Proxy, depending on the configured access rights, redirects or blocks the request
Validation Assessment	**Pass.** Reverse Proxy is installed on Edge Nodes of UC3. Users initiating a request at an internal service are redirected to Keycloak login page before the request can be fulfilled. Also, services that are configured as trusted clients on Keycloak, can retrieve their access token and make requests to internal services. This is validated on UC3 testbed and considered completed.

Requirement	TR.77
Owner	UOULU
Evaluation Scope	UC1
Components(s)	AI for Obstacle Bypassing
Measurement Point(s)	AI for Obstacle Bypassing
Goal (Success Criteria)	Obtain pretrained model for obstacle bypassing
Validation Strategy	Demonstration
Validation Plan	Given the labelled dataset for training: 1. Offline training over the labelled dataset 2. Offline validation over a labelled testing dataset 3. Online validation over the demo setup
Validation Assessment	**Pass.** A lab setting is developed with an off-the-shelf robot to collect data and generate an initial AI model to detect and bypass the obstacle. The pretrained model performance is validated and the demonstration is included in the final demo video.

Requirement	TR.78
Owner	AVL, TTC, HOLO, UOULU
Evaluation Scope	UC1
Components(s)	AI for Obstacle Bypassing

Measurement Point(s)	AI for Obstacle Bypassing
Goal (Success Criteria)	Labelled data available, using the linear and angular velocity provided by the tractor associated with each frame from the camera.
Validation Strategy	Demonstration
Validation Plan	1. Frames coming from the cameras of the tractor 2. Assign linear and angular velocity information to each individual frame. All this data coming from the tractor 3. Validation of the data by the trained AI model if the tractor can pass by the obstacle, based on the available, labelled data
Validation Assessment	**Pass with some alterations**: Successfully validated in a sandbox scenario by replacing the eTractor from an off-the-shelf mobile robot (Waveshare Jetank). The scenario mimics the real-world setting by designing a miniature environment. Demonstration shows that the mobile robot can successfully bypass an obstacle placed on its path. The AI model is integrated with the tractor controller and validated with its ability to generate control commands.

Requirement	TR.79
Owner	Siemens
Evaluation Scope	UC3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	Edge apps must be orchestrated with docker-compose as one or multiple docker containers.
Validation Strategy	This is a precondition for edge apps to be deployed; all edge apps comply to this
Validation Plan	*It is implicitly validated by the presence of running apps, other orchestration formats cannot be configured*
Validation Assessment	**Pass.** As it is a precondition for edge apps to be deployed; all edge apps comply to this

Requirement	TR.80
Owner	Siemens
Evaluation Scope	UC3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	Edge apps requiring to exposing ports to outside on the host must be able to set the port number within the port range [32768, 60999]
Validation Strategy	This is a precondition for edge apps to be deployed; all edge apps comply to this
Validation Plan	Edge apps cannot be deployed, which offend this rule. The deployment tool provides an error message.

Validation Assessment	**Pass**. As it is a precondition for edge apps to be deployed; all edge apps comply to this.

Requirement	TR.81
Owner	Siemens
Evaluation Scope	Use Case 3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	Edge App container port numbers must be at least 8000
Validation Strategy	This is a precondition for edge apps to be deployed; all edge apps comply to this
Validation Plan	Edge apps cannot be deployed, which offend this rule. The deployment tool provides an error message.
Validation Assessment	**Pass**. As it is a precondition for edge apps to be deployed; all edge apps comply to this.

Requirement	TR.82
Owner	Siemens
Evaluation Scope	Use Case 3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	Edge apps must be x86 compatible and Linux-based
Validation Strategy	This is a precondition for edge apps to be deployed; all edge apps comply to this
Validation Plan	Edge apps cannot be deployed, which offend this rule. The deployment tool provides an error message.
Validation Assessment	**Pass**. As it is a precondition for edge apps to be deployed; all edge apps comply to this.

Requirement	TR.83
Owner	Siemens
Evaluation Scope	Use Case 3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	Edge apps must be ready to be stopped and started at any time.
Validation Strategy	Trials and analysis of edge app architecture
Validation Plan	Sudden restart of edge apps (e.g. with power cycles) at random point in time

Validation Assessment	**Pass**. No app is sensitive to unforeseen restarts

Requirement	TR.84
Owner	Siemens
Evaluation Scope	UC3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	Edge apps must be able to self-reliantly recreate their state after start (local state might be destroyed after each shutdown)
Validation Strategy	Trials and analysis of edge app architecture
Validation Plan	Sudden restart of edge apps (e.g. with power cycles) at random point in time
Validation Assessment	**Pass**. No app is sensitive to unforeseen restarts

Requirement	TR.85
Owner	Siemens
Evaluation Scope	UC3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	Edge app developer must comply to the guidelines for the Hello Edge App (https://gitlab.eurecom.fr/intelliot-project/edge/apps/spec-iedge/hello-edge-app)
Validation Strategy	This is a precondition for edge apps to be deployed; all edge apps comply to this
Validation Plan	Edge apps cannot be deployed, which offend this rule. The deployment tool provides an error message.
Validation Assessment	**Pass**. As it is a precondition for edge apps to be deployed; all edge apps comply to this.

Requirement	TR.86
Owner	Siemens
Evaluation Scope	Use Case 3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	The configuration files for provisioning of Edge Apps have to be available in the edge app's corresponding GitLab repository.
Validation Strategy	Manual check
Validation Plan	Check if specification is there and if the edge app can be provisioned

Validation Assessment	**Pass**. All corresponding gitlab repos are there.

Requirement	TR.87
Owner	Siemens
Evaluation Scope	Use Case 3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	Edge Apps must not be depending on mounted volumes on the host file systems
Validation Strategy	This is a precondition for edge app GNFR.15 to be deployed; all edge apps comply to this
Validation Plan	Edge apps cannot be deployed, which offend this rule. The deployment tool provides an error message.
Validation Assessment	**Pass**. As it is a precondition for edge apps to be deployed; all edge apps comply to this.

Requirement	TR.88
Owner	Siemens
Evaluation Scope	Use Case 3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	Edge Apps must not depend on special hardware equipment on the edge device (dongles, GPUs)
Validation Strategy	This is a precondition for edge apps to be deployed; all edge apps comply to this
Validation Plan	Edge apps cannot be deployed, which offend this rule. The deployment tool provides an error message.
Validation Assessment	**Pass**. As it is a precondition for edge apps to be deployed; all edge apps comply to this.

Requirement	TR.89
Owner	Siemens
Evaluation Scope	Use Case 3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	Edge apps must not expose ports on number 80, 443, 1883
Validation Strategy	This is a precondition for edge apps to be deployed; all edge apps comply to this
Validation Plan	Edge apps cannot be deployed, which offend this rule. The deployment tool provides an error message.

Validation Assessment	**Pass**. As it is a precondition for edge apps to be deployed; all edge apps comply to this.

Requirement	TR.90
Owner	Siemens
Evaluation Scope	Use Case 3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	Edge apps must not rely on Internet uplink
Validation Strategy	Demonstration
Validation Plan	Interrupt the internet access, e.g. by removing cable
Validation Assessment	**Pass**. The apps still work on network isolation.

Requirement	TR.91
Owner	Siemens
Evaluation Scope	Use Case 3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	Edge app must not rely on the environment where they are executed
Validation Strategy	Demonstration
Validation Plan	Execute app in multiple environments, e.g. on a linux pc and/or on an edge device
Validation Assessment	**Pass**. As in the scope of the UC3 demonstrator the apps ran successfully in all relevant environments.

Requirement	TR.92
Owner	Siemens
Evaluation Scope	Use Case 3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	Edge apps are head-less (no GUI)
Validation Strategy	Demonstration
Validation Plan	Validation on edge app inspection
Validation Assessment	**Pass**. All edge apps won't depend on GUI operation except it is explicitly required

Requirement	TR.93
Owner	Siemens
Evaluation Scope	Use Case 3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	Images for edge app container must be available in an accessible container registry
Validation Strategy	This is a precondition for edge apps to be deployed; all edge apps comply to this
Validation Plan	*Evaluate all app specifications used in UC3 automatically*
Validation Assessment	**Pass**. All specifications comply to this requirement as the container registry could be accessed with the build tool.

Requirement	TR.95
Owner	Siemens
Evaluation Scope	Use Case 3
Components(s)	Engraver App
Measurement Point(s)	Engraver App
Goal (Success Criteria)	Text engraving requests must include the text to be engraved (single line, no formatting), the font size (or alternatively the text width and/or the text height), the font name, the northwestern corner of the text
Validation Strategy	Demonstration
Validation Plan	Execute text engraving job and see if it is engraved correctly
Validation Assessment	**Pass.** We have demonstrated that the engraver app works as specified.

Requirement	TR.96
Owner	Siemens
Evaluation Scope	Use Case 3
Components(s)	Engraver App
Measurement Point(s)	Engraver App
Goal (Success Criteria)	Text engraving requests should include a maximal bounding box with x, y, width and height coordinates
Validation Strategy	Demonstration
Validation Plan	Execute text engraving job and see if it is engraved correctly

Validation Assessment	**Pass.** We have demonstrated that the engraver app works as specified.

Requirement	TR.97
Owner	Siemens
Evaluation Scope	Use Case 3
Components(s)	All edge apps
Measurement Point(s)	All edge apps
Goal (Success Criteria)	An application based on the service of an edge app must cope that an edge app can be instantiated only once in the edge infrastructure
Validation Strategy	This is a precondition for edge apps to be deployed; all edge apps comply to this
Validation Plan	Doesn't need to be validated
Validation Assessment	**Pass.** Implicitly validated as we have shown that the whole use case works without having any edge app instantiated multiple times.

3.4 Validation and Evaluation of Open Call 1 Technical Requirements

During the first Open Call, the IntellIoT consortium worked with four SMEs which integrated their tools and technologies each one on a different domain. They defined their own set of requirements and KPIs and provided reports on their progress. Their results can be found on their respective final reports.

iKnowHow provided their autonomous self-navigating all-terrain ground vehicle, GreenBot. GreenBot is comprised of a moving platform and a robotic arm. They extended the Agriculture Use Case functionality by contributing to all three IntellIoT Pillars. They interfaced with HyperMAS, used a VR application to control GreenBot and connected their Farmer's logbook to DLTs.

More specifically, iKnowHow contributed to UC1 by demonstrating the scalability of the solutions developed for the agricultural domain. Their robotic solution was integrated into the HyperMAS, demonstrating that multiple systems can easily be the agents in a similar way. Additionally, the VR application was not only used to control the actual robot in a similar way as the eTractor from the use case, but, as an additional feature, they were able to control the actuator that was mounted on their robot. This was a robotic arm, with a sprayer system and camera mounted on the end effector of the arm. With the aid of the VR application, they demonstrated the capability of controlling the robot arm remotely using the VR application, using the data stream from the camera as input for the VR system. Such a system can provide more functionality to the remote operator by being able to perform tasks in the field (e.g., pruning, harvesting, cutting, etc.) without being present in the field.

Furthermore, iKnowHow introduced their Farmer's logbook to the overall system, using this as a way to log the data coming from the tractor, exploiting the DLT solution from IntellIoT for secure storage of the data. Different forms of data can be stored, varying from actuation data to machine data, to make sure that the machine is doing what it is supposed to do up to predictive maintenance to ensure the lifetime of the machine. During the cooperation with iKnowHow in the use case integration, it was discussed how the logbook could be more exploited within the use case. The logbook provided an additional feature, namely that it provided a map, where the tractor was going to be deployed. This feature enabled the use case team to use the logbook as a simplified Farm Management System (FMS), which enables the farmer to define the task the tractor has to perform and where this task has to be performed. In cooperation with the use case partners, it was decided to extend this feature. The farmer selects the area where the tractor must perform its task, and the logbook uses a trajectory calculation algorithm to calculate a coverage algorithm over the whole selected field. This would then be recalculated in specific way points and function as input for the trajectory of the tractor. Through this functionality, the logbook was integrated in the overall solution and becoming more as a single logbook, but a tool for the farmer to easily program the tractor and have it traverse the selected field.

iKnowHow has fulfilled their part according to what they promised into the Open Call application, by providing the robot platform, and integrating it with the HyperMAS, controlling the platform and the actuator with the VR application and merging the logbook with the DLT functionality. Additionally, they even went beyond their application, but making the logbook an FMS, with which the farmer can define the activity of the agricultural vehicle and have it perform the task assigned to it.

Vidavo enhanced the Healthcare Use Case by providing a number of components. At the patient side, they provide the Vidavo SDK and Vida24 mobile app which allows the gathering of patient data from smart devices. IntellIoT AI models, running locally on the mobile device, process that information and sends alarms to the corresponding physician. At the other end, Vida24 web application provides an interface to the physician who can assess the received data and issue reminders towards patients.

Careplans are integral to the effective monitoring and management of patient health and an important component of UC2. Medical experts develop and assign careplans for each patient, outlining the prescribed level of physical activity per week, which includes detailed instructions on the number of repetitions, type, and the duration of each exercise session. Furthermore, these careplans define thresholds for vital signs both at rest and during physical exertion. Vidavo has enhanced this personalized approach by designing and implementing an advanced notification algorithm. This notification algorithm, engineered in collaboration with medical specialists from PAGNI, processes vital sign data alongside the directives set forth in careplans and identifying and signaling medical experts about critical events thereby ensuring targeted medical interventions.

Trilogis extended the IntellIoT UC3 towards the real-time representation of the digital twin of the manufacturing process. The indoor localization and tracking of workpieces have been integrated combining both hardware and software components. In more detail, the digital model of the UC3 manufacturing cell has been integrated on a real-time map for continuous geolocation of workpieces. A set of sub-areas delimited by the integrated geofence engine of the software TRACCIA® by Trilogis has been deployed to collect the in/out events of each workpiece handled by the robot. The RFID technology has been selected among others to easily identify proximity areas using low-cost and scalable RFID transponders attached on the workpieces. Workpieces are localized when they are moved within the coverage of an RFID reader, that has been configured to properly calibrate the size of the detection area.

The integration of the information available in HyperMAS that is useful for the computation of KPIs has been investigated. In particular, the information about the status of the machines (e.g., busy and percentage of completion) has been exploited for the analysis of the working phases. The KPIs based on spatiotemporal data of workpieces have been computed. Dedicated REST API have been developed to expose toward HyperMAS in which proximity area are located the workpieces. The historical analysis of the KPI values and trend is available using the interactive dashboard on the web application TRACCIA®. This analysis is useful to evaluate the performance of the process and identify strengths and weaknesses of the process for continuous improvement.

Finally, selected MYWAI technologies have been integrated into the design of the IntellIoT architectural framework and mapped to each of the three use case storylines. With specific reference to UC3-Manufacturing, MYWAI has showcased open interoperability of production technical data within the IntellIIot framework by enabling the flexible integration of equipment, sensor systems and data intelligence modules to offer new services and business models in the emerging Equipment-as-a-Service market. MYWAI has defined the project's context and identified the MYWAI Enablers, conducting a thorough analysis of functional, non-functional, and technical requirements. MYWAI has mapped the MYWAI Enablers into the IntellIoT Framework. Additionally, there has been a particular emphasis on UC3-Manufacturing, MYWAI Enablers have been adapted and tailored for the UC3 demonstrator, followed by their integration and rigorous testing as integral components of the UC3 demonstrator.

In UC3, the MYWAI Enablers allow automatic detection of anomalies, malfunctioning and failures potentially affecting mechanical/electrical components of the robot arm and/or other manufacturing equipment through AI-based algorithms running on an edge/very edge device and applied to data time series gathered by heterogeneous sensors placed on the monitored equipment. Anomaly/failure models have been designed by domain experts according to available real time and/or historical data collected by the sensors. Detected anomalies/failures are promptly notified to human operators to trigger their intervention, either from remote or in the industrial plant. To ensure transparency and knowledge transfer, MYWAI has diligently documented the work performed along with the associated achievements.

4. VALIDATION AND EVALUATION OF AGRICULTURE USE-CASE REQUIREMENTS

This section provides the validation and evaluation report for the requirements that are associated with IntellIoT's Agriculture use-case (UC1) and covers both functional and non-functional requirements as specified in Deliverable D2.5.

4.1 Validation and Evaluation of Functional Requirements

Requirement	FR.UC1#1
Owner	HSG
Evaluation Scope	UC1.1.1
Components(s)	Logbook
Measurement Point(s)	Goal Creation of the Logbook
Goal (Success Criteria)	The user can select the fields to be processed, the process to be executed, sends the goal to the HyperMAS by clicking a button
Validation Strategy	Demonstration
Validation Plan	1. Selection of the fields to be processed 2. Selection of the process to be executed 3. Send the goal to the HyperMAS
Validation Assessment	**Pass**. This process has been shown in the technical video for UC1, where the Logbook is used to create a generate a goal as a path for the tractor that is sent to the agent.

Requirement	FR.UC1#2
Owner	HSG
Evaluation Scope	UC1.1.1
Components(s)	HyperMAS
Measurement Point(s)	Agents in HyperMAS
Goal (Success Criteria)	The agents select the vehicle to be used and provides waypoints towards fulfilling the goals.
Validation Strategy	Demonstration
Validation Plan	1. The agent selects the vehicle 2. The agents send waypoints
Validation Assessment	**Pass.** This has been shown during integration meeting in Steyr.

Requirement	FR.UC1#3
Owner	AVL
Evaluation Scope	UC1.1.2
Components(s)	eTractorControlSystem

Measurement Point(s)	Movement of the eTractor, which can be determined via the current GPS coordinates (provided as a topic), or via the GPS coordinates over time.
Goal (Success Criteria)	Reaching a waypoint goal and simple collision avoidance during the action. If the destination cannot be reached due to an obstacle, the ETCS (eTractor Control System) goes into a defined error state and waits for new instructions.
Validation Strategy	Demonstration
Validation Plan	Passing of an incorrect waypoint -> ETCS.state = error (eTractor stops/not moving) Passing a correct waypoint -> ETCS.state = moving and eTractor starts to move Time to arrive waypoint, assurance that no objects have been bypassed
Validation Assessment	**Pass**. An RTK system for better position resolution was implemented. First a low-level simulation of GPS and radar sensor was done, then the system was deployed on a mini demo model truck for evaluation of the algorithm. When appearing an obstacle, the velocity was reduced in relation to the obstacle's distance and finally the model truck stopped. The AVL eTractor currently needs to be reworked. When this is done the evaluation of the collision avoidance shall be repeated.

Requirement	FR.UC1#5
Owner	AVL
Evaluation Scope	UC1.1.1
Components(s)	eTractorControlSystem
Measurement Point(s)	Sensor data are clearly visualized in the respective formats.
Goal (Success Criteria)	Scan environment, Maximum FOV (field of view) with minimum blind areas
Validation Strategy	Demonstration with various objects and rosbag (rosbag - ROS Wiki)
Validation Plan	Positioning of various objects (e.g., boxes) in a defined test arrangement (angle and distance, discretely divided)
Validation Assessment	**Pass**. Cameras installed, rosbags recorded; Additional radar sensor installed and sensor fusion done: various objects in different positions could be recognized.

Requirement	FR.UC1#7
Owner	AVL
Evaluation Scope	UC1.2.1
Components(s)	eTractorControlSystem
Measurement Point(s)	Parking brake of the eTractor as an essential feature of the safe position of the eTractor
Goal (Success Criteria)	Via the interface between TTTech controller and the eTractorControlSystem, the eTractor can be set to a safe state.
Validation Strategy	Demonstration and rosbag
Validation Plan	Sending error and E-stop commands to the ETCS Validate duration from sending the error command to reaching the safe state
Validation Assessment	**Pass**. An E-Stop command from the TTTech tractor controller to eTractorControlSystem is forwarded to the Vehicle Control Unit and causes the tractor to stop.

Requirement	FR.UC1#8
Owner	UOULU
Evaluation Scope	UC1.2.1
Components(s)	AI for Obstacle Bypassing
Measurement Point(s)	Communication with other devices and/or human operator
Goal (Success Criteria)	Obtain support for obstacle bypassing when the AI on the tractor fails
Validation Strategy	Demonstration
Validation Plan	Send requests for updated AI models and/or human takeover
Validation Assessment	**Pass**: AI as a service is integrated into the tractor controller. It is capable of informing the need of human assistance when the obstacle detection confidence is below a predefined threshold.

Requirement	FR.UC1#10
Owner	UOULU
Evaluation Scope	UC1.2.1
Components(s)	AI for Obstacle Bypassing
Measurement Point(s)	Receiving updated AI model from the network to the AI on the eTractor
Goal (Success Criteria)	Maintaining up-to-date AI models at all eTractors
Validation Strategy	Demonstration
Validation Plan	When AI fails to make confident decisions, it contacts to other AI components through IAKM and receives an updated AI model
Validation Assessment	**Pass with alterations**. Two different AI models are trained for well-lit and poorly lit situations. The integrated AI service on the tractor controller communicates with IAKM agent to obtain the relevant model when needed.

Requirement	FR.UC1#11
Owner	HOLO
Evaluation Scope	UC1.2.5
Components(s)	VR Glasses (Oculus Quest 2), HIL-VR-Application
Measurement Point(s)	VR Glasses
Goal (Success Criteria)	VR Glasses displays a 180° view of camera stream to the Operator
Validation Strategy	Demonstration
Validation Plan	Operator is able to see a 180° video display in VR space
Validation Assessment	**Pass**. User can see 180° video display in VR space.

Requirement	FR.UC1#12

Owner	HOLO
Evaluation Scope	UC1.2.5
Components(s)	HMI, VR Glasses, HIL-VR-Application, Tractor, Interface
Measurement Point(s)	Tractor, HIL-VR-Application, Interface
Goal (Success Criteria)	Operator can remotely control the tractor using the HMI
Validation Strategy	Demonstration
Validation Plan	1. Operator sends movement commands using HMI 2. HIL-VR-Application receives commands 3. HIL-VR-Application sends commands to the Tractor using the interface 4. Robot receives and performs the commands
Validation Assessment	**Pass**. During the integration, it was confirmed that commands can be passed to the tractor, which will control its movements. Tested directions: forward, backwards, left, and right, as well as everything in between. Level of acceleration was controlled by the joystick interaction. Pushed fully back, the tractor drove in full speed.

Requirement	FR.UC1#14
Owner	HOLO
Evaluation Scope	UC1.2.5
Components(s)	VR Glasses
Measurement Point(s)	VR Glasses
Goal (Success Criteria)	Display surroundings of tractor in VR
Validation Strategy	Demonstration
Validation Plan	Tractor environment is overlayed in VR for the Operator
Validation Assessment	**Partial.** The video feed displays the environment around the tractor. Sensor data, such as point cloud was not displayed, as it was not useful and caused too much stress on the network.

Requirement	FR.UC1#18
Owner	HOLO
Evaluation Scope	UC1.2.6, UC1.2.7
Components(s)	Entity, HIL-VR-Application, VR Glasses, Interface
Measurement Point(s)	VR Glasses
Goal (Success Criteria)	Display data or video stream from connected entities
Validation Strategy	Demonstration
Validation Plan	1. HIL-VR-Application connects to selected entities through an interface 2. Entities stream data or video through established interface 3. VR Glasses displays the information to Operator through a user interface
Validation Assessment	**Pass**. Operator sees video content in VR headset.

4.2 Validation and Evaluation of Non-Functional Requirements

Requirement	UCNFR.2
Owner	HOLO
Evaluation Scope	UC1
Components(s)	HMI, HIL-VR-Application
Measurement Point(s)	HIL-VR-Application
Goal (Success Criteria)	HIL-VR-Application provides Operator with all the driving capabilities (e.g., steering, accelerating, and breaking) through the HMI (i.e., VR Controllers)
Validation Strategy	Demonstration
Validation Plan	Operator can perform all driving functionality remotely
Validation Assessment	**Pass**. The HIL application supports full control of the tractor once the help request was accepted by the Human Operator. The Human Operator can steer the tractor in all directions, they can control acceleration through the controllers and can observe the tractor's environment via the VR headset. This functionality was tested during the integration meeting in Steyr.

Requirement	UCNFR.3
Owner	TTC, EURECOM, AVL
Evaluation Scope	UC1: Key Scene 2.2
Components(s)	Tractor, 5G Infrastructure
Measurement Point(s)	Tractor
Goal (Success Criteria)	Tractor must stop as fast as possible if requested by the human operator
Validation Strategy	Demo
Validation Plan	1. The human will control the tractor 2. Based on the speed of the tractor, it will be defined how fast the tractor needs to stop, i.e., it is only allowed to move a maximum distance. 3. The distance will be measured and when needed updated.
Validation Assessment	**Pass with alterations:** The tractor is successfully being controlled by the human operator and stops as soon as the operator stops giving commands. The requirement has been validated in a real-world demonstration scenario using the back-up solution of available WiFi, as the 5G infrastructure was not up and running at the final tests on the field.

5. VALIDATION AND EVALUATION OF HEALTHCARE USE-CASE REQUIREMENTS

This section provides the validation and evaluation report for the requirements that are associated with IntellIoT's Healthcare use-case (UC2) and covers both functional and non-functional requirements as specified in Deliverable D2.5.

5.1 Validation and Evaluation of Functional Requirements

Requirement	FR.UC2#1
Owner	Philips
Evaluation Scope	UC2.2.1, UC2.2.2, UC2.2.3
Components(s)	Local AI
Measurement Point(s)	Local AI, Healthcare AI Models, Patients Data Repository, Human Machine Interface
Goal (Success Criteria)	The Local AI collects measurements of steps, heart rate, oxygen saturation, blood pressure, temperature and body weight. The Healthcare AI Models can use the measurements for training and inference. A subset of the measurements is stored in the Patients Data Repository. The Human Machine Interface can display the measurements.
Validation Strategy	Demonstration
Validation Plan	1. A patient takes a new measurement. 2. For Collaborative IoT, when triggered or scheduled by the Local AI, the Healthcare AI Models use the new measurement. 3. For Human-in-the-Loop, the new measurement is stored in the Patients Data Repository and is displayed by the Human Machine Interface.
Validation Assessment	**Pass**. In Cycle 1, the Local AI and Healthcare AI Models use new measurements taken by the smartwatch in the lab setup. In Cycle 2, the rest of the smart devices measurements from the pilot study are used. Also, storing data in the Patients Data Repository and displaying it in the Human Machine Interface is already implemented.

Requirement	FR.UC2#2
Owner	Philips
Evaluation Scope	UC2.2.1, UC2.2.2, UC2.2.3
Components(s)	Local AI
Measurement Point(s)	Local AI, Patients Data Repository, Human Machine Interface
Goal (Success Criteria)	The Local AI collects measurements while the patient is both resting and performing physical exercise. A subset of the data is stored in the Patients Data Repository and displayed in the Human Machine Interface.
Validation Strategy	Demonstration
Validation Plan	1. A patient takes a new measurement while resting. Also, a patient takes a new measurement while exercising. 2. A subset of the measurements is transferred to and stored in the Patients Data Repository.

	3. The measurements are processed and displayed by the Human Machine Interface, where the physician can access them.
Validation Assessment	**Pass**. In Cycle 1, the Local AI uses new measurements taken by the smartwatch in the lab setup. In Cycle 2, the rest of the smart devices are used in the pilot by patients. Also, storing data in the Patients Data Repository and displaying it in the Human Machine Interface is already implemented.

Requirement	FR.UC2#3
Owner	Philips
Evaluation Scope	UC2.2.1, UC2.2.2, UC2.2.3
Components(s)	Local AI
Measurement Point(s)	Local AI, Patients Data Repository, Human Machine Interface
Goal (Success Criteria)	The times that a patient initiates and terminates an exercise session are registered by the Local AI. Also, the times and corresponding measurements are stored in the Patients Data Repository. The Human Machine Interface displays the times and corresponding measurements.
Validation Strategy	Demonstration
Validation Plan	1. A patient initiates an exercice session. 2. The time of initiation is registered by the Local AI. 3. The corresponding measurements are stored in the Patients Data Repository and can be displayed by the Human Machine Interface. 4. At the end of the exercise session, the patient terminates it. 5. The time of termination is registered by the Local AI.
Validation Assessment	**Pass**. Start time and end time of exercise can be viewed in HMI.

Requirement	FR.UC2#4
Owner	Philips
Evaluation Scope	UC2.2.1, UC2.2.2, UC2.2.3
Components(s)	Smartwatch, Patients Data Repository, Human Machine Interface
Measurement Point(s)	Patients Data Repository, Human Machine Interface
Goal (Success Criteria)	The patient can trace their electrocardiogram (ECG) on demand. The ECG is subsequently stored in the Patients Data Repository and displayed by the Human Machine Interface.
Validation Strategy	Demonstration
Validation Plan	1. A patient traces their ECG, for example when they experience symptoms. 2. The ECG is stored in the Patients Data Repository and displayed by the Human Machine Interface.
Validation Assessment	**Partial**. The Human Machine Interface has a functionality for uploading smartwatch-derived ECG files (PDF) to the mobile app and web platform, which can then be stored in the patient data repository. Throughout the duration of the 2nd phase of the pilot study conducted in the context of UC2 (device-facilitated follow-up phase), enrolled patients did indeed obtain ECG tracings using the smartwatch, primarily when experiencing palpitations with or without dizziness. However, for the scopes of this pilot, it was decided

	that smartwatch-derived ECG files should not be transmitted to the physicians by direct upload into the mobile app, but rather via e-mail. Then, the physician has the option to redact personal data from the file and subsequently upload it into the patient-physician interface and patient data repository. The rationale for this decision lies primarily in the fact that ECG files generated by the manufacturer's health app also display patients' names and/or other personal data such as birth dates, which few patients would have been able to redact before uploading. Thus, upload of such files would constitute a deviation of de-identification requirements set for this study (D1.6), according to which any patient-related data entering the mobile app should be already pseudonymized. The only theoretical way in which this could have been achieved would be direct access to- and transmission of- raw ECG data (rather than PDF flies), a functionality that could not be developed within the project period.

Requirement	FR.UC2#5
Owner	Philips
Evaluation Scope	UC2.2.1, UC2.2.2, UC2.2.3
Components(s)	Smartwatch, Local AI, Patients Data Repository, Human Machine Interface
Measurement Point(s)	Local AI, Patients Data Repository, Human Machine Interface
Goal (Success Criteria)	The Local AI can register distance covered in meters, steps, and calories consumed. This data is stored in the Patients Data Repository and displayed by the Human Machine Interface.
Validation Strategy	Demonstration
Validation Plan	1. A patient starts an exercise session. 2. Measurements of distance covered in meters, steps and calories consumed are registered by the Local AI. 3. The measurements are stored in the Patients Data Repository and displayed by the Human Machine Interface.
Validation Assessment	**Pass**. A functionality has been added to the Local AI with the potential to register distance covered in meters, step count and calories consumed, all of which are readily provided by the smartwatch and are first registered into the manufacturer's mobile app. However, for the purposes of the UC2 pilot we opted to only utilize step count as a measure of physical activity, as a potentially more representative index of patient effort and has been consistently used in clinical trials evaluating dose-response relationships between physical activity and health outcomes. Actual distance covered (in meters) would then depend also on each patient's somatometric characteristics. Calories consumed, on the other hand, is an index that was deemed highly unlikely to have any effect on our cohort's patient evaluation and management. Thus, to improve "signal-to-noise" ratio" for the physician evaluating transmitted data by navigating through the web platform, this parameter – although directly visible to the patient on smartwatch screen- was omitted. The Patients Data Repository has been implemented and functionality for displaying the measurements has been added to the Human Machine Interface.

Requirement	FR.UC2#6
Owner	Philips
Evaluation Scope	UC2.2.1, UC2.2.2, UC2.2.3

Components(s)	Smart weight scale, Local AI, Patients Data Repository, Human Machine Interface
Measurement Point(s)	Local AI, Patients Data Repository, Human Machine Interface
Goal (Success Criteria)	The Local AI can register body composition, such as percentage of body fat and body water. This data is stored in the Patients Data Repository and displayed by the Human Machine Interface.
Validation Strategy	Demonstration
Validation Plan	1. A patient takes measurements of body composition with the weight scale. 2. The measurements are registered by the Local AI. 3. The measurements are stored in the Patients Data Repository and displayed by the Human Machine Interface.
Validation Assessment	**Pass.** A functionality has been added to the Local AI for registering weight scale-derived measurements, i.e., body weight and details on body composition. Even though weight scales used in UC2 readily provide body composition details –such as body water content-it was eventually decided to only focus on body weight, which is the sole weight scale-derived parameter to be registered by the local AI. The rationale of this decision lies in the fact that measurements of body composition are conducted via bioelectrical impedance analysis, which is not indicated in the sub-population of patients enrolled in UC2 that carry implantable electronic devices (pacemakers / defibrillators). In addition, daily body weight (and not body water content) monitoring is the most consistently used method in patients with heart failure for early detection of fluid retention on a home-based setting, as well as during hospitalizations. The Patients Data Repository has been implemented and functionality for displaying the measurements has been added to the Human Machine Interface.

Requirement	FR.UC2#7
Owner	Philips
Evaluation Scope	UC2.2.1, UC2.2.4
Components(s)	Smart device, Local AI, Patients Data Repository, Human Machine Interface
Measurement Point(s)	Local AI, Patients Data Repository, Human Machine Interface
Goal (Success Criteria)	The Local AI can register biometric data. This data is stored in the Patients Data Repository and displayed by the Human Machine Interface.
Validation Strategy	Demonstration
Validation Plan	1. A patient takes some sort of biometric measurement. 2. The measurements are registered by the Local AI. 3. The measurements are stored in the Patients Data Repository and displayed by the Human Machine Interface.
Validation Assessment	**Pass**. This functionality has been added to the Local AI for registering biometric measurements, like blood pressure, blood oxygen saturation, heart rate, steps and body weight. The Patients Data Repository has been implemented and these measurements are presented to the Human Machine Interface.

Requirement	FR.UC2#8
Owner	Philips

Evaluation Scope	UC2.2.5, UC2.2.6
Components(s)	Local AI
Measurement Point(s)	Local AI, Healthcare AI Models, Patients Data Repository, Human Machine Interface
Goal (Success Criteria)	When abnormal measurements or significant changes on biometric parameters occur, these will be detected either by the Local AI or by the Healthcare AI Models. The Local AI will use rule-based algorithms, while the Healthcare AI Models will use neural networks, for the detections. The triggered alarms will be stored in the Patients Data Repository and displayed by the Human Machine Interface for analysis by the physicians.
Validation Strategy	Demonstration
Validation Plan	1. A patient takes measurements at prescribed times using different types of smart devices. 2. The measurements are analysed by the Local AI and the Healthcare AI Models. 3. In case of abnormal measurements or significant changes, for example when compared to past measurements or to similar patients, an appropriate alarm will be triggered. 4. The alarms will be stored in the Patients Data Repository and displayed by the Human Machine Interface for analysis by the physicians.
Validation Assessment	**Pass**. An initial Healthcare AI Model has been implemented, which provides its output to the Local AI. The Local AI uses rule-based algorithms, while the Healthcare AI Models uses neural networks, for the detections and predictions. Also, the functionality has been implemented for storing alarms in the Patients Data Repository and displaying them by the Human Machine Interface.

Requirement	FR.UC2#10
Owner	SANL
Evaluation Scope	UC2.3.3, UC2.3.4
Components(s)	SAP
Measurement Point(s)	PDF report extraction feature of SAP
Goal (Success Criteria)	Ensure that SAP can export reports that include all evidence of monitored assets (e.g., access requests to patient database).
Validation Strategy	Demonstration
Validation Plan	1. Setup monitoring of rule for specific action/asset 2. Generate events to trigger monitoring rule violation & satisfactions 3. Export PDF report, reviewing to verify all evidence and associated details are included in said report.
Validation Assessment	**Pass.** Successful demonstration in UC2.

Requirement	FR.UC2#11
Owner	Philips
Evaluation Scope	UC2.2.5
Components(s)	Human Machine Interface
Measurement Point(s)	Local AI, Human Machine Interface

Goal (Success Criteria)	The physician can manually set values, such as a maximum or minimum value, for all parameters measured as well as for different activity periods, such as exercising or resting, that will in turn trigger alarms when the parameters surpass them.
Validation Strategy	Demonstration
Validation Plan	1. The physician sets a maximum or minimum value for a specific parameter and activity period. 2. When a measured parameter surpasses the value, an alarm is triggered.
Validation Assessment	**Pass**. The implementation of setting values on a care plan in the Human Machine Interface has been completed for all key parameters of interest. With respect to exercise regimen-related care plan –which is highly individualized- parameters of interest include type of exercise, prescribed exercise time per session, number of exercise sessions per day and per week, minimum target heart rate (minTHR) and maximum permitted target heart rate (THR). If all target values are achieved and no safety threshold is deviated, the care plan is characterized as successful.

Requirement	FR.UC2#12
Owner	Philips
Evaluation Scope	UC2.2.4
Components(s)	Human Machine Interface
Measurement Point(s)	Human Machine Interface, Patients Data Repository
Goal (Success Criteria)	Diagrams with mean values of biometric parameters and their variation with the previous ones, which have been stored in the Patients Data Repository, are sent weekly to the clinicians through the Human Machine Interface.
Validation Strategy	Demonstration
Validation Plan	1. Biometric parameters are measured by patients and stored in the Patients Data Repository. 2. The Human Machine Interface calculates mean values and the variation in the parameters and displays them as diagrams.
Validation Assessment	**Pass**. Automated emails are sent to physicians with statistics and graphs.

Requirement	FR.UC2#13
Owner	Philips
Evaluation Scope	UC2.2.4
Components(s)	Local AI, Healthcare AI Models
Measurement Point(s)	Local AI, Healthcare AI Models
Goal (Success Criteria)	The Local AI and Healthcare AI Models provide personalized recommendations to patients based on previous behaviour and/or the behaviour of similar patients.
Validation Strategy	Demonstration
Validation Plan	1. The Local AI collects measurements of patients through time. 2. The Local AI processes the measurements using rule-based algorithms. The Healthcare AI Models process the measurements using neural networks. The output produces personalized recommendations.

	3. The recommendations are provided to patients.
Validation Assessment	**Fail.** Given the issues with the collected patient data, even getting a model that yields acceptable general predictions was impossible. Let alone getting a model in place that proves an even better, personalized output. See Appendix A: Healthcare Use Case – Data analysis and model development for detailed data analysis and model development efforts.

Requirement	FR.UC2#14
Owner	Philips
Evaluation Scope	UC2.1.1
Components(s)	Healthcare AI Models
Measurement Point(s)	Local AI, Global AI, Healthcare AI Models
Goal (Success Criteria)	Base Healthcare AI Models are stored in the Global AI and initially at the Local AI. The Healthcare AI Models implement neural networks and use data collected by the Local AI for training these.
Validation Strategy	Demonstration
Validation Plan	1. A Healthcare AI Model implements a neural network, which uses a given loss function. It is stored in the Global AI and initially also in the Local AI. 2. The Local AI collects data, which is then used by the Healthcare AI Model to train the neural network. 3. During training, the loss function is monitored for a number of training rounds.
Validation Assessment	**Pass**. A Healthcare AI Model has been implemented, which trains a neural network and provides its output to the Local AI.

Requirement	FR.UC2#15
Owner	Philips
Evaluation Scope	UC2.1.1
Components(s)	Global AI, Local AI
Measurement Point(s)	Global AI, Local AI
Goal (Success Criteria)	The Local AI can send training results, I.e., model weights, to the Global AI, which will aggregate the results and as appropriate update a model.
Validation Strategy	Demonstration
Validation Plan	1. A base Healthcare AI Model is available at the Global AI. 2. The Local AI collects measurements that the Healthcare AI Models use for training neural networks. 3. The training results, I.e., model weights, are sent by the Local AI to the Global AI. 4. The Global AI aggregates the results and updates as appropriate the base model.
Validation Assessment	**Pass**. A Healthcare AI Model has been implemented, which provides its output to the Local AI. The Local AI sends training results to the Global AI, where the results are aggregated to update the (base) model.

Requirement	FR.UC2#16
Owner	Philips
Evaluation Scope	UC2.1.1
Components(s)	Global AI
Measurement Point(s)	Global AI, Local AI
Goal (Success Criteria)	The Global AI performs model evaluation on a curated dataset to determine if the updated model performs better than the base or current model. If so, the updated model is distributed to the Local AI.
Validation Strategy	Demonstration
Validation Plan	1. A number of Local AI send training results, I.e., model weights, to the Global AI. 2. The Global AI aggregates the model weights, updates the model, and evaluates the model on a curated dataset. 3. If the updated model performs better than the current model, it is sent to the Local AI.
Validation Assessment	**Pass**. The Local AI can send training results to the Global AI. The Global AI can aggregate the results and perform model evaluation.

5.2 Validation and Evaluation of Non-Functional Requirements

Requirement	UCNFR.4
Owner	Philips
Evaluation Scope	UC2
Components(s)	Human Machine Interface
Measurement Point(s)	Human Machine Interface, Local AI, Patients Data Repository
Goal (Success Criteria)	The Human Machine Interface shall provide alarms to the physician if there are daily changes that fluctuate more than 20% in measurements taken by the smart blood pressure monitors.
Validation Strategy	Demonstration
Validation Plan	1. The Local AI collects measurements over time from the smart blood pressure monitor. 2. The measurements are stored in the Patients Data Repository. 3. The Human Machine Interface can generate alarms when the daily measurements fluctuate more than 20%.
Validation Assessment	**Pass**. These alarms are displayed on IntellIoT notifications page.

Requirement	UCNFR.5
Owner	Philips
Evaluation Scope	UC2

Components(s)	Local AI, Healthcare AI Models
Measurement Point(s)	Local AI, Healthcare AI Models
Goal (Success Criteria)	The Local AI provides notifications to the patients when weight, blood pressure and activity level are not the appropriate or when the patients are not adherent to their exercise regimen, lifestyle and use of the monitoring system. The Local AI and Healthcare AI Models implement algorithms to detect the levels and adherence.
Validation Strategy	Demonstration
Validation Plan	1. The Local AI collects measurements over time from the smart weight scale, smart blood pressure meter, and smartwatch. 2. The Local AI processes the measurements using rule-based algorithms. The Healthcare AI Models process the measurements using neural networks. 3. In case the measurements deviate from historical measurements, from those of similar patients, or from thresholds, a notification is provided to the patient.
Validation Assessment	**Pass.** Rule based processing of collected measurements and generation of alerts has been implemented and demonstrated. AI model output is sent to the clinician's dashboard for further evaluation. It was decided that AI model output never gets alerted to a patient without this clinical evaluation in the loop.

6. VALIDATION AND EVALUATION OF MANUFACTURING USE-CASE REQUIREMENTS

This section provides the validation and evaluation report for the requirements that are associated with IntellIoT's Manufacturing use-case (UC3) and covers both functional and non-functional requirements as specified in Deliverable D2.5.

6.1 Validation and Evaluation of Functional Requirements

Requirement	FR.UC3#1
Owner	HSG
Evaluation Scope	UC3.2.1
Components(s)	HyperMAS, Goal Specification Interface
Measurement Point(s)	Communication HyperMAS, End User Goal Specification Front End
Goal (Success Criteria)	Agents in the HyperMAS Infrastructure can send messages to the Goal Specification Interface to tell the user to reformulate the goal and the Goal Specification Interface displays this message
Validation Strategy	Demonstration
Validation Plan	1. Agent in HyperMAS Infrastructure sends a message 2. Goal Specification Interface receives the message 3. Goal Specification Interface displays the message
Validation Assessment	**Pass.** This has been shown during an integration meeting in Siemens Munich**.**

Requirement	FR.UC3#2
Owner	HSG
Evaluation Scope	UC3.2.1, UC3.2.2
Components(s)	Goal Specification Interface
Measurement Point(s)	Goal Creation of the Goal Specification Interface
Goal (Success Criteria)	The Goal Specification Interface displays the workpiece template and provides operations to change it, then confirm the input to start the creation process.
Validation Strategy	Demonstration
Validation Plan	1. The Goal Specification Interface displays the workpiece template 2. The user changes the workpiece template 3. The user confirms the input 4. The creation process starts
Validation Assessment	**Pass.** This has been shown during an integration meeting in Siemens Munich.

Requirement	FR.UC3#5
Owner	HOLO

Evaluation Scope	UC3.2.5
Components(s)	Interface, HIL-AR-Application, Robot Controller
Measurement Point(s)	Robot Controller, HIL-AR-Application
Goal (Success Criteria)	Provide the Robot Controller with the handedness (Right or Left) of the robot
Validation Strategy	Demonstration
Validation Plan	Operator provides the information to the Robot Controller through the established interface
Validation Assessment	**Passed.** HOLO stylus can be held with both hands.

Requirement	FR.UC3#6
Owner	HOLO
Evaluation Scope	UC3.2.5
Components(s)	AR Application, UR5 Controller
Measurement Point(s)	AR Application
Goal (Success Criteria)	Rendered robot is depicted in the same position as the real robot
Validation Strategy	Demonstration
Validation Plan	Initialize HIL operation
Validation Assessment	**Pass.** Real robot and digital twin appear in same position enabling the control feature.

Requirement	FR.UC3#7
Owner	Siemens
Evaluation Scope	Setup
Components(s)	• Workpiece storage • Workpieces
Measurement Point(s)	Workpiece storage
Goal (Success Criteria)	Workpiece storage shall be reachable by operator without entering the hazards area where the robot is working. The process always checks the state of the storage before accessing it, to that placing or removing workpieces in the storage does not trouble the production flow.
Validation Strategy	Demonstration
Validation Plan	Try to remove and add workpieces without crossing the safety barriers. Add and remove workpieces at arbitrary times during the running process.
Validation Assessment	**Passed.** The workpiece table can be reached from the back without accessing the robot's operation area. HyperMAS always uses AI to detect engraving areas and grab spots before it attempts to select a workpiece. After a workpiece was transported to a machine, AI is used again to detect the engraving area within the machine. Thus, even if the workpiece would be removed while the robot is already approaching it, the production process would fail without further threads or issues.

Requirement	FR.UC3#8
Owner	AAU
Evaluation Scope	UC3.3.3
Components(s)	DLT client/manager
Measurement Point(s)	DLT client
Goal (Success Criteria)	Check customer has installed MetaMask on his personal device such as laptop or mobile phone to communicate with DLT via Smart Contract to get the logs and data
Validation Strategy	Demonstration
Validation Plan	Check customer has installed MetaMask
Validation Assessment	**Pass.** MetaMask is installed in a laptop and can get the full content of the blockchain.

6.2 Validation and Evaluation of Non-Functional Requirements

Requirement	UCNFR.6
Owner	Siemens
Evaluation Scope	UC3
Components(s)	• Robot • Safety barriers
Measurement Point(s)	Observation
Goal (Success Criteria)	Avoid injury of employees or visitors
Validation Strategy	Demonstration
Validation Plan	Verify, that the robot cannot reach over the safety barriers
Validation Assessment	**Pass.** Safety barriers have been installed in the UC3 demo setup which avoid that employees or visitors accidentally enter the robot's operation area. Emergency buttons are available to stop the robot any time from outside the barriers.

Requirement	UCNFR.7
Owner	HSG
Evaluation Scope	UC3
Components(s)	Hypermedia MAS
Measurement Point(s)	Hypermedia MAS
Goal (Success Criteria)	Maximum reaction time Hypermedia MAS takes to acknowledge an update from a machine or robot shall be short enough to avoid resending of the update.
Validation Strategy	Demonstration
Validation Plan	Verify that the net response time (without network) it less than 500ms

Validation Assessment	**Pass.** Test have shown that the communication between machines and HyperMAS work flawlessly.

Requirement	UCNFR.9
Owner	All Edge App developers
Evaluation Scope	UC3
Components(s)	All Edge Apps
Measurement Point(s)	Edge App logs
Goal (Success Criteria)	The startup time of an edge app should be below 10 seconds
Validation Strategy	Observation
Validation Plan	Measure the start time, e.g. by examining the logs
Validation Assessment	**Pass.** The startup time can be read from the logs. On the further assessment of the application no malfunction was recognized which directs on slow initialization. Further, depending services are designed in a fashion that they can deal with temporary unavailability of single services.

7. VALIDATION AND EVALUATION OF OBJECTIVE KPIS

This section describes the evaluation criteria defined in IntellIoT's Description of Action to measure the success of the project. The evaluation is based on two main pillars: The first is a set of Key Performance Indicators (KPIs) that aim to assess the project's performance and ensure it meets the main objectives that have been set for the project. A second set of KPIs is defined in order to relate the project's achievements to the impacts that have been expected by the call.

In the subsequent sections, the KPIs for each objective are going to be presented and described, providing insights on how each KPI is associated with the different components of the IntellIoT framework, how the achievement of every KPI is going to be measured and validated and which are the baseline performances that need to be matched or surpassed. Lastly, an assessment will be provided to indicate whether the KPI target has been reached. Section 7.7 provides in a similar manner, a description of the impact KPIs and the assessment of their achievement or not.

7.1 Objective 1: Creation of a self-aware and semi-autonomous multi-agent system

The first objective of IntellIoT aims at the creation of a self-aware and semi-autonomous multi-agent system over an optimized computation and communication infrastructure that manages compositions of IoT/edge devices in closed loop with the network. The following KPIs are associated with this specific objective.

KPI-ID	1.1
Name	Open-Source HyperMAS
Description	Open-source software components for HyperMAS, including libraries and tooling for researchers and practitioners to design, deploy, and manage such systems.
Responsible partner(s)	HSG
Scope	IntellIoT Framework and UCs
Mapping to components & measurement points	This maps to the three main components that are created in the scope of WP3 T3.1: - Per use case, the End User Goal Specification Frontend (for Farmer/Doctor/Customer) - Across use cases, the Infrastructure for Hypermedia MAS - Across use cases, the Web-based Integrated Development Environment for Hypermedia MAS - Support libraries, e.g., for the handling of W3C WoT Thing Descriptions
Baseline	No HyperMAS software components that are published as open source.
Means of verification	Documentation
Methodology / tools	Publicly accessible repositories (e.g., GitLab/GitHub) with the code and instructions for the listed software components. Where applicable, whitepapers or research articles that document the listed software components.
Final Assessment/Evaluation	**KPI achieved**. Our components are publicly available at our GitHub project (https://github.com/Interactions-HSG/yggdrasil for the Hypermedia MAS Infrastructure, https://github.com/Interactions-HSG/goal-interface for the Goal Interface in UC3, https://github.com/Interactions-HSG/intelliot-hypermas-explorer for the Agent IDE). The final versions will also be put in the IntellIoT Gitlab project.

KPI-ID	1.2
Name	Vertical and Horizontally Scaling HyperMAS
Description	Reach high vertical and horizontal scalability of the HyperMAS with respect to increasing numbers of agents, artifacts, and devices with a benchmark against existing MAS
Responsible partner(s)	HSG
Scope	IntellIoT Framework
Mapping to components & measurement points	Ability of the Hypermedia MAS Infrastructure and the Web-based IDE for Hypermedia MAS to support large numbers of agents and artifacts (i.e., devices and services) compared to existing MAS
Means of verification	Testing
Methodology / tools	Evaluation of the relevant components in the laboratory (using simulated services) as well as in the context of the IntellIoT use cases.
Baseline	Baseline values depending on what MAS framework is taken as reference.
Final Assessment/Evaluation	**KPI target achieved**. The HyperMAS Infrastructure can handle the number of agents and artifacts required by the different use cases. It can also support many more agents and artifacts (potentially unlimited by deploying many different interoperable instances).

KPI-ID	1.3
Name	HyperMAS Deployments
Description	Provide deployments (within the three use cases) of the HyperMAS and evaluate (with existing benchmarks) system flexibility & evolvability (e.g., adapt to dynamic environments).
Responsible partner(s)	HSG
Scope	IntellIoT Framework and Use Cases
Mapping to components & measurement points	- Running demonstrators across all three IntellIoT use cases, given the scope of each UC and the applicability of HyperMAS components in this scope. - Ability of the Hypermedia MAS Infrastructure and the Web-based IDE for Hypermedia MAS to apply across the heterogeneous use cases of IntellIoT.
Means of verification	Testing
Methodology / tools	Ability of the Hypermedia MAS Components (Web-based IDE and Infrastructure) to support the IntellIoT use cases without any changes, i.e., without use-case-specific adaptation. Tested through the implementation of demonstrators for the IntellIoT use cases that use equivalent and unchanged Hypermedia MAS Components.
Baseline	N/A
Final Assessment/Evaluation	**KPI achieved**. Our integration with UC1 and UC3 shows that the HyperMAS system can be extended with additional services across use cases without modifications to the infrastructure. This has specifically been demonstrated in the scope of a test of the

	Agent IDE in our lab that showed that the time effort to deploy agents on a fictional scenario that unites mobile robots (from UC1) and a robot arm (from UC3) using our systems is under 5 minutes.

KPI-ID	1.4
Name	HyperMAS Reconfiguration
Description	Initial configuration as well as reconfiguration of the HyperMAS (e.g., in case of failures) will be *real-time* enabled, where concrete real-time requirements will be defined per use case.
Responsible partner(s)	HSG
Scope	IntellIoT Framework
Mapping to components & measurement points	- Ability of the Web-based IDE for Hypermedia MAS to allow engineers to design and re-design agent organizations and agent procedural knowledge while the system is running - Extent of the ability to deploy such changes while the system is running (not all changes can be deployed in this case)
Means of verification	Testing
Methodology / tools	Based on tests for KPI 1.3, ability of the Hypermedia MAS Components to allow engineers to design and re-design agent organizations based on available services and ability of the Hypermedia MAS Infrastructure to run the resulting organizations. Tested by adding/removing heterogeneous services at run time and verifying that engineers can update the agent organization to keep a running system (as long as this is feasible). This should be refined given the results from WP5.
Baseline	N/A
Final Assessment/Evaluation	**KPI partially achieved**. Changes to the system can be made at runtime. However, this feature has not been shown during the integration meetings where changes to the system at runtime were not considered.

KPI-ID	1.5
Name	Optimized allocation
Description	Optimized allocation of IoT application functions with a maximum optimality gap of 15% and implementation for 3 optimality criteria: reliability, response time and energy consumption.
Responsible partner(s)	Siemens
Scope	UC3
Mapping to components & measurement points	Determining the optimal allocation of IoT application functions will be done by the Computational Resource Manager. It comprises an algorithm for the optimized allocation of IoT application functions on IoT devices / edge resources while considering the network configuration. The deployment of such functions will be tuned towards different criteria, e.g., reliability of compute nodes, response time of the application, or energy consumption. Together

	with the Edge Orchestrator, this will automatically deploy functions of a composed IoT application to the Edge infrastructure.
Means of verification	Testing
Methodology / tools	A stepwise testing approach will be chosen: 1. Model based test: relying on models for network, devices and application functions upon which the algorithm for optimal allocation is applied. 2. Simulation based test: based on a selected simulation environment (e.g., Omnet++) network and device infrastructure will be simulated and allocation options will be tested. 3. Operational test: using the actual network/device infrastructure of the UC3 demonstrator, tests will be conducted with different component allocations.
Baseline	The baseline is a set of randomly allocated IoT application functions.
Final Assessment / Evaluation	We developed the Computational Resource Manager to optimize the allocation of Edge Apps by minimizing energy consumption and response time between applications. This optimization process considers various factors such as device resources, Edge App requirements, network connections, and communication needs to find the most optimal solution for the task allocation problem. Within the scope of D4.5, we present and explain three algorithms for solving the task allocation problem: Integer Linear Programming (ILP), Particle Swarm Optimization (PSO), and Deep Reinforcement Learning (DRL). While PSO and DRL do not achieve the desired 15% optimality gap for large networks, ILP does, as it always returns the most optimal result. However, we find that the execution time of the ILP solution is over 14 days in the worst-case scenario with around 140 nodes in the network, whereas PSO took only 15.9 seconds and DRL completed in 189 milliseconds. These findings have been published by Buschmann et al. [1]. Since our ILP approach focuses on minimizing energy consumption and does not consider response times, we created a Mixed Integer Programming (MIP) approach. This MIP approach reduces the execution time in comparison to ILP to approximately five days in the worst-case scenario and incorporates the response time constraints between Edge Apps. This improves significantly upon the ILP solution while still providing the most optimal solution (0% optimality gap). **KPI target partially achieved.** [1] P. Buschmann, M. H. M. Shorim, M. Helm, A. Bröring, and G. Carle, "Task Allocation in Industrial Edge Networks with Particle Swarm Optimization and Deep Reinforcement Learning," in *Proceedings of the 12th International Conference on the Internet of Things*, in IoT 2022. New York, NY, USA: Association for Computing Machinery, Jan. 2023, pp. 239–247. doi: 10.1145/3567445.3571114.

7.2 Objective 2: Enable ultra-reliable low-latency communication over heterogeneous networks

The second objective of the project aims to enable dynamic network planning/management for ultra-reliable low-latency communication schemes over heterogeneous networks (LTE, 5G NR, Cellular IoT, D2D) to achieve tactile (real-time) and contextual (adaptive) interaction between IoT devices, humans and services. The KPIs that follow are associated with the specific targets.

KPI-ID	2.1
Name	5G URLL Communications
Description	Extending 5G network functionalities supporting URLL & eMBB for the needs of the three use cases.
Responsible partner(s)	EURECOM
Scope	UC1 / UC3
Mapping to components & measurement points	The OAI FLEXRIC component is in charge of configuring the required radio resources from the Communication Resource Manager & Edge controller. The FlexRIC itself contains the configuration parameters and a data-driven optimizer will need to dynamically adjust the right parameters to the wireless conditions.
Means of verification	Testing
Methodology / tools	• Unit tests (PHY layer simulation) • Laboratory test (1-2 devices)
Baseline	Delay <10ms on the RAN
Final Assessment/Evaluation	**KPI target achieved.** 5G RAN URLL for UC1 and UC3 showed <10ms latency for FR1. FR2 tests also showed latency <<10ms.

KPI-ID	2.2
Name	TSN functions integration
Description	TSN functions integration in computation & communication infrastructure (combined with 5G).
Responsible partner(s)	Siemens, EURECOM, HOLO
Scope	IntellIoT framework and UC3
Mapping to components & measurement points	TSN functions are implemented in TSN switches and endpoints and controlled by the Network Controller. Network controller is mainly triggered by HIL service. HIL service integrates TSN with 5G communication services. Measurement points: • Low latency and low jitter communication between robot controller and robot arm. • High bandwidth and low latency for camera stream to human operator. • Dynamic allocation of communication services between robot and human operator.
Means of verification	Testing

Methodology / tools	Test probes are deployed to relevant end stations. These allow for end-to-end measurement of Quality-of-Service parameters.
Baseline	Current TSN-based networks are vendor-locked and have very limited support for dynamic reconfiguration. No close interaction between 5G and TSN networks.
Final Assessment/Evaluation	**KPI partially achieved**. TSN controller is deployed as an edge app. It discovers all devices supporting LLDP and derives actual topology information. Via a REST interface to the 5G core, it even discovers 5G UEs, e.g. the mobile phone used in UC3. According to communication service requests received from its northbound interface, e.g. from the edge orchestrator, it computes a network schedule and reserves network resources. If 5G links are involved in a path, it triggers communication resource manager to reserve a 5G slice via the aforementioned REST interface. All of the above are implemented and tested, and we have shown that TSN schedules as well as 5G slices improve delay, jitter and packet loss for the reserved communication services. Unfortunately, the 5G core in use gets instable some minutes after a slice is established, that's why we currently cannot use slicing in the final demo.

KPI-ID	2.3
Name	5G Multi-RAN
Description	Enabling heterogeneous networking technologies: LTE, 5G NR, Cellular IoT, D2D.
Responsible partner(s)	EURECOM
Scope	UC1 / UC3
Mapping to components & measurement points	The 5G Communication Resource Manager can support various RAN technologies. The 5G RAN and 5G Core can operate in a stand-alone or none-standalone mode (either with a 4G backend or with a full 5G backend). The non-standalone mode supports both 5G and LTE UEs.
Means of verification	Testing
Methodology / tools	Cycle 1 • 5G non-standalone mode – unit test + lab test Cycle 2: • 5G standalone and non-standalone modes – unit test & demonstrators.
Baseline	As function of the configured connection choice, the right mode is used by the UE.
Final Assessment/Evaluation	• **KPI achieved** - 5G non-standalone & 5G standalone FR1 isolated (one or the others) • **KPI achieved** – 5G Standalone as function of Communication Manager request

KPI-ID	2.4
Name	Industrial D2D
Description	Enabling wireless URLL D2D scheduler for decentralized computing in IoT context.

Responsible partner(s)	EURECOM
Scope	UC1 / UC3
Mapping to components & measurement points	A D2D RAN controller (called ProSE controller in 3GPP) configures the D2D wireless link. A 5G D2D scheduler is provided as an external module to the 5G 3GPP architecture. The 5G D2D scheduler provides quasi-deterministic resource allocation on the D2D link respecting the requirements configured by the D2D RAN controller
Means of verification	Testing
Methodology / tools	• Simulation - first D2D scheduler is developed and the feasibility to provide URLL D2D is tested on a network simulator supporting the 3GPP 5G stack (rel.16). • Simulation - A 5G D2D Slice Admission Control is developed to limit the number of participating 5G UE in D2D communication to keep the URLL KPIs. • Simulation – a 5G D2D RAN controller creates a 5G URLL slice and coordinate group management • Simulation – an improved scheduler is developed that can jointly handle group and scheduling • Simulation – realistic challenging mobility and topology (channel, density) is created to evaluate the 5G URLL D2D capability.
Baseline	D2D resources are allocated with delay and reliability guarantees
Final Assessment/Evaluation	**KPI target achieved.** 5G URLL supporting < 1ms RAN delay with 10-4 reliability with 8 robots; Reliability increases to 10-5 if delay is <10ms

KPI-ID	2.5
Name	Application-tailored reliability
Description	Application-tailored definition and fulfilling of a reliability requirement for the three use cases, maintained under challenging network conditions and based on a data-driven prediction.
Responsible partner(s)	AAU
Scope	UC3
Mapping to components & measurement points	Dynamically determining the communications resource allocation of IoT application functions will be done by the Communications Resource Manager. It comprises an algorithm for the optimized allocation of VR/AR. The deployment of such functions will be tuned towards different reliability-related criteria, e.g., information freshness, or latency-reliability curves.
Means of verification	Testing
Methodology / tools	Two steps: 1. Model test with simulations: relying on models for network, devices, upon which the algorithm for optimal allocation is applied. Selection of a simulation environment (e.g., Python)

	2. Operational test: using the actual 5G infrastructure, tests will be conducted with different traffic and allocations.
Baseline	The baseline is a set of randomly allocated resources.
Final Assessment/Evaluation	**KPI target achieved**. The communication resource manager allows setting the throughput and latency requirements, and the functionality has been demonstrated with one mission-critical user (latency 10 ms) and a broadband user (10 Mbit/s of average throughput) sharing resources. The results are published in a conference paper.

7.3 Objective 3: Semi-autonomous IoT applications with distributed AI while keeping human-in-the-loop

The third objective of the project is to enable semi-autonomous IoT applications by leveraging distributed AI algorithms under compute, storage, mobility, and bandwidth constraints and by integrating the human-in-the-loop for safety, assistance and continuous improvement of AI. The following KPIs are associated with this specific objective.

KPI-ID	3.1
Name	Distributed training guarantees
Description	With a centralized training model as benchmark for training accuracy, enable distributed solutions achieving an accuracy beyond 95% of benchmark within 50 inter-device communication rounds.
Responsible partner(s)	UOULU
Scope	At least one out of UC1 and UC3
Mapping to components & measurement points	Resource-aware re-trainer, AI model aggregator, Agriculture/Manufacturing AI model
Means of verification	Testing
Methodology / tools	Proposed: Implementation of federated learning algorithm and carry out distributed training over the system. Benchmark: All devices upload training data to a server where centralized training is carried out. V1: training over artificial data, V2: training over UC-specific data
Baseline	Training accuracy of centralized training where the training takes place in a server after importing data from all devices.
Final Assessment/Evaluation	**KPI target achieved.** Validated with MNIST, CIFAR, and our own simulator-based datasets. The results are published in several articles. During Cycle 2, the validations are carried out with UC3 based augmented data. Here, the augmented dataset is partitioned and divided among 10 virtual FL clients. Then all clients train a single model with the aid of a server using the FL algorithm (frequent training on local datasets while occasionally sharing the local models with the server for model averaging). The accuracy of FL model is compared with the accuracy of centralized model after 50 model averaging steps. The results are shared in D3.6.

KPI-ID	3.2
Name	ML model reduction
Description	At least 10 times ML model size reduction during knowledge distillation for constrained IoT devices.
Responsible partner(s)	UOULU
Scope	At least one out of UC1 and UC3
Mapping to components & measurement points	Resource-aware ML inference, Agriculture/Manufacturing AI model
Means of verification	Testing
Methodology / tools	Use the pre-trained model as the baseline Minimize ML model size subject to ML model accuracy above a predefined target
Baseline	Pre-trained model size prior to model reduction
Final Assessment/Evaluation	**KPI target achieved.** Validated with the UC3 specific model training and UC1 related model training at the lab setting. The minimized models are integrated in the demo settings. The size of the original model in UC3 was about 1.2GB and as for the first step, the size is reduced to 350MB. The size of the final integrated UC3 model is 138MB.

KPI-ID	3.3
Name	Enabling autonomy
Description	Ensure the frequency of necessary human interventions reduces at an exponential rate over time.
Responsible partner(s)	UOULU
Scope	At least one out of UC1 and UC3
Mapping to components & measurement points	Resource-aware re-trainer, AI model aggregator, Agriculture/Manufacturing AI model
Means of verification	Testing
Methodology / tools	At the interventions, human inputs are added into the training dataset Retrain the ML models with modified training data
Baseline	None
Final Assessment/Evaluation	**KPI target achieved with alterations.** Validated with the augmented data from UC3 and model training therein in a simulated environment. The choice of the evaluation in a simulation setting is due to the need of large number of interventions (minimum of few hundreds of failures) is impractical to conduct in the demo setting. Here, the training dataset is partitioned into subsets. The initial model is trained with a subset and sequentially tested until about 100-150 mistakes are made. The samples correspond to the mistakes along their labels are used as new data for the retraining. This procedure is carried out several times and the results correspond to the exponential decay of the need of interventions are disseminated in D3.6.

KPI-ID	3.4
Name	ML energy reduction
Description	At least 10% of energy reduction during distributed training while ensuring high training accuracy compared to centralized training methods.
Responsible partner(s)	UOULU
Scope	At least one out of UC1 and UC3
Mapping to components & measurement points	Resource-aware re-trainer, AI model aggregator, Agriculture/Manufacturing AI model
Means of verification	Testing
Methodology / tools	Proposed: Implementation of federated learning algorithm and carry out distributed training over the system. Benchmark: All devices upload training data to a server where centralized training is carried out. V1: training over artificial data, V2: training over UC-specific data
Baseline	Total energy consumption of centralized training where the training takes place in a server after importing data from all devices.
Final Assessment/Evaluation	**KPI target achieved.** Validated with MNIST, CIFAR, and our own simulator-based datasets. The results are published in several articles. In Cycle 2, the validation is carried out with UC3 model training. The trained model is integrated in the demo setting. During the validation, the model training is allowed to be carried out by splitting the first half to be trained at the client and the rest at the server. The reduction of energy is measured with respective the client end.

KPI-ID	3.5
Name	Increase accuracy of HOLO Stylus
Description	Reach an accuracy of the Holo-Stylus below 5 mm in the field of view of the operator.
Responsible partner(s)	HOLO
Scope	UC3
Mapping to components & measurement points	HMU application on HoloLens; ISAR;
Means of verification	Testing
Methodology / tools	Normative-Actual value comparison.
Baseline	N/A
Final Assessment/Evaluation	**Pass.** The accuracy of the Holo-Stylus accuracy is 3 mm, which is below the 5 mm KPI mark.

7.4 Objective 4: Enable security, privacy and trust-by-design

IntellIoT aims to enable security, privacy and trust-by-design with continuous assurance monitoring, assessment and certification as an integral part of the system, providing trustworthy integration of third party IoT devices and services. The following KPIs are associated with the effort to achieve those goals.

KPI-ID	4.1
Name	Continuous Assurance
Description	Delivery of a continuous assurance and certification component supporting: (a) individual risk assessment schemes; (b) incremental risk assessment schemes, and; (c) hybrid risk assessment schemes to estimate risk by combining the outcomes of the schemes in (a) and (b).
Responsible partner(s)	SANL
Scope	IntellIoT framework
Mapping to components & measurement points	Security Assurance Platform
Means of verification	Testing
Methodology / tools	Testing will cover each and every one of the core features of the Security Assurance platform that will be integrated into IntellIoT to satisfy this KPI. In more detail, it will cover: (i) each one of the individual risk assessment schemes, including: (a) vulnerability assessments; (b) static analysis; (c) penetration testing, and; (d) continuous runtime monitoring. (ii) Incremental risk assessment schemes, featuring mechanisms to will allow to build and elaborate upon previous assessments. (iii) Hybrid risk assessment, which will combine and correlate results from all the above assessments, providing a multi-perspective analysis of the security and privacy posture of the protected systems. For testing purposes, a mock IoT infrastructure model will be developed and used, while a comprehensive validation of the above will also follow, in the context of all three of the project's use cases.
Baseline	N/A
Final Assessment / Evaluation	Delivery, assessment & demonstration of (i), (ii) & (iii) above; see D5.5 & UC demos. **KPI target achieved**.

KPI-ID	4.2
Name	DLT implementations
Description	Delivery of at least 2 DLT implementations that can adjust level of trust to capabilities of devices, can integrate proxies, and can conform to certain latency and reliability requirements, such that the level of decentralization of device participation is proportional to its computation-communication capabilities.
Responsible partner(s)	AAU

Scope	UC1/UC2/UC3
Mapping to components & measurement points	Resource aware DLT-manager
Means of verification	Simulations + experimental testing
Methodology / tools	Two phases: 1. Simulation test: relying on models for network, devices and application functions. The steps are: a. Test with different BC platforms: Ethereum, Hyperledger fabric, IOTA and find the proper one for IntellIoT. b. Check that the system works correctly. c. Implement two protocol options: (1) lightweight nodes; (2) standard devices (no DLT-capable). d. Do performance measures: latency, throughput, robustness. e. Design APIs and integrate with other partners. 2. Experimental test: using the actual nodes for each UC (tractor, robots, wearables…). The same steps as for the simulation test phase are followed here.
Baseline	Not applicable: the baseline is not having DLT
Final Assessment / Evaluation	**KPI target achieved.** There are 3 DLT implementations, one per UC, and the functionality and transactions registered in the UCs have been demonstrated.

KPI-ID	4.3
Name	Secure Routing for Ad-Hoc IoT Networks
Description	Delivery of at least two trust-based secure routing algorithms, applicable for the IoT system, which cover the design requirements of i) relatively static networks with low mobility and ii) open networks with high mobility nodes, respectively.
Responsible partner(s)	TSI
Scope	UC1
Mapping to components & measurement points	Secure Routing / Intrusion Detection System components
Means of verification	Testing
Methodology / tools	The goal is to deliver routing algorithms for adhoc networks that can integrate trust metrics for their routing decisions. Methodology and tools employed: In the first stage of development, different routing algorithms for adhoc networks are evaluated both by what is reported in the literature and through simulations performed in a state-of-the-art network simulator (NS3). Once a routing algorithm is chosen, the simulation topology is transferred to a real-world setup comprised of Raspberry Pi devices to be validated in a testbed.

	As part of the development of other components within IntellIoT, a trust-based IDS has been developed. This has been chosen as the primary source of trust measurement, as it can provide trust indicators for all directly connected nodes to the routing algorithm. The routing algorithm will be either directly or indirectly connected to the trust-based IDS system, so that malicious nodes will not be used to relay traffic. Direct connection: the routing algorithm will be modified to account for the trust-related information as part of its decision process. Indirect connection: the routing algorithm works in parallel with the trust-based IDS. The latter is employed as "transparent" firewall (L2 firewall) functioning at the same level with the routing algorithm, thus blocking access to malicious nodes and therefore forcing the algorithm to update routing decisions as connection is blocked with the offending nodes. The functionality is tested by hosting multiple entities to the network in an Ad-Hoc topology. Most of the entities generate legitimate traffic, and some entities work as malicious entities and try to compromise the network. The Intrusion Detection System will discriminate legitimate networking entities from malicious ones. The chosen routes for each legitimate entity are monitored and analyzed to verify that malicious nodes are skipped and there is no data loss or breach. Several different attacks will be performed, and the successful identification to each attack and rerouting of traffic will be tested.
Baseline	N/A
Final Assessment / Evaluation	After completing the first stage of deployment, testing and evaluation of relevant routing algorithms, we adopted the B.A.T.M.A.N.-advanced (batman-adv) routing algorithm for adhoc networks. The trust measurement mechanism is offered by the trust-based IDS system developed within IntellIoT. The indirect connection approach mentioned above has been adopted, in which the trust-based IDS system influences the operation of the routing algorithm by removing untrusted nodes from the active communications (through use of ebtables – L2 firewall – that block all message exchanges with the blacklisted nodes). In absence of communication ability with those nodes, the routing algorithm updates its operation to use different routes that do not include the offending nodes. Overall, the tightly integrated system comprised of the Trust-based IDS, the L2 firewall and the B.A.T.M.A.N. routing scheme, form a secure routing protocol based on trust metrics. It should be noted that based on the principles of the trust-based IDS system, the secure routing scheme adopted makes localized decisions. That is, each node that participates on the network makes blocking decisions for itself and does not exchange this information with other nodes. The routing scheme was deployed in Raspberry Pi 4 devices to build a real-world low mobility/stationary testbed. Currently, Ethernet and Wi-Fi have been successfully tested as physical layers. The secure routing protocol has not been tested in an actual real-world scenario that involves high mobility nodes. The principles of the secure routing protocol as it has been implemented do not demonstrate any real open issues that may prove problematic in its deployment in that specific scenario. However, there are trust metrics that may be affected, e.g. an increased packet loss rate, as a result of disconnections due to the high mobility. As such, the network measurements that are

	taken into account to compute trust metrics (and therefore the potential attacks that the trust-based IDS system is tuned to detect) may have to be adjusted. Since the high mobility scenario has not been extensively tested, the **KPI can be considered partially achieved**.

KPI-ID	4.4
Name	Moving Target Defenses for IoT systems
Description	Development of at least 2 MTD algorithms for: i) local decision making by an individual agent for its underlying system, and ii) horizontal incorporation of trusted agents in the IoT system.
Responsible partner(s)	TSI
Scope	IntellIoT framework
Mapping to components & measurement points	Moving Target Defenses component - Security Assurance Platform
Means of verification	Testing
Methodology / tools	Two cases will be tested: one during normal operation of the overall system and one for operation of the system under ongoing attacks. During normal operation the MTDs will proactively change the system's configuration at predefined or random intervals. The goal is to increase the required effort for an attacker to analyze a specific system configuration, exploit potential vulnerabilities or gather enough information about the system. During ongoing attacks, the agents must enforce pre-defined defense strategies that can counter malicious actions or at least mitigate their side effects. The generation of new configurations will be tested to verify the validity of each new configuration. The predefined set of defense strategies will be tested for the corresponding attacks and the mitigation will be verified.
Baseline	Not applicable
Final Assessment / Evaluation	The MTD system that has been implemented, is comprised of an MTD server module and MTD clients that can be installed in all devices participating in the network. The MTD modules are dockerised to facilitate deployment. Native implementations are also available. The MTD system enables changes in the network configuration and topology (e.g. IP addresses, change of subnets etc) as well as changes in communication aspects (for example using unencrypted or encrypted communications or hoping between different encryption algorithms). A communication mechanism with other security components (namely SAP and trust-based IDS) through a broker system has been implemented and the relevant integration has been completed. An API has been implemented to enable integration with the TSN controller. The MTD performs proactive or reactive mechanisms to counter threats. Proactive mechanisms change network parameters periodically, while reactive mechanisms

	force changes according to security events (signaled by the other security components). Completed testing using a large number of docker instances to simulate a large network and validated the correct functioning of the system. Performance estimations and measurements have been performed both in the simulated scenarios setups as well as in real-world deployments (in IntellIoT's Use Case 1 and in a lab environment). The MTD system has been integrated into IntellIoT's real-world use cases. Through this process it has been successfully tested and validated. Additionally, the MTD has also been employed in systems not associated with the IntellIoT project through the Open Call 2 process. The related work has been published in a conference paper (T. Kyriakakis, S. Ioannidis, "A Moving Target Defense Security Solution for IoT Applications", 19th International Conference on Design of Reliable Communication Networks (DRCN), Spain, 2023) **KPI Target Achieved**

7.5 Objective 5: Development of a reference implementation of the IntellIoT framework

Under the specific objective, the goal is to develop a reference implementation of the IntellIoT framework, demonstrated and evaluated in three use-case areas: agriculture, healthcare and manufacturing. The KPIs that are provided below are used as a measurement to assess the success of this effort.

KPI-ID	5.1
Name	IntellIoT Framework
Description	Delivery of the integrated *IntellIoT* framework.
Responsible partner(s)	AVL
Scope	IntellIoT framework
Mapping to components & measurement points	All components
Means of verification	Testing
Methodology / tools	The three Use Cases provide the testing ground for the overall IntellIoT framework. Each Use Case is designed to make use of a large number of IntellIoT components (if not all) and the usage scenarios per use case are designed to demonstrate the functioning of the integrated framework as they depend on the collaborative functioning of multiple components to be successfully completed.
Baseline	N/A
Final Assessment/Evaluation	**KPI target achieved**. All components were integrated into their respective target environments and tested as single framework in all three use cases.

KPI-ID	5.2
Name	End-user workshops

Description	Involve end-users of three domains and conduct (at least) two requirements' workshops.
Responsible partner(s)	STARTUPC
Scope	UC1 / UC2 / UC3
Mapping to components & measurement points	Not applicable
Means of verification	Documentation
Methodology / tools	The end-user workshops were conducted in a co-creative setup, using a design thinking approach to capture end-user's needs, pain points, requirements and expectations regarding the IntellIoT applications. Input from these workshops was collected and synthesized in a deliverable report and used for the refinement of the Use Cases.
Baseline	Not applicable
Final Assessment/Evaluation	During the first cycle of the project three end-user workshops took place: Healthcare (16. April 2021), Manufacturing (15. April 2021) & Agriculture (12. April 2021). Fifteen experts attended the end-user workshops where information about their special needs were gathered. All online workshops were recorded with the consent of the attendees, transcribed and its content reviewed to align it with upcoming activities with the aim to apply the IntellIoT framework in the most user-centric way. Learnings from the end-user workshops were also applied during Open Call 1. At a later stage, during the second cycle of IntellIoT, end users were involved in two ways: (i) they were invited to join special exploitation workshops to share the applicability of the IntellIoT framework from a business point of view and (ii) end-user feedback was also collected via interviews at different events, like the Agritechnica Event. Based on the above the **KPI target can be considered achieved**.

KPI-ID	5.3
Name	Framework validation
Description	Validation of the *IntellIoT* framework in three use cases each in relevant environment.
Responsible partner(s)	TSI
Scope	IntellIoT framework
Mapping to components & measurement points	All IntellIoT framework components and KPIs
Means of verification	Demonstration
Methodology / tools	Within IntellIoT, three real-world use cases were designed, each one with a number of usage scenarios aiming to test and validate the overall IntellIoT framework, as their successful completion required the collaborative operation of numerous IntellIoT components. A series of functional, non-functional and technical requirements were derived to design and validate the framework. While each scenario and use case had its own intricacies and dedicated requirements, a significant overlap both in the technical requirements and the integration aspects exists, aiming to prove the generality of the

	approach and that the developed framework can be successfully applied in multiple application domains. Therefore, the process to validate the overall IntellIoT framework relies on the validation of all the requirements and KPIs presented in deliverable D5.6 and the final successful deployment and demonstration of the three use cases and their scenarios.
Baseline	N/A
Final Assessment	The project has set 115 general requirements for all aspects of the framework (including 18 functional, 16 non-functional and 81 technical requirements), along with 40 additional requirements specifically targeting the different domains of each use case. According to what is reported in the previous Sections (Sections 3 to 6), 16 out of the 18 general functional requirements were successfully covered by the final implementation of the framework and its components, while two (2) requirements were partially achieved. The vast majority of the general non-functional requirements is also successfully covered by the final framework (13/16), while only 3 requirements were partially achieved, and none failed to be covered. Concerning the technical requirements, 72 out of 81 requirements were successfully validated, with the remaining nine (9) partially covered and none failed. Similarly for the use case specific requirements, the vast majority is successfully covered (12/14 passed and two partial accomplishments for UC1, 15/17 passed, one partial and one fail for UC2, all requirements passed for UC3). Concerning the KPIs related to the components and the overall framework (reported in Objectives 1, 2, 3, 4 and 5), all targets set have been achieved either in their full context (17/21) or at least partially (4/21) and no KPI target has been missed. According to the above, the IntellIoT framework can be considered validated and the **KPI can be considered achieved**.

7.6 Objective 6: Promotion and exploitation of the IntellIoT framework

This objective is associated with the promotion and exploitation of the IntellIoT framework. This has been primarily done through contribution to standards and delivery of open-source components as well as by building an active IoT ecosystem (supported by three Open Calls) and focused dissemination and exploitation activities. The following list provides the KPIs associated with these goals.

KPI-ID	6.1
Name	Open Calls
Description	Successful conduction of two Open Calls with a minimum of 40 applicants per Open Call.
Responsible partner(s)	STARTUPC
Scope	IntellIoT framework and UC1 / UC2 / UC3
Mapping to components & measurement points	Not applicable
Means of verification	Documentation
Methodology / tools	Applications to the two large open calls were managed through a central tool (F6S), which included the applicant's general information, detailed description of the planned

	contribution and technology, as well as the evaluation by external experts. The special Open Call 3 war executed as a hackathon via the platform taikai.
Baseline	N/A
Final Assessment	Open Call 1, Open Call 2 and Open Call 3 have been conducted successfully in 2021, 2022 and 2023 with 400+ applications in total. The final results are reported in D6.11, final report Open Call coordination & results. Thus, the **KPI can be considered achieved**.

KPI-ID	6.2
Name	Dissemination and communication
Description	Achieve dissemination and communication targets as defined in Description of Action.
Responsible partner(s)	STARTUPC
Scope	IntellIoT framework and UC1 / UC2 / UC3
Mapping to components & measurement points	Not applicable
Means of verification	Documentation
Methodology / tools	All dissemination and communication activities are monitored on a constant basis and results captured in a central dissemination file.
Baseline	Not applicable
Final Assessment	All dissemination and communication activities have been reported in Deliverable D6.2 and Deliverable D6.9 (Dissemination and Ecosystem Building, final version) and are inline or exceeding set targets. The **KPI has been achieved**.

7.7 KPIs related to expected impact of the project

IntellIoT strived to address all impacts expected by the Next Generation Internet of Things call and therefore defined a set of impact KPIs (iKPIs) for transparent performance assessment of the project. More specifically, these are six main impacts expected to be addressed, namely:

1. Contribution to human centered IoT evolution improving usability and user acceptance, notably through strengthened security and user control
2. Contribution to emerging or future standards and pre-normative activities
3. Log-term evolution of next-generation IoT infrastructures and service platform technologies and contribution to scientific progress enabling novel, future semi-autonomous IoT applications
4. Proposal of novel and disruptive business models
5. Mobilization of key IoT players in security and privacy
6. Maintenance of an active ecosystems of all relevant IoT stakeholders

For each of the above six expected impacts a number of iKPIs have been defined. The iKPI IDs are used to identify uniquely each iKPI and the numbering associates each iKPI with a specific impact (e.g., iKPI-ID i2.1 is the first iKPI associated with expected impact 2: contribution to emerging or future standards and pre-normative activities).

iKPI-ID	i1.1
Description	At least 70% of end-users of the developed IoT applications (based on IntellIoT framework) state they trust and accept the smart system, based on established metrics
Lead partner(s)	PAGNI
Mapping to components & measurement points	IoT applications in the use cases
Means of verification	Inclusion & analysis of questionnaire and/or online poll results in D5.6.
Methodology / tools	Questionnaires prepared by the consortium and distributed during the end-user workshops. Post-workshops' analysis of results.
Baseline	N/A
Final Assessment	**UC2-specific data**: For the evaluation of this KPI, the end-user group consisted of the patients enrolled in the UC2 pilot study, as they were the only persons to actually use the e-health system developed during the project. As of 10/1/2024, a questionnaire that was specifically designed to assess the degree of acceptance of- and trust in the e-health system on behalf of the patients has been distributed to- and answered by 18/19 participants. This questionnaire consists of 10 multiple choice questions, defined to address: i. perceived user-friendliness, ii. perceived system complexity, iii. perceived responsiveness, iv. sense of personal data security, v. effect on management of patients' health, vi. effect on patients' relationship with their attending physician, vii. perceived potential of the system to prevent unnecessary healthcare visits, viii. effect on overall sense of safety, ix. users' estimations regarding trends in use of similar e-health systems in the future, x. likelihood of recommending the system to other patients. Each question could be answered by selecting one out of 5 possible answers, graded from 1 to 5 points, and the total score was summed (minimum score: 10/50; maximum score: 50/50). A score of at least 30/50 was set as the lowest threshold of system acceptance and trust. *Results:* In 17/18 completed questionnaires, a score of >30/50 has recorded, thus KPI i1.1 is already successfully assessed in this end-user category (>70% of users have stated

	they trust and accept the smart system, even if all pending questionnaires register scores <30/50). The full questionnaire, as the comprehensive outline of answers and the respective score per patient is presented in Appendix B: Questionnaires . **Trustworthiness-related data**: Using a different questionnaire (see also Appendix B) targeting industry stakeholders and researchers working on the field of NG-IoT, all of the survey participants (25 out of 25) believe that the inclusion of trustworthy IoT capabilities, as proposed by IntellIoT, would facilitate the deployment of Next Generation IoT applications in verticals domains. Furthermore, 23 out of 25 correspondents (92%) agree or strongly agree that the trust-related technologies provided by IntellIoT help cover the security & privacy needs of NG-IoT applications. According to the above results, we consider that the **iKPI has been achieved**.

iKPI-ID	i2.1
Description	Influence standardization with at least 4 contributions
Lead partner(s)	EURECOM
`	3GPP D2D; 3GPP RAN controller;
Means of verification	Contributions or comments added to WI documents; leadership in WI; creation of new WI; Deliverable D6.10
Methodology / tools	Identifying the key WI and SDO, where IntellIoT contributions could be beneficial; participations to WI meetings; contributions to the meetings; presentations of IntellIoT framework to the WI meetings and stakeholders;
Baseline	N/A
Final Assessment	**iKPI achieved**. Creation of a new W3C WG, integration of two IntellIoT concepts in W3C standards, two 3GPP standards contributed. More details in D6.10.

iKPI-ID	i2.2
Description	Development and offering of at least three security assessment and certification models tailored to specific standards, validated in the context of IntellIoT's use cases
Lead partner(s)	SANL
Mapping to components & measurement points	Security Assurance Platform
Means of verification	Deliverable D4.8
Methodology / tools	The assessment and certification models will be specified through the Security Assurance Platform and documented in the relevant deliverable (D4.8). They will encompass vulnerability, dynamic, and hybrid assessments.

	All three models will be validated in each of the three use cases.
Baseline	N/A
Final Assessment / Evaluation	Successful delivery of 3 assessment models (Vulnerability & Dynamic assessment models, Cycle 1 & Hybrid in Cycle 2); see D5.5 & UC demos. **iKPI target achieved**.

iKPI-ID	i3.1
Description	At least 80% of the stakeholders involved in the use cases (without direct project involvement) express a positive view on: (a) IntellIoT facilitates deployment of semi-autonomous IoT applications; (b) IntellIoT increases trustworthiness of NG IoT infrastructures
Lead partner(s)	SANL / TSI
Mapping to components & measurement points	Trustworthiness Components
Means of verification	Inclusion & analysis of questionnaire results in D5.6.
Methodology / tools	Questionnaires circulated by the IntellIoT consortium members at end-user workshops and other events (as well as contacts from other EU projects and/or partnerships) to assess the perception of industry stakeholders and researchers associated with the field of NG-IoT applications on the related subjects of this iKPI.
Baseline	N/A
Final Assessment	Data extracted from Questionnaires (see also Appendix B for a more thorough analysis): 1. 60% of the correspondent's state that security and privacy concerns hinder the adoption of NG-IoT applications in their specific domain 2. All correspondents state that the inclusion of trustworthy IoT capabilities, as proposed by IntellIoT, would facilitate the deployment of Next Generation IoT applications in verticals domains 3. 92% of the correspondents agree or strongly agree that the Trust-related technologies provided by IntellIoT help cover the security & privacy needs of NG-IoT applications. 4. 80% of the correspondents agree or strongly agree that IntellIoT's security & trustworthiness technologies offer a specific advantage in comparison to the relevant solutions that they already employ Based on the aforementioned results, we consider that external stakeholders express an overly positive view of the advantages the IntellIoT trust-worthiness components provide. As such we consider that the **iKPI has been achieved**.

iKPI-ID	i3.2
Description	At least two external organizations (beyond Open Call partners) use (on trial-basis) the IntellIoT framework and build IoT applications on top
Lead partner(s)	TTC

Mapping to components & measurement points	Out of the three use-case domains at least one technical building block described in Section 3 is used by external organizations apart from the IntellIoT framework.
Means of verification	Description of the applications that are using the IntellIoT framework and its technological building block which are built on top of it in deliverable D6.8. An evaluation report from the external organizations about the IntellIoT framework and a brief assessment of the business case will complement the verification.
Methodology / tools	Technical description report on the IoT application and an assessment of the business case (e.g., business canvas or similar) from an end-user of the IntellIoT project results and/or technologies.
Baseline	State of the Art IoT applications at the projects starting date.
Final Assessment	During the first cycle of the project the initial version of the IntellIoT framework has been delivered to selected Open Call 1 winners to build applications or extend the framework. During the Open Call 2 process, four selected SMEs extended their real-world applications by using components of the IntellIoT framework. Finally, within Open Call #3 (from M35 to M39) 30 Micro-SMEs applied to enter a special Hackathon. Within the Hackathon Week of five days, 15 eligible SMEs that fulfilled the requirements of Micro-SMEs worked closely with IntellIoT experts to explore the applicability of the IntellIoT framework in 11 projects. As such we consider that the **iKPI has been achieved**.

iKPI-ID	i4.1
Description	At least 80% of expert stakeholders participating in exploitation workshops (externals and use case owners) are engaged in deriving and approving novel business models
Responsible partner(s)	TTC
Mapping to components & measurement points	The exploitation workshops will identify the specific technology or the overall system approach (IntellIoT framework) as component or system, depending on the product/service under investigation.
Means of verification	Exploitation Workshop organized and results summarized and reported in the deliverable D6.8. Involvement of IoT business experts in external events such as Data Natives and IoT Tech Expo as well as hackathons organized by IntellIoT.
Methodology / tools	Report on novel business models and its product/service identified by participating parties. Questionnaire prepared by the consortium and online surveys to structure the outcome of the exploitation workshops.
Baseline	N/A
Final Assessment	During the second half of the project the business models foreseen by the result owners (during the exploitation workshops held on a UC level) have been discussed and explored during several events (e.g. Data Native Conference, IoT Tech Expo) where Business Angels, Investors have been participating. A more detailed description can be found in the respective Deliverables D6.8 and D6.9. Thus, we consider that the **iKPI has been achieved.**

iKPI-ID	i5.1

Description	Mobilise at least 5 key IoT players external to the consortium, through their participation in security and privacy-related dissemination activities organised by IntellIoT
Lead partner(s)	STARTUPC
Mapping to components & measurement points	N/A
Means of verification	Ongoing inclusion of involved stakeholders in Dissemination file, final list in D6.9
Methodology / tools	Organisation of own Meetups, inclusion of articles by IoT stakeholders in Medium Magazine, engagement with external players via Social Media Channels. Involvement of security and IoT experts in external events such as Data Natives and IoT Tech Week as well as hackathons.
Baseline	N/A
Cycle 1 Assessment and Future Updates	More than 10 key IoT players have been involved in dissemination activities. A list if stakeholders will be reported in D6.9, final report on Dissemination & Ecosystem building. We consider that the **iKPI has been achieved.**

iKPI-ID	i5.2
Description	Highlight the potential of security and privacy as an enabler for NG IoT applications, through validation of associated business models (see i4.1) with all involved IoT players
Lead partner(s)	TTC / SANL
Mapping to components & measurement points	Security, Privacy & Trust components of IntellIoT used in the individual application domains of the project.
Means of verification	Inclusion & analysis of questionnaire and/or online poll results in D6.8.
Methodology / tools	Questionnaires circulated by the IntellIoT consortium members at end-user workshops and other events (as well as contacts from other EU projects and/or partnerships) to assess the perception of industry stakeholders and researchers associated with the field of NG-IoT applications on the related subjects of this iKPI.
Baseline	N/A
Final Assessment	Based on the questionnaires that have been circulated (see also Appendix B), 96% of the correspondents have stated that from a business (added value perspective), the increased security, privacy & overall trustworthiness emphasised by IntellIoT can be considered an important competitive advantage for NG IoT market offerings. As such, we consider that the **iKPI has been achieved**.

iKPI-ID	i5.3
Description	At least 80% of surveyed end-users (without direct project involvement) confirm that IntellIoT increases security and privacy protection and thereby alleviates an important adoption barrier
Lead partner(s)	SANL / TSI

Mapping to components & measurement points	Security, Privacy & Trust components of IntellIoT
Means of verification	Inclusion & analysis of questionnaire results in D5.6.
Methodology / tools	Questionnaires circulated by the IntellIoT consortium members at end-user workshops and other events (as well as contacts from other EU projects and/or partnerships) to assess the perception of industry stakeholders and researchers associated with the field of NG-IoT applications on the related subjects of this iKPI.
Baseline	N/A
Final Assessment	Data extracted from Questionnaires (see also Appendix B for a more thorough analysis): 1. 60% of the correspondents state that security and privacy concerns hinder the adoption of NG-IoT applications in their specific domain 2. All correspondents state that the inclusion of trustworthy IoT capabilities, as proposed by IntellIoT, would facilitate the deployment of Next Generation IoT applications in verticals domains 3. 92% of the correspondents agree or strongly agree that the Trust-related technologies provided by IntellIoT help cover the security & privacy needs of NG-IoT applications. 4. 80% of the correspondents agree or strongly agree that IntellIoT's security & trustworthiness technologies offer a specific advantage in comparison to the relevant solutions that they already employ Based on the aforementioned results, we consider that external stakeholders confirm that IntellIoT trustworthiness components increase security and privacy protection and alleviate an important adoption barrier. As such we consider that the **iKPI has been achieved**.

iKPI-ID	i6.1
Description	Demonstrate interaction with at least 20 different IoT stakeholders who are external to the IntellIoT consortium, in the context of the ecosystem building activities of the project
Lead partner(s)	STARTUPC
Mapping to components & measurement points	N/A
Means of verification	Ongoing inclusion of involved stakeholders in Dissemination file, final list in D6.9
Methodology / tools	Organisation of own Meetups, inclusion of articles by IoT stakeholders in Medium Magazine, engagement with external players via Social Media Channels
Baseline	N/A
Final Assessment	**Pass**. The goal of involving 20 different IoT stakeholders have been achieved. The final results are reported in D6.9, final report on Dissemination & Ecosystem building

iKPI-ID	i7.1

Description	Successful safety and security assessment of the developed IoT environment for semi-autonomous behaviour of the farming vehicle
Lead partner(s)	SANL / TSI
Mapping to components & measurement points	Security & Trust components of IntellIoT
Means of verification	Deliverable D5.6
Methodology / tools	Verification through satisfaction of security and safety -related requirements specified for UC1, and the tractor in specific.
Baseline	Security and safety threats identified in UC1 scenarios.
Final Assessment	Cycle 1 & Cycle 2 deployment security assessments successfully completed. This **fully** satisfies the iKPI in terms of security & **partially** satisfies the iKPI in terms of safety (as security is one of the pre-requisites of safety that falls within the scope of the technologies covered within IntellIoT). **iKPI can be considered partially achieved**.

iKPI-ID	i8.1
Description	Clinicians time saving of 5%, without any loss of information for patient or specialist
Lead partner(s)	PAGNI& VIDAVO
Mapping to components & measurement points	Healthcare AI models; Patients Data Repository, Patient-physician interface
Means of verification	Summary: *Approach 1 (initial):* Comparison of time dedicated by physicians to enrolled patients' cardiovascular care during the 1st phase of the study (conventional follow-up) to total time dedicated by physicians for patients' cardiovascular care *and* e-health system-related activities (accessing the system, evaluation of available data / alerts, support of patients in system use) during the 2nd phase of the study (device-facilitated follow-up). *Approach 2 (additional):* Comparison of actual total time dedicated by physicians for patients' cardiovascular care *and* e-health system-related activities during the 2nd phase of the study to estimated total time that would have been dedicated to patients' cardiovascular care during the same period without the use of the system (actual time *plus* estimated time of physical healthcare visits that would have been mandated -per physicians judgement- if system-derived medical data had not been available). Rationale for addition of a second verification approach: Differences in actual time dedicated to patients' cardiovascular care between 1st and 2nd phase of the study (each approximately 47 weeks long) could be attributable to factors such as optimization of medical care (leading to fewer unscheduled visits), disease progression (leading to more unscheduled visits), or merely timing of events such as acute coronary syndromes that would be unlikely to occur more than once in a given patient throughout the study period, rather than to the use of the e-health system per se. Detailed description of verification procedure:

	Nineteen patients with a diagnosis of heart failure, regardless of current functional class, were enrolled in the pilot study conducted in the context of UC2. During 1st phase of the study (conventional clinical follow-up), the patients were treated as per standard of care, without the aid of the e-health system (smart devices, mobile app and patient-physician interface), while during the 2nd phase the e-health system-facilitated remote follow-up took place. For each patient, the duration of participation in each phase of the study was identical, while any healthcare encounters (scheduled visits on an outpatient basis, unscheduled visits, urgent phone calls, emergency department visits, hospitalizations) for a primarily cardiovascular cause were systemically recorded in a prospective throughout participation (1st and 2nd phase). Additionally, a retrospective search for any missed visits was performed by interrogating the hospital's visit registration database at the end of the study. *Calculation of actual time dedicated by physicians at each study phase:* Time dedicated by physicians to patients' care was calculated as follows: -Outpatients' department visit for comprehensive clinical cardiovascular evaluation: 30 minutes; -Outpatients' department visit for implantable defibrillator interrogation only: 10 minutes; -Emergency department visit not leading to hospitalization: 120 minutes; -Emergency department visit leading to hospitalization: 240 minutes (regardless of eventual length of hospitalization). *Calculation of estimated additional time that would have been dedicated to patients' cardiovascular care during the 2nd study phase if the system had not been available:* As above (e.g. if it is judged that an emergency visit would be mandated if device-derived data had not been available, 120 minutes are added). Regarding time dedicated by physicians for evaluation of device-derived data and system-generated alerts, 13 minutes were accounted for each week after initiation of device-facilitated follow-up; this value was determined by directly measuring the time a physician requires -after logging into the system's web platform (patient-physician interface)- for a comprehensive evaluation of all device-derived biological parameter values and measurement-generated alerts of a one-week period. Time required for support of patients in system use during the 2nd phase was taken into account; however, time dedicated to installation of the initial mobile app and its updated versions into patients' phones and to initial education of patients on the use of the system was not taken into consideration. Lastly, visits to healthcare facilities for clearly non-cardiovascular causes -although recorded in study-specific documents- were not included in the evaluation of this KPI, because development of the e-health system was specifically targeted to aiding cardiovascular care.
Methodology / tools	Prospective evaluation with additional retrospective validation (pilot study).
Baseline	-Approach 1: Time dedicated to patients' cardiovascular care during the 1st phase of the study (conventional follow-up). -Approach 2: Estimated time that would have been dedicated to patients' cardiovascular care during the 2nd phase of the study (device-facilitated follow-up) without the use of the system.

Final assessment	**iKPI achieved, but limitations in evaluation methodology acknowledged.** Total actual time dedicated by physicians to enrolled patients' cardiovascular care during the 1st phase of the study was calculated at 1770 minutes, whereas total time spent for patients' cardiovascular care plus for system-related activities (as defined above) during the 2nd phase was estimated at 1601 minutes (patients' care: 850 minutes; time for data / alerts evaluation: 611 minutes; time for patient support on system use: 140 minutes). The difference between the 2 periods in total time dedicated for patients' care was driven by emergency department visits and hospitalizations that occurred in 3 specific patients during the 1st phase of the study, whereas no significant differences were observed in the remaining 16. Estimated total time that would have been dedicated to patients' care during the 2nd phase without the system, on the other hand, was calculated at 1750 minutes. In particular, in 4 patients it is judged that unscheduled visits to the outpatients' clinic and/or the emergency department were obviated due to the availability of device-derived medical data of sufficient quality for management decisions to be taken; palpitations with or without associated dizziness was the primary complaint in most of these cases. Thus, physician time saving by verification approach 1 is estimated at 9.5%, with the respective value being 8.5% by approach 2. Conclusively, despite the fact that methodological limitations clearly exist (small sample size, non-randomized / non-parallel-arm study, conventions in estimating clinical visit duration), this iKPI is considered as successfully assessed in this specific patient cohort.

iKPI-ID	i8.2
Description	Higher data quality and volume compared to the current context. We will reach 90% increase in data points collection in out-of-home settings compared with current situation (measured baseline).
Lead partner(s)	PAGNI & VIDAVO
Mapping to components & measurement points	Patient data repository; patient-physician interface
Means of verification	Comparison of amount of data collected by physicians during the 1st phase of the pilot study (conventional patient follow-up), to the amount of data collected during the 2nd phase (device-facilitated follow-up).Relevant data of interest for the verification of this KPI include values of biological parameters obtained from: -Blood pressure monitors (systolic/diastolic blood pressure, heart rate); -Pulse oximeters (haemoglobin oxygen saturation, heart rate); -Weight scales (body weight); -Thermometers (body temperatures). On the other hand, smartwatch-derived biological parameter data are not included in this analysis. The rationale for this lies in the fact that only 4/19 patients participating in the UC2-pilot study had been using smartwatches before enrolment, and thus a >90% increase in data quantity would have been attributable to a great degree merely to the availability of this device type from the onset of the 2nd phase, rather to the system as a

	whole. Contrary to this, all other device types are routinely used both by healthcare personnel at clinic visits and by patients in the community setting.
Methodology / tools	Prospective collection of biological parameter data throughout both study phases; data can include measurement values recorded at healthcare facility visits by healthcare professionals, as well by the patients on an out-of-office/hospital basis.
Baseline	Number of data points collected during the 1st phase of the pilot study (conventional follow-up).
Final assessment	**iKPI achieved.** During the baseline period, 240 data points were collected from enrolled patients, the majority of which were recorded by healthcare personnel during physical visits. Only 2 patients provided hand-written biological parameter data during this period, namely systolic and diastolic blood pressure values and blood pressure monitor-derived heart rate values, with their quality adjudicated as "suboptimal", based on legibility, inconsistency of availability of time of measurement, as well as lack of details on device used. During the 2nd phase of the pilot, despite the suboptimal compliance of most patients to recommendations regarding frequency of system use, a total of 11896 data points (4857% increase) were registered to the patient-physician interface, all judged to be of optimal quality (time of measurement recorded, measurement unit always present, legible and derived from calibrated devices).

iKPI-ID	i8.3
Description	Patients achieve a 20% increase in their activity level (steps) and out of home walking time
Lead partner(s)	PAGNI
Mapping to components & measurement points	Smartphone Applications for Patients' Health and Fitness
Means of verification	Retrieval of physical activity data (steps) recorder in the smartwatch manufacturer's mobile application in patients' smartphone: In each patient, average daily step count of the 1st 30-day interval since onset of smartwatch use (study month) is compared to the highest daily average value observed in subsequent study months.
Methodology / tools	Exercise and activity applications in smartphone. Messages that encourage people to achieve health and wellness goals and remind them when they do not meet minimum goals. Exercise care plan visible on mobile app start screen.
Baseline	Activity level data from the first month of device-facilitated patient follow-up.
Final assessment	**iKPI achieved; limitations in evaluation methodology acknowledged.** As of 30/11/2023, physical activity data from 16/17 patients that did eventually use the smartwatch are available. A baseline-to-peak increase in physical activity level of at least 20% was observed in 9/16 patients, with 3 more demonstrating a baseline-to-peak increase in the order of magnitude of 15-19.99%. Only 3 patients failed to achieve any increase in daily step count average. In this 16-patient cohort, mean baseline-to-peak increase in daily step count was calculated at 23.34% (-20.33 - 103.63%). It should be noted, however, that, after the study month in which the peak physical activity value was observed (typically in the 2nd-4th study month), mean daily step count declined again in almost all patients. This could suggest that the positive effect of the system as an aid to optimize physical activity level –although present- was short lived.

	On the other hand, should also be emphasized that: 3/16 patients developed activity-limiting orthopaedic or rheumatological issues during the study period, b) 3/16 patients underwent non-cardiac surgery during the 2nd study phase, c) limitation of outdoor physical activity is strongly indicated during the summer months in Greece, where very high temperatures can prevail. Taken together, the data outlined above suggest that this KPI can be considered as successfully assessed, albeit with clear methodological limitations (small sample size, non-randomized / non-parallel-arm study, evidence of inconsistent smartwatch use in a minority of patients).

iKPI-ID	i9.1
Description	External stakeholders develop five additional manufacturing IoT applications based on the IntellIoT framework, e.g., through Open Calls or hackathons
Lead partner(s)	Siemens
Mapping to components & measurement points	Overall IntellIoT framework are used as basis for application development.
Means of verification	Deliverable D5.6
Methodology / tools	Final demonstration will include showcasing external applications.
Baseline	N/A
Final Assessment	Through the Open Call 1, the new partners Trilogis and myw.ai have been on-boarded and have contributed solutions to the manufacturing use case, by integrating their own applications. Within Open Call 2 we identified another new partner - Pumacy - with a strong applicability of the IntellIoT framework in manufacturing and connected industries such as construction, energy and smart city. In the Applied Data Hackathon the IntellIoT partner HSG provided a relevant Hackathon challenge on "Object Classification in Real Environments" that hackers used to explore the applicability of IntellIoT as well. In our Open Call 3, altogether 11 Micro-SME qualified and nine of them worked on solutions solving challenges in Industrial Metaverse applying concepts of the IntellIoT framework. Four SME were selected for the Open Call 3 and developed applications using the IntellIoT framework. **iKPI target achieved**.

iKPI-ID	i10.1
Description	Delivery of open-source software components for the HyperMAS including libraries and tooling for researchers and practitioners to design, deploy, and manage IoT/edge infrastructures
Lead partner(s)	HSG

Mapping to components & measurement points	Web-based IDE for Hypermedia MAS Hypermedia MAS Infrastructure
Means of verification	This impact is reached if the software from these two components is available as open-source and together with documentation that enables outside users to utilize the software for running hypermedia-based multiagent systems in their own context. Deliverable D3.5.
Methodology / tools	Verification of access to source code and documentation of these components via a public code sharing platform such as GitLab/GitHub.
Baseline	N/A
Final Assessment	**iKPI achieved.** Our code is Open Source, available at: - Hypermedia MAS Infrastructure: https://github.com/Interactions-HSG/yggdrasil - Web-based IDE for Hypermedia MAS: https://github.com/Interactions-HSG/intelliot-hypermas-explorer

iKPI-ID	i11.1
Description	Delivery of open-source components for 5G communication and dynamic network management supporting context-based and data-driven ultra-reliable low-latency communication for the NG IoT, as defined in Obj. 2
Lead partner(s)	EURECOM
Mapping to components & measurement points	5G RAN/CN; 5G FlexRIC; 5G FlexCN;
Means of verification	Contributions available on the OSA software library (https://gitlab.eurecom.fr/oai), either integrated or as stand-alone modules.
Methodology / tools	Access to OAI (RAN, CN, Mosaic5G) code on EURECOM hosted/administrated GitLab https://gitlab.eurecom.fr/ ; CD/CI methodology to keep new contributions in line with the stable code. Deliverable D4.7.
Baseline	OAI GitLab devel branch
Final Assessment	5G OAI including O-ran agent, FlexRIC and FlexCN in baseline release; Also available in startup BubbleRAN as baseline product in 1st quarter 2024. **iKPI achieved.**

iKPI-ID	i12.1
Description	Delivery of open-source AI algorithms and IoT/edge infrastructure components
Lead partner(s)	UOULU
Mapping to components & measurement points	AI components of UC1 and UC3
Means of verification	D3.6 and dissemination

Methodology / tools	Delivery of algorithms and source codes
Baseline	N/A
Final Assessment	AI components are available in https://gitlab.eurecom.fr/intelliot while related research is published in https://github.com/ICONgroupCWC/DistributedAI **iKPI achieved.**

iKPI-ID	i13.1
Description	Delivery of more than three standalone and re-usable innovative security tools and technologies developed within IntellIoT, in a form ready-to-be-adopted in other domains (in addition to the domains covered by use cases)
Lead partner(s)	TSI / SANL / AAU
Mapping to components & measurement points	Security, Privacy & Trust components of IntellIoT
Means of verification	Security, Privacy and Trust enablers detailed within D3.8 & D4.8.
Methodology / tools	Delivery of algorithms, source code and/or binary files and deployment and usage documentation for the core security, privacy and trust enablers of IntellIoT.
Baseline	N/A
Final Assessment	Successful delivery & demonstration of four innovative and re-usable security tools (DLTs, MTDs, SAP, Trust IDS). See D5.5 & UC demos. Licensing information provided at the end of D4.8. Exploitation details provided in D6.7. **iKPI target achieved.**

8. CONCLUSIONS

The current deliverable provides the final validation and evaluation report of the IntellIoT project. As such, it includes the validation and evaluation of the final versions of the components that have been developed during the project as well as the integrated IntellIoT framework that represents the final outcome of the project.

The validation and evaluation process is based on the activities carried out within the technical workpackages (WP3, WP4 and Tasks 5.1 and 5.2 of WP5). In these WPs, the different components have been developed and tested, while their integration, deployment and demonstration in the context of the three use cases has been carried out. Using this information, in this deliverable the successful satisfaction and validation of the final requirements set in deliverables D2.5 and D2.6 has been assessed. On top of that, using the results of the validation process of said requirements, as well as the outcomes of all other relevant activities (such as dissemination efforts, standardizations processes and other), the assessment of achievements of KPIs and impact KPIs described in the DoA has also been carried out.

In summary, the results of the validation report presented in Sections 3, 4, 5 and 6 concerning the functional, non-functional, technical and use case specific requirements, demonstrate that:

- **137 requirements out of 154 in total (89%) have been successfully validated.**
- **16 requirements have been partially validated**.
- **1 requirement has not been met.**

It should be noted that the partially validated requirements (as well as the single one that has not been met) are not considered crucial and in most cases, they are associated with a specific aspect of a feature that has not been thoroughly tested in the use case deployment (while lab tests have been successfully completed). Therefore, these partially accomplished requirements do not lead to an overall failure of a critical component or integrated entity and do not manifest as a show -stopper to achieve KPIs.

Indeed, this is demonstrated by the assessment of the KPI targets. Out of the 24 KPIs set at the beginning of the project to measure its success, 21 have been fully achieved and only three are partially accomplished (while none is evaluated as a failure to reach). This has further led to almost complete success (18/19) to reach the impact KPI targets.

APPENDIX A: HEALTHCARE USE CASE – DATA ANALYSIS AND MODEL DEVELOPMENT

Based on available data and clinical or operational relevance, the cardiologist and the data scientists in the team agreed to focus on the prediction of heart rate at rest (hrar). The prediction target can be the hrar in two days from now, or the average hrar in the upcoming week. The average heart rate from the smartwatch measurements is used as the hrar value.

Section 1 describes alternative prediction targets, and reasons why they were not pursued.

In Section 2, the results of data analysis are described. It mentions which features are measured, and how often they occur, both individually and in pairs. Graphs related to prediction target, the heart rate at rest, show clear discrepancies between patients. The discrepancy is reduced for the difference between the heart rate at rest of a patient at day x+2 and day x. Subsequently, it is shown by ways of scatter plots that there is a clear relationship between heart rate at rest now and in near future. A relationship between heart rate at rest and number of steps is unclear.

Section 3 reports on tests with several prediction models, using the complete data set, without any federated learning (FL) technique. The rationale for doing so is that FL models can only perform well if similar models without the restrictions imposed by FL perform well. The models and their performance are described in Section 0.

Unfortunately, none of the models significantly outperformed the naïve predictor, which sets the prediction target equal to the most recent hrar.

Obstacles for successful prediction of hrar include:

Data issues

For each patient, at most one feature report is available per day. In particular, although a smart watch measures heart rate frequently, the data that is available for analysis contains at most three heart rate related features from smart watch measurements per patient per day (viz. average heart rate, minimum heart rate and maximum heart rate). The number of feature reports available for analysis is limited to only 4137. Moreover, the number of feature reports containing the average heart rate is only 2404. From the 19 patients, four patients do not have any heart rate data from the smart watch; two patients have 10 and 50 reports containing such data, respectively. The 4+2=6 patients mentioned above therefore are excluded from analysis.

Moreover, data availability need not be consistent over time: patients can have series of consecutive days without any data report, and some features occur in more reports than others.

Inherent unpredictability of heart rate at rest

In the IntellIoT consortium October 30 and 31, 2023, external advisor Prof. Ch. Thümller commented that in his medical experience, predicting heart rate at rest over a long period, e.g., one in day in advance, is notoriously difficult. Moreover, relevant data such as per-patient medication or indoor or outdoor temperature are unavailable.

This comment was supported by our data analysis. In a scatter plot, a relation between the heart rate at rest at day x and the heart rate at rest at day x+2 can be observed, but scatter plots of the number of steps taken versus the heart rate at rest at day x+2 does not show such a relation. This makes it unlikely that number of steps is a very relevant feature for predicting the heart rate at rest.

Discrepancy between heart rate at rest of different patients

The heart rate at rest of different patients may have widely different distributions. For example, the median heart rate at rest of some patients in the population is about 53 beats per minute (bpm), while it is about 75 bpm for another patient. The discrepancy between the distributions of the heart rate at rest of different patients makes it more challenging to build a model for predicting the heart rate at rest for *all* patients. The effect of the discrepancy can be mitigated by not directly predicting the (average of) the heart rate at rest, but rather the *difference* between the (average of) the heart rate at rest in future and the current heart rate at rest. The reason is that the distributions of the difference are much more similar for different patients than the distributions of the heart rate at rest themselves. Our experiments show that this mitigation does have an advantageous effect, but insufficiently so to obtain an appreciable advantage over the naïve predictor.

Selection of the prediction target

We start with an edited version of the document discussing the prediction target prepared by the cardiologist.

Introduction / Definitions:

MET

At exercise, the most objective way to assess performance (work performed) in everyday practice is direct measurement of Maximal Oxygen Consumption (VO_2peak) by performing cardiopulmonary exercise testing (CPET). However, work performed at exercise can also be expressed in Metabolic Equivalents (METs), *and, in fact, this is by far the most common way by which physicians grade estimated exercise intensity in everyday practice.*

1 MET represents *oxygen consumption at rest*, and it equals to *3.5 ml / min / kg.* For example, if an individual achieves a VO_2 peak value of 17.5 ml / min / kg at an exercise test, this has been a 5MET-exercise (17.5/3.5), meaning that oxygen consumption was increased fivefold during exercise compared to rest.

The estimated METs achieved at exercise are also automatically provided by the software at every conventional exercise stress test (without using the CPET mask), but on CPET measurement is direct.

HRI

Heart Rate Index (HRI) = Maximal Heart Rate achieved at a given exercise session (HRmax) / Resting Heart Rate (HRrest)

HRI=HRmax/HRrest

MET VS HRI

The key point of interest in this easily calculated index is that it has been shown to correlate well to METs achieved during exercise; more precisely, the METs achieved at a given exercise session can be estimated satisfactorily by the formula:

METs = 6 x HRI - 5

Example: A person with a resting heart rate of 60bpm who achieves a maximal HR of 120bpm at an exercise session, is estimated to have achieved a work equivalent to:

6 x 120/60 – 5 = 7 METs

In the context of Use case 2 of the IntellIoT project, the importance of this index (HRI) lies in the fact that it is easy to calculate only from Heart rate data for every single exercise session the subjects will be performing, giving a reasonably accurate estimate of exercise performance.

BMI

Body Mass Index (BMI):

BMI equals to Body Weight / $Height^2$

Proposals for predictions:

Table 3. Proposals for predictions

	Input (parameters based on which predictions are to be based)	**Output (parameters to be predicted)**	**Comment**
1	-Physical activity (steps) during the last 7 days -Body weight / BMI prior to exercise (last available value) -Pre-exercise SaO2 (oxygen saturation) -HRrest -HRI of the previous exercises	-HRI (Heart Rate Index), or estimated METs at exercise Endpoint for prediction here is HRI.	This target was not selected due to insufficient data. It is hard for patients to comply with the required workflow, there were connectivity/communication issues (multiple platforms)
2	-Physical activity (steps) during the last 7 days -Body weight / BMI prior to exercise (last available value) -Pre-exercise SaO2 (oxygen saturation) -HRrest -Successful previous exercise session (Yes/No)	-Success of next exercise session (pre-specified criteria)	Criteria not specified. Unclear if SaO2 measurement we obtain is always performed at the correct time.
3	-Average number of steps from day (X-7) until day X Reason for selecting weekly average over weekly sum of steps: If average step count is the input, then potential gaps in data (steps) will not impede the sliding window. In addition, weekly step sum can easily be calculated from average step count. -Last available BW (kg) measurement -Last available SaO2	-HRrest on day(X+2) T -HRminimum on day (X+2) -Average of HRrest / HRminimum / HRmaximum on week (X+1 to x+8)	Only considered HRrest on day X+2 and average of HR rest on weeks X+1 to X+8. Why not HRmin and HRmax? HR min under investigation now. No fundamentally new results for day X+2.
4	- % Weekly compliance with system use (measurements performed / measurements recommended)	-Successful conduct of exercise regimen at weekly level	Discarded: only one measurement per patient per week, too few datapoints

5	- % Weekly compliance with system use (measurements performed / measurements recommended) - Successful conduct of exercise regimen at weekly level	*-Score at Quality-of-Life Questionnaire (to be confirmed)*	Discarded: Quality-of-Life Questionnaire not available
6	Heart rate minimum average, week (X-7 to X): Absolute value, but also any Δ from last weekly sliding window. Heart rate max (same) Blood pressure (systolic): Same Body weight: Same Oxygen saturation (rest): Same Weekly steps: Same	Any clinically relevant unscheduled patient-physician interaction (comment field completed by physician): Phone calls for symptoms, phone calls for worrying measurements, non-predefined medication modification, unscheduled visits, measurement-based automatic alarm generation, hospitalizations	Discarded: few events, expectation was that it was unlikely to obtain statistically relevant results.

Data analysis

DATA DESCRIPTION

Analysis was performed on the data available on November 13, 2023. It consists of 4137 reports. Each report contains measurement data from one day from one patient. Each patient has at most one report per day. Table 4 shows the number of reports in which each particular feature occurs. Some features are obtained from the same measurement and therefore are simultaneously absent or present in a report. These features are shown in the same row of Table 4.

Table 4. Availability of features

Feature	Count	Measuring device
hr_average, spo2_average	2645	Oximetry meter
hr_min, hr_max, spo2_min, spo2_max	2613	Oximetry meter
heart_rate, systolic_pressure, diastolic_pressure	2290	Blood pressure meter
maximumheartr, minimumheartr	2414	Smartwatch
averagehr	2404	Smartwatch
weight	1586	Weighing scale
steps	3088	Smart watch

Please note that although a smart watch monitors heart rate at a fine time scale, per patient and day at most three heart rate related features measured by a smartwatch are available for analysis. These features are daily aggregates. It was decided to use heart rate information from the smartwatch only, because according to the clinician's expert opinion, the values from this data source most accurately reflected the *actual* heart rate at rest.

As shown in Table 5, the amount of available data varies over patients. The column 'total' refers to the number of reports for a patient. The column 'both' denotes the number of reports containing both averagehr and steps.

Table 5. Availability of features per patient

Patient	total	averagehr	steps	both
patient_6	310	258	259	257
patient_19	321	240	244	238
patient_17	231	217	215	210
patient_5	229	195	194	193
patient_15	299	221	185	183
patient_3	241	170	170	170
patient_14	271	171	267	169
patient_20	246	180	190	161
patient_8	261	149	238	149
patient_9	242	156	160	145
patient_4	253	169	160	115
patient_16	256	112	197	112
patient_11	272	106	196	105
patient_10	156	50	142	49
patient_18	86	10	79	7
patient_1	138	0	0	0
patient_2	284	0	192	0
patient_12	37	0	0	0
patient_7	4	0	0	0

Patients 1,2,7 and 12 do not have any data from the smartwatch related to heart rate, and patients 10 and 18 just have 50 and 10 such data points. These six patients can thus be removed from the analysis. According to the cardiologist's judgment, the medical condition of some patients makes it infeasible to predict their heart rate at rest. As a result, the analysis only includes the date of nine patients (viz. patients 3,4,5,6,9,15,17,19 and 20).

The reports also contain a 'comments' field, which is hardly ever used, *viz.* 29 times. Its usage varies over patients, e.g., it is used as much as seven times with patient 14. Five times, the field recorded when an exercise started and ended, and two times, it recorded an unscheduled event.

HEART RATE AT REST DISTRIBUTIONS FOR DIFFERENT PATIENTS.

Figure 2 shows the heart rate at rest over time for three patients. Apart from the variation over time per patient,

Figure 2 also shows a marked difference between patients. For example, Patient 3 has a much higher heart rate at rest than patient 20.

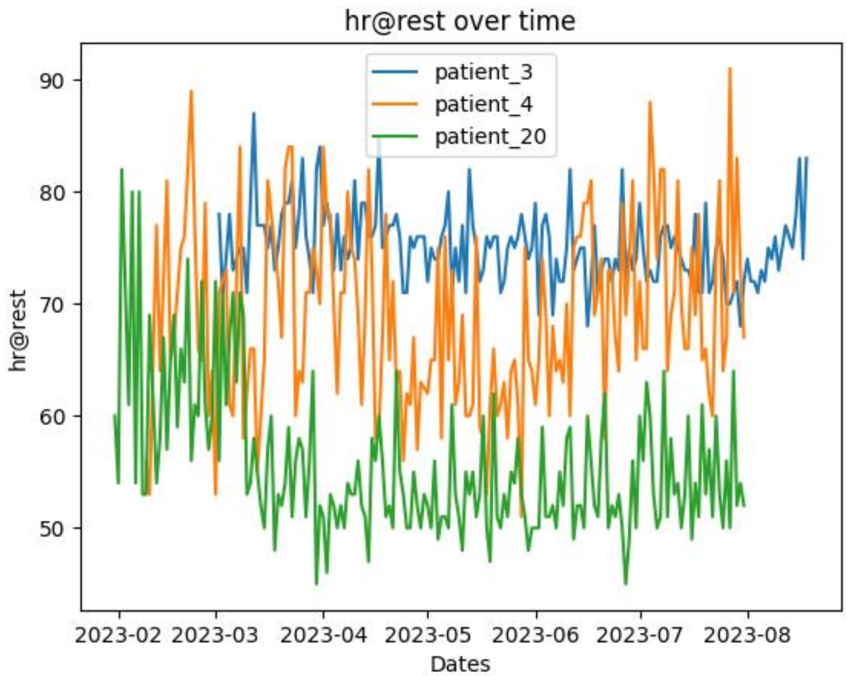


Figure 2: heart rate at rest over time

The difference of the heart rate at rest over patients is further illustrated in the boxplots of the distribution of the heart rate at rest in Figure 3. As usual, the upper and lower lines of the boxes correspond to the 75% and 25% quartile, respectively. The middle line in each box corresponds to the median. The numbers below the patient labels indicate the number of measurements. For example, there were 221 measurements for heart rate at rest (with the smartwatch) for patient 15.

In order to have a target with a distribution that is more similar amongst patients, we determined for each patient the difference of heart rate at rest at day x+2 and heart rate at rest at day x (whenever both measurements for heart rate at rest are available). In this way, a patient-dependent offset of heart rate at rest is mitigated. Boxplots for the distribution of the difference are shown in Figure 4.

For all patients (except patient 17), the median difference is approximately 0. The distributions of the difference are much more similar for different patients than for the heart rate at rest. As a result, predicting the difference is better suited for federated learning, which has as basic assumption that datasets are independently and identically distributed.

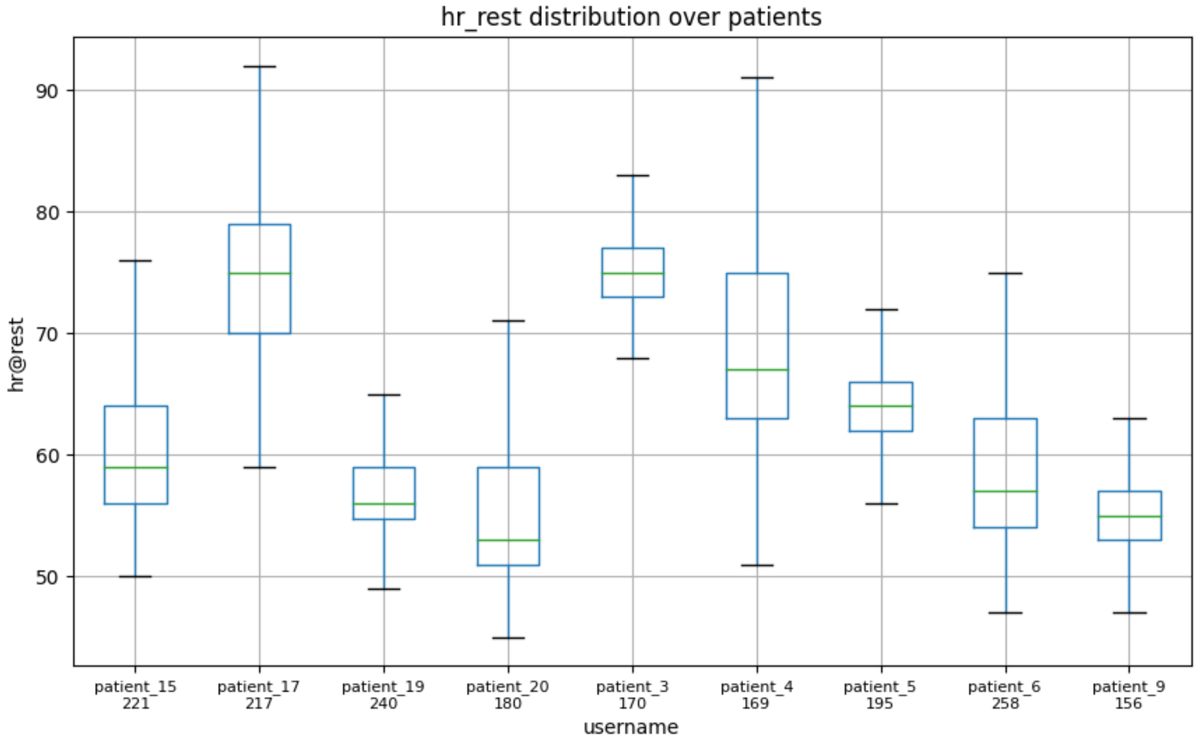


Figure 3: Heart rate at rest distribution over patients

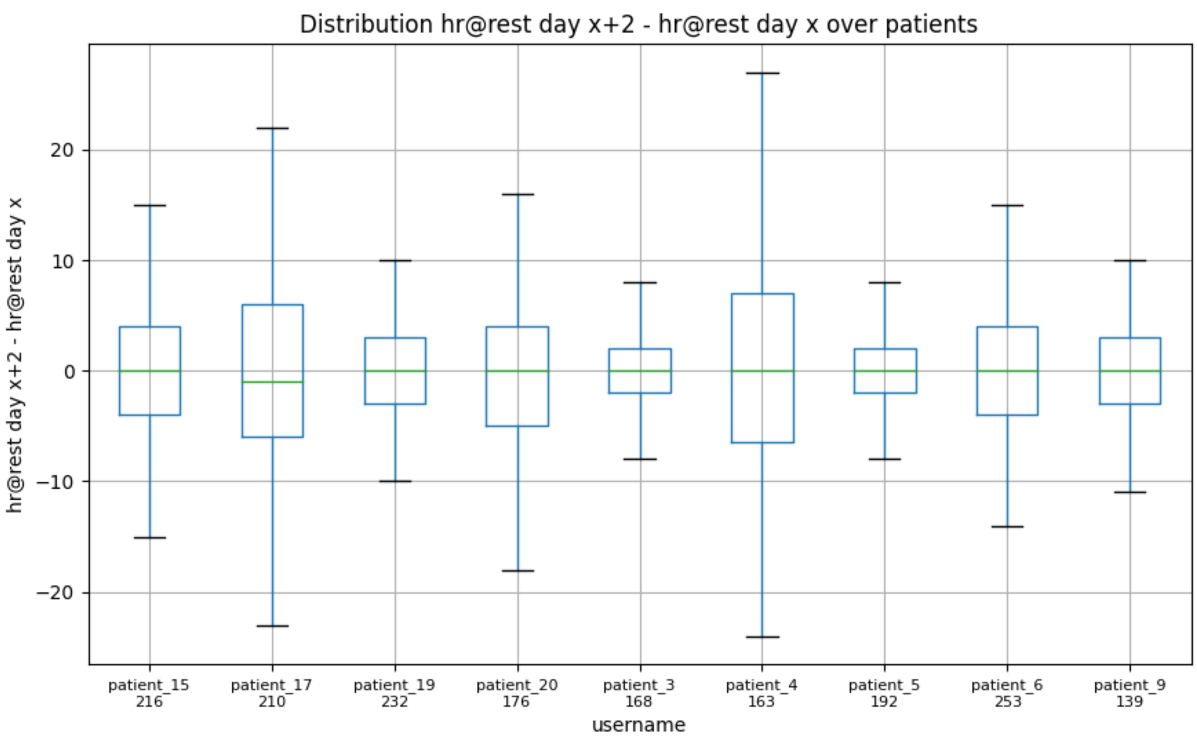


Figure 4: Distribution of hr@rest day x+2 – hr@rest day x

RELATIONSHIP BETWEEN FEATURES AND HEART RATE AT REST

This section shows results from the analysis of relations between various features and the heart rate at rest from smartwatch data. It has three subsections, dealing with three different prediction targets on day x: the heart rate at rest at day x+2, the average heart rate at rest in days x+1, x+2, ..., x+7, and the "residual" average heart rate with the upcoming week, that is, the difference between the average heart rate at rest in days x+1,..,x+7 and the heart rate at rest at day x. In each subsection, the features are: heart rate at rest at day x, average heart rate at rest in days x-6, x-5, ... , x, number of steps at day x, and average number of steps in days x-6, x-5, .., x.

The results are shown in scatterplots, where different patients are indicated by different colors.

Relations between heart rate measured by oximetry meter or blood pressure and other features were investigated at well, e.g. the relation between heart rate at rest measured by the smartwatch and measured by the oximeter. For conciseness, the results of this analysis are not shown in this report.

PREDICTION TARGET: HEART RATE AT REST IN TWO DAYS

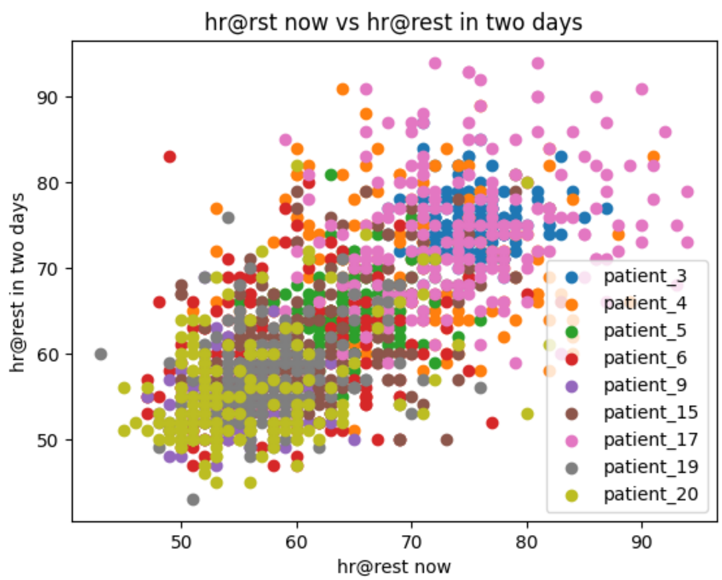


Figure 5: Heart rate now versus heart rate in two days

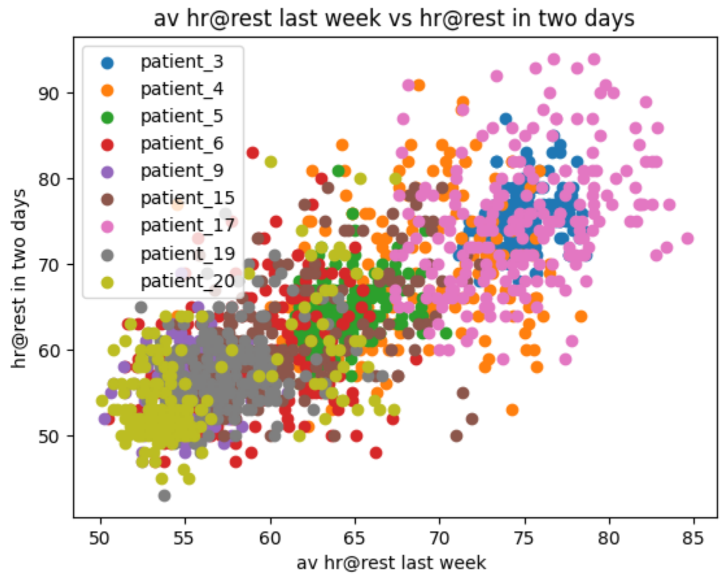


Figure 6: Average heart rate last week versus heart rate in two days

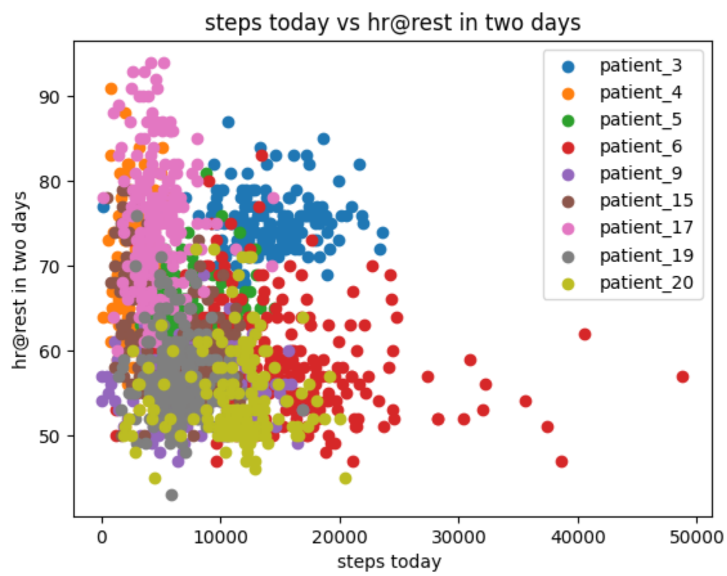


Figure 7: Number of steps versus heart rate in two days

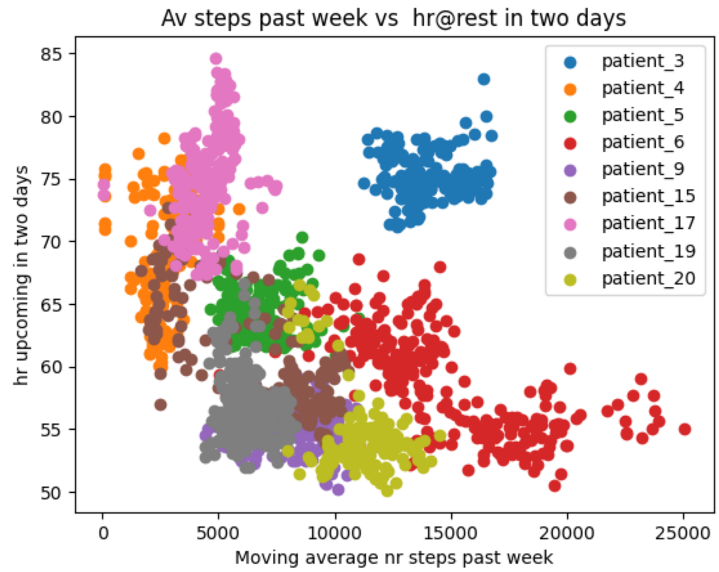


Figure 8: Average number of steps last week vs heart rate in two days

Conclusion: Figure 5 the datapoints cluster around the diagonal where the heart rate at rest in two days and the heart rate at rest are equal. The same is true for the data points Figure 6. This shows that predicting the heart rate to stay the same on average is not unreasonable. Figure 7 does not show a clear relationship between the heart rate at rest in two days and today's number of steps. In Figure 8, we see that for some patients (e.g. patients 6 and 20), a larger number of steps tends to correspond to a lower heart rate at rest. However, this is not true for all patients, e.g., with patients 3 and 17. It is therefore questionable if the number of steps is a useful feature for predicting the heart rate at rest in two days.

PREDICTION TARGET: AVERAGE HEART RATE AT REST IN UPCOMING WEEK.

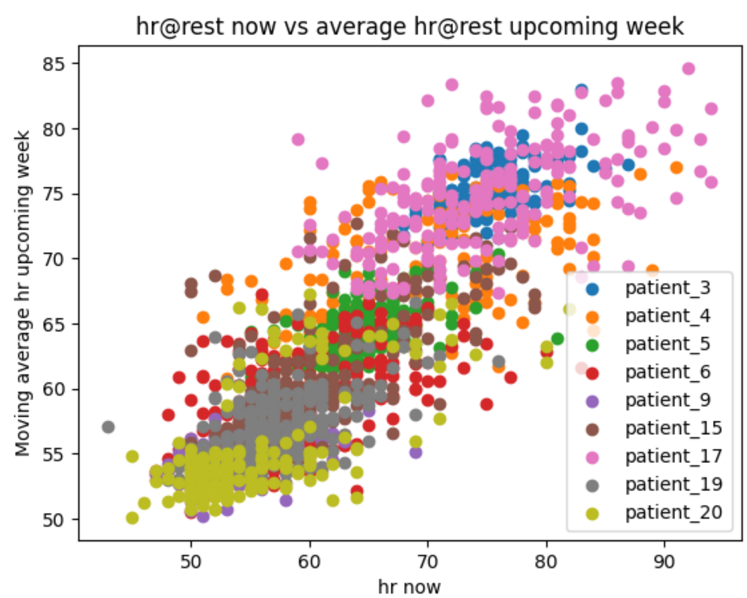


Figure 9: Current heart rate at rest versus average heart rate at rest next week

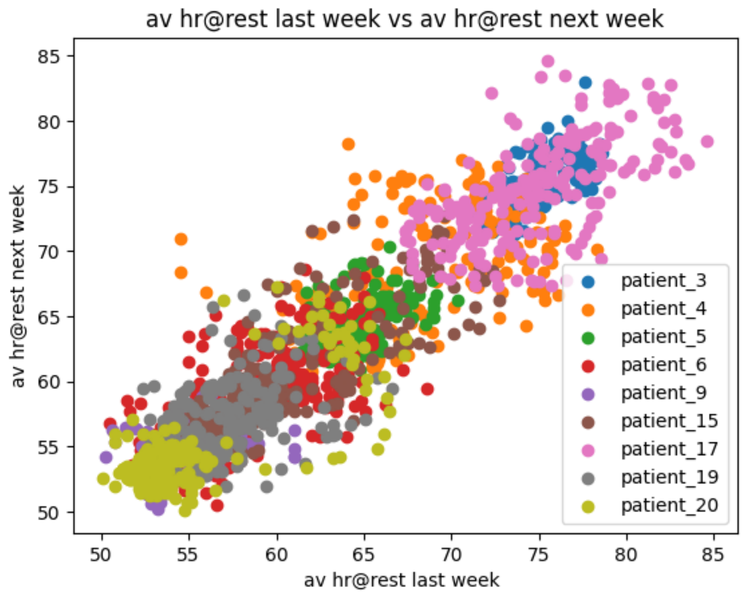


Figure 10: Average heart rate at rest last week versus average heart rate at rest next week

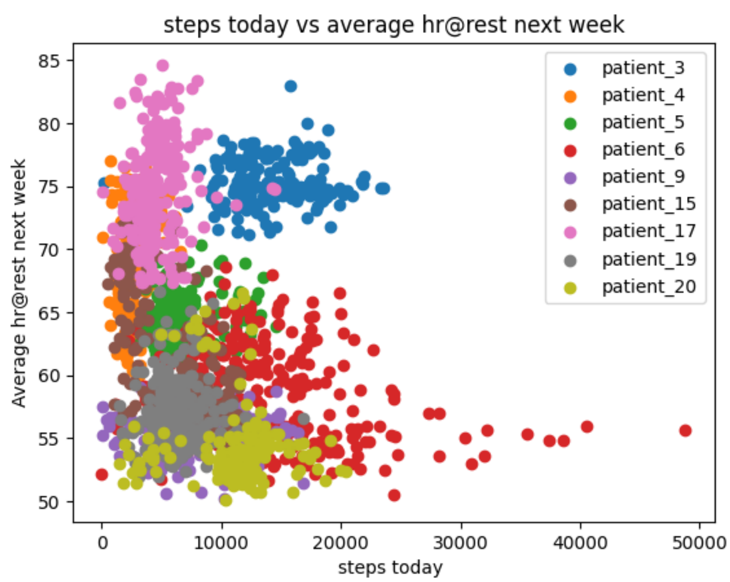


Figure 11: Number of steps versus average heart rate at rest next week

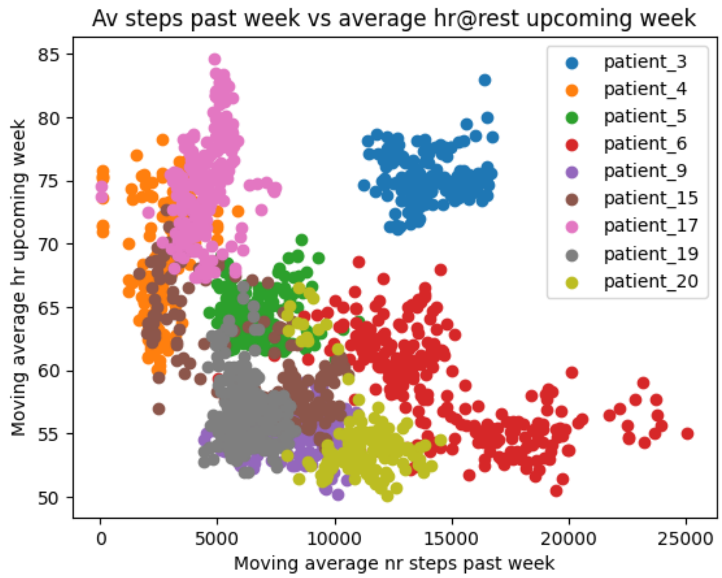


Figure 12: Average number of steps last week versus average heart rate at rest next week

The conclusion is the same as for predicting the heart rate at rest in two days: a clear relationship of the average heart rate at rest in the upcoming with heart rate at rest from today, or its average from the last week; no clear relationship between number of steps from today, or its average from the last week.

PREDICTION TARGET: RESIDUAL HEART RATE AT REST

In this subsection, the prediction target under consideration is the residual heart rate at rest. That is, at day x, the prediction target is the difference between the average heart rate at rest at days x+1,x+2,..,x+7 and the heart rate at rest at day x.

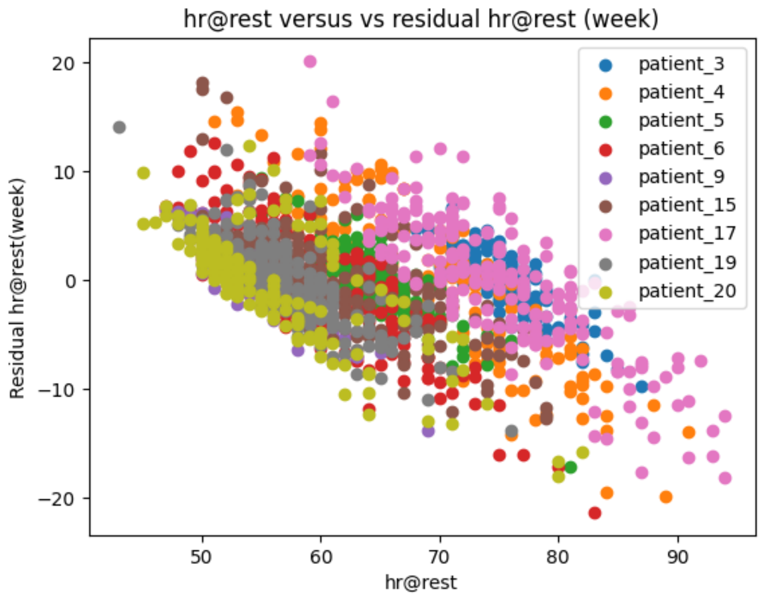


Figure 13: Current heart rate versus residual (week)

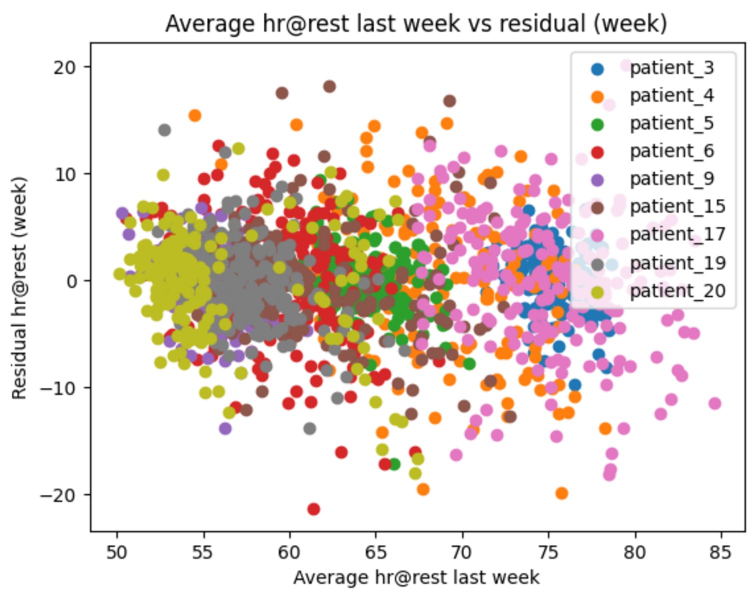


Figure 14: Average heart rate last week versus residual heart rate (week)

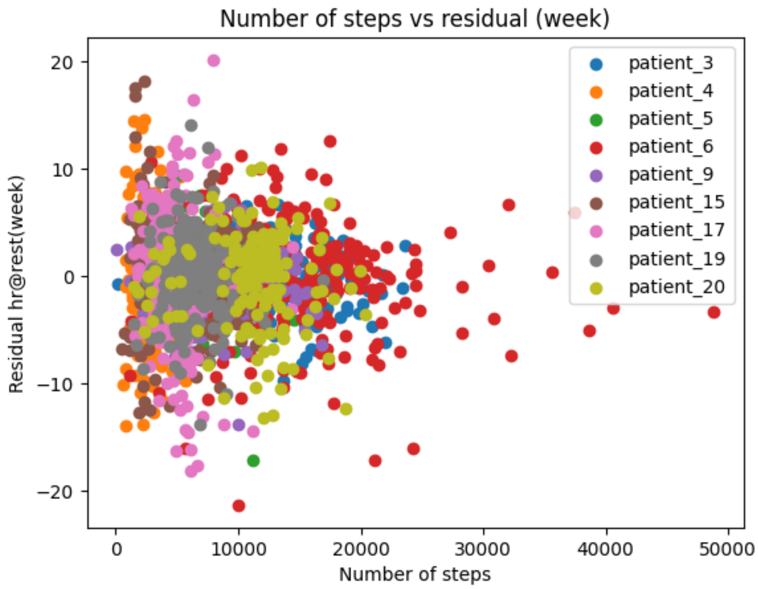


Figure 15: Number of steps today versus residual (week)

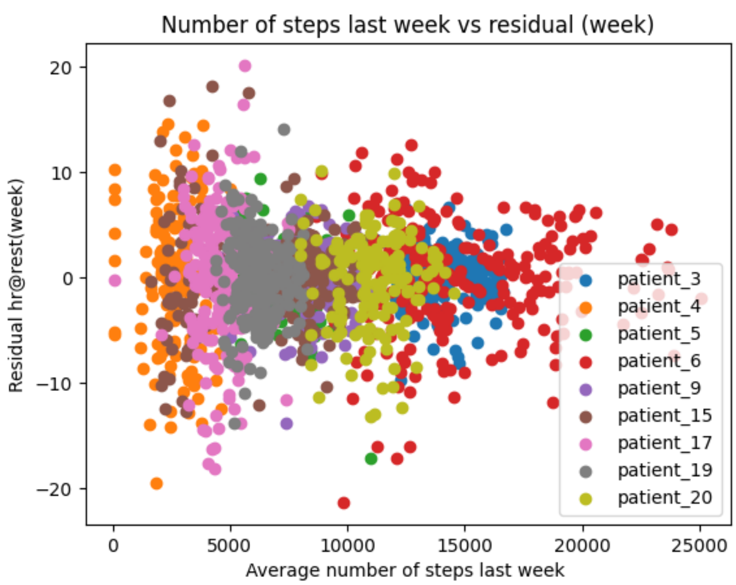


Figure 16: Number of steps last week versus residual (week)

Conclusion:

Figure 13 shows that the residual heart rate tends to decrease if the current heart rate at rest decreases. This is not illogical: if the current heart rate at rest is already high, it is expected that it will not increase further. This tendency is much less prominent in Figure 14, presumably because the average heart rate at rest last week and the residual involve days that can be up to 14 days apart.

The (average) number of steps does not have a clearly visible relationship with the residual heart rate at rest.

Prediction models and simulations

OVERVIEW TRAINING MODELS

We decided to compare four models, each increasing in complexity (and thus in theory could figure out more complex patterns):

Naïve predictor: this predictor will simply predict the last known measured value (e.g., the predicted heartrate of day x+2 is the currently measured heartrate of day x).

Average naïve predictor: This predictor will instead predict the average resting heart rate of the last 2 weeks of heart rate data.

Simple linear model: a linear 1-layer neural network

Recurrent Neural Network / GRU: More sophisticated neural network taking previous measurements into account by updating an internal memory state.

We can use the naïve predictor as a benchmark for the more complicated prediction models. Note that both the naïve models do **not** have any learnable parameters and do not require any training. We validate the model performance by computing the Mean Squared Error (MSE) loss.

DATA OVERVIEW

As described above, not all patients have consistent enough data to properly train a model on. Hence, we have decided to only train (and validate) on patients: 3, 5, 6, 8, 9, 15, 16, 19 & 20.

As for the label to predict, we compute the average of the heart rate values (measured by the smartwatch) over the upcoming 7 days. Obviously, we will exclude data with NaNs while calculating this average (e.g. if 6 of 7 days have data, we will compute the average by adding and dividing over those 6 days that do have data).

For the pre-processing of the data for model training, we windowed the time-series data into windows of 14 days. To fill the training dataset, we selected the first window of 14 days with good data coverage, this is the first entry into the dataset. Next, we shift the window by one day and is the next entry into the dataset. We keep shifting the window by one day until we reach the end of the good data. In other words, we use a shifting window of length 14 days with step-size of 1.

To combat missing data measurements (NaN) in the input, we simply put a 0 in all the entries without a measurement. Moreover, we remove all windows with any NaN labels, as the model will not have any ground truth to train on.

VERIFYING MODEL TRAINING

To verify that the model can train on the data, we will overfit heavily on the data of a single patient. The model should be able to 'learn' the trend of a single patient and follow the predictions near perfectly. This will show that weights of the model are optimized correctly given the training framework. We will validate this for the models with trainable parameters, i.e. the linear and GRU.

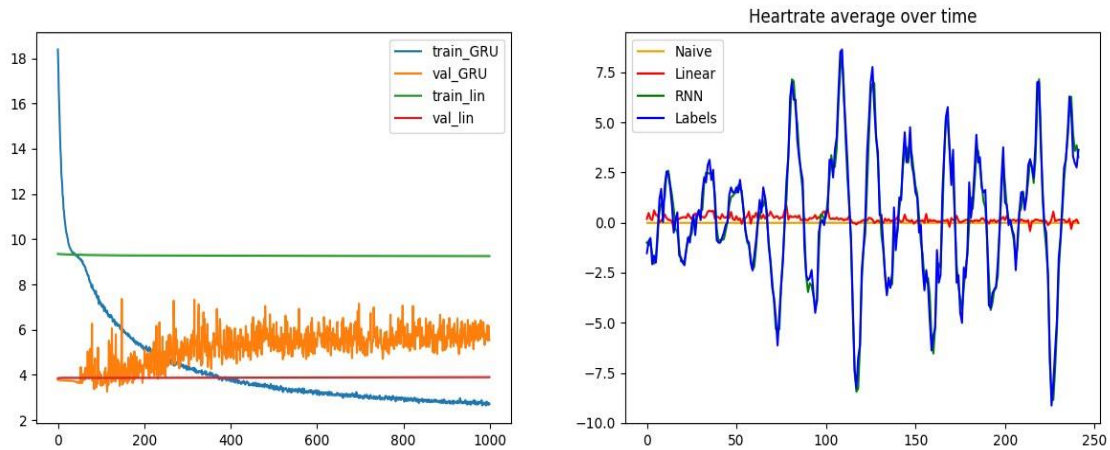


Figure 17 The loss function (left) and prediction plot (right) of models trained to overfit on a particular patient.

Figure 17 shows the training results (left) and prediction results (right). The labels here are the calculated residual losses (described above in the data section). Note that the GRU training loss keeps converging and nicely follows the labels on the right side, but also note that the linear loss is stagnant and appears to not be learning a lot on the residual losses.

TRAINING – AVERAGE HEART RATE AT REST NEXT 7 DAYS

For the first model, we select a single patient as the validation patient to monitor overfitting and the ability for the model to generalize on unseen data. We then train on the data of the remaining patients. When completed, we compute the MSE loss of the final model on the validation patient, reset the model weights, select a new validation patient, and repeat the training procedure. This results in the following MSE loss values:

Table 6. Validation MSE values for 4 different models. Note that the average naïve model outperforms all other models

Validation #	Naïve	AVG Naïve	GRU	Linear
Patient_3	8.7086	**2.7027**	127.9528	157.0122
Patient_5	10.1367	**3.5137**	7.7268	19.1228
Patient_6	28.6260	**8.0901**	21.2730	46.7929
Patient_8	50.8722	**27.7464**	38.7570	200.5317
Patient_9	10.7230	**3.7559**	10.3197	25.3531
Patient_15	21.4486	**6.6591**	10.7011	22.6176
Patient_16	6.8466	**5.8985**	38.8419	41.0851
Patient_19	13.7987	**7.4027**	11.0098	19.5955
Patient_20	16.9641	**7.5804**	10.1056	96.9589

From this table, we can observe that the GRU slightly outperforms the Naïve model. However, the AVG Naïve model performs by far the best on the validation data. We can visualize some patients to get insight into the training procedure and predictions.

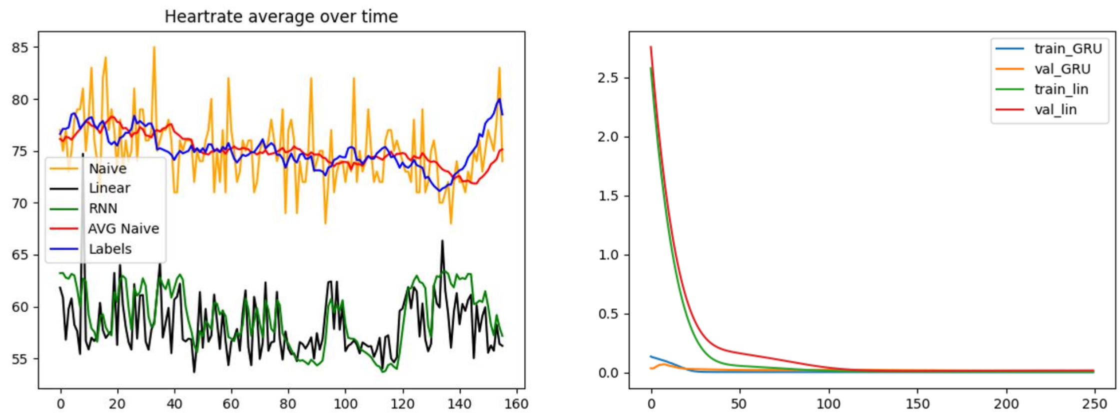


Figure 18 Patient 3 predictions (left) and training results (right)

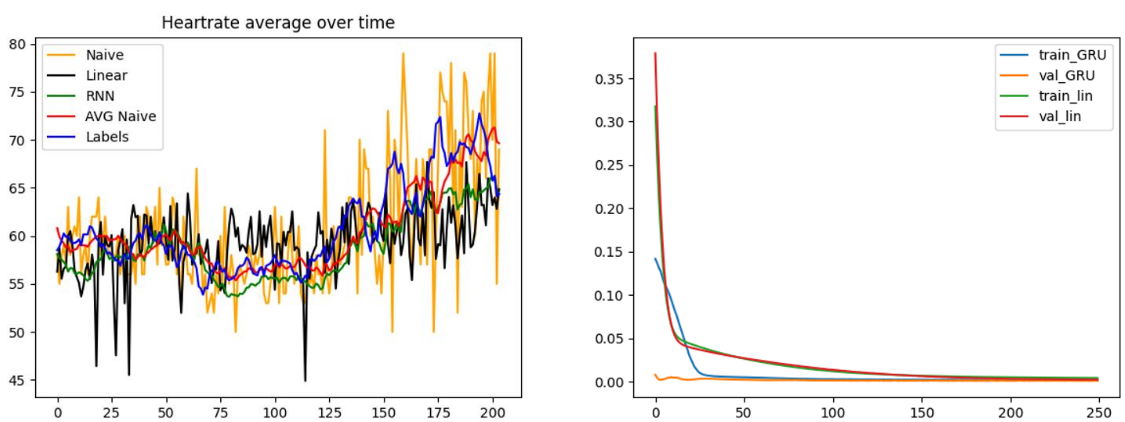


Figure 19 Patient 15 predictions (left) and training results (right)

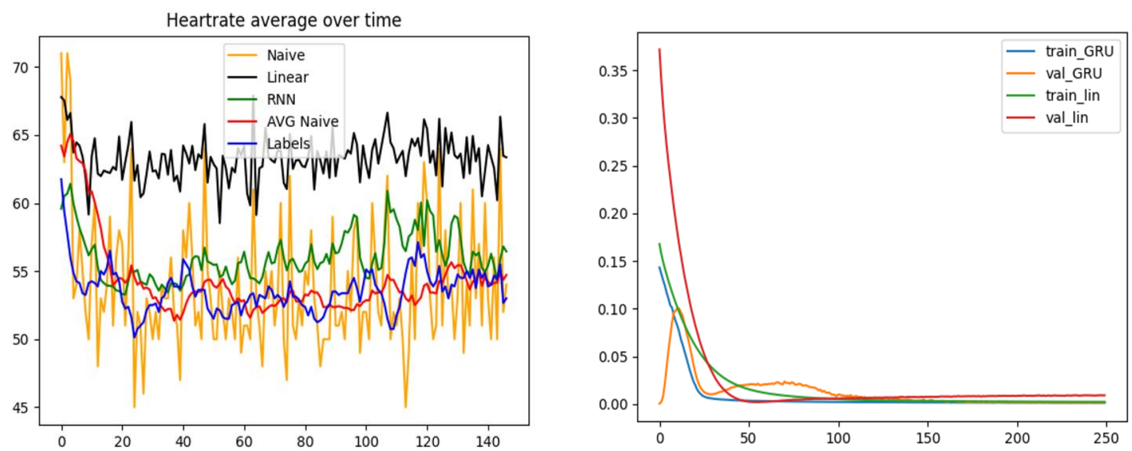


Figure 20 Patient 20 predictions (left) and training results (right)

Notice that in all cases, both the linear and GRU models converge nicely to low MSE loss values. Overall, the models can predict the rough trend of the heart rate data, however they cannot track the peaks of this trend. Notice in Figure 18 that the linear and GRU models consistently under-predict the heart rate values. This might be because the training dataset does not contain patients with a resting heart rate this high.

TRAINING – RESIDUAL AVERAGE HEART RATE NEXT 7 DAYS

Next, we trained a model on the residual average heart rate data. The residual was calculated by taking the average heart rate of the last 7 days and subtracting it from the average heart rate of the upcoming 7 days. Note that as we compute the residual now, the naïve predictor now simply predicts 0. We ran the same training experiment and below you can find the MSE loss on the validation patients. Notice that now depending on the patient, the GRU or Linear models perform slightly better (although all losses are incredibly close to each other). Below you will also find three figures of the model predictions and training results. Notice that again for patient 3, like the model trained in the previous section, the RNN has trouble converging to the correct values (Figure 21). Figure 22 shows that the RNN model is trying to predict some of the peaks and valleys (especially around the 150+ indices). Figure 23 shows that for some validation patients, the RNN still is hovering around predicting 0 residual and thus has trouble learning the trend for this patient.

Table 7 Validation MSE values for 3 different models. Note that the average naïve model outperforms all other models

Validation #	Naïve	GRU	Linear
Patient_3	**2.3642**	29.5286	2.6723
Patient_5	**3.8591**	6.5298	3.9208
Patient_6	10.5805	12.8735	**10.5659**
Patient_8	32.4681	**30.1358**	32.6384
Patient_9	**3.8373**	6.9956	3.8542
Patient_15	**7.7070**	8.7100	7.7285

Patient_16	5.8367	**5.6032**	5.7110
Patient_19	7.7202	**7.6460**	7.7602
Patient_20	**7.5656**	7.7148	7.7581

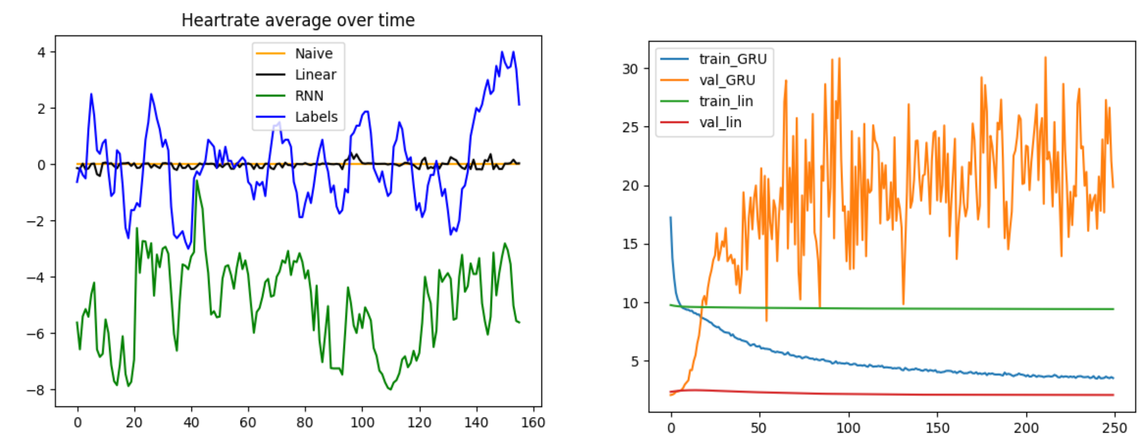


Figure 21 Patient 3 prediction (left) and training results (right)

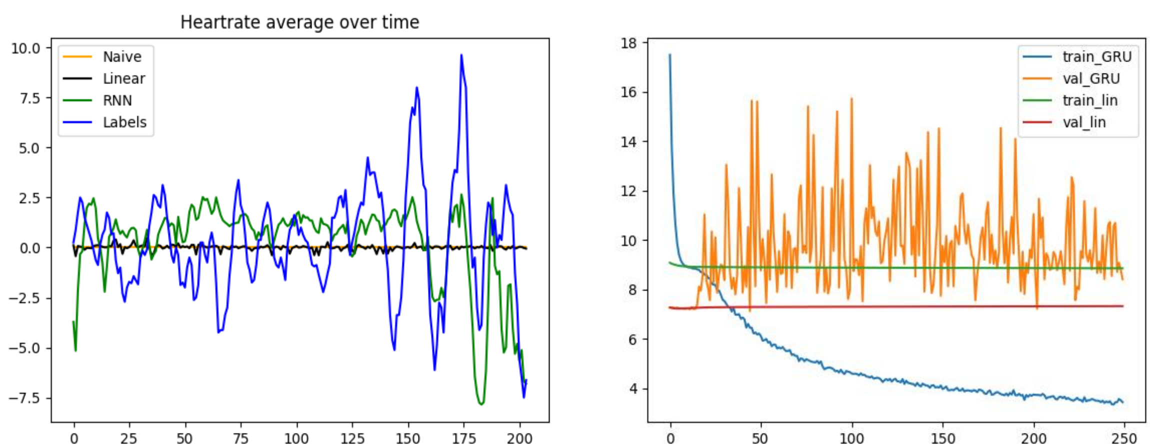


Figure 22 Patient 15 prediction (left) and training results (right)

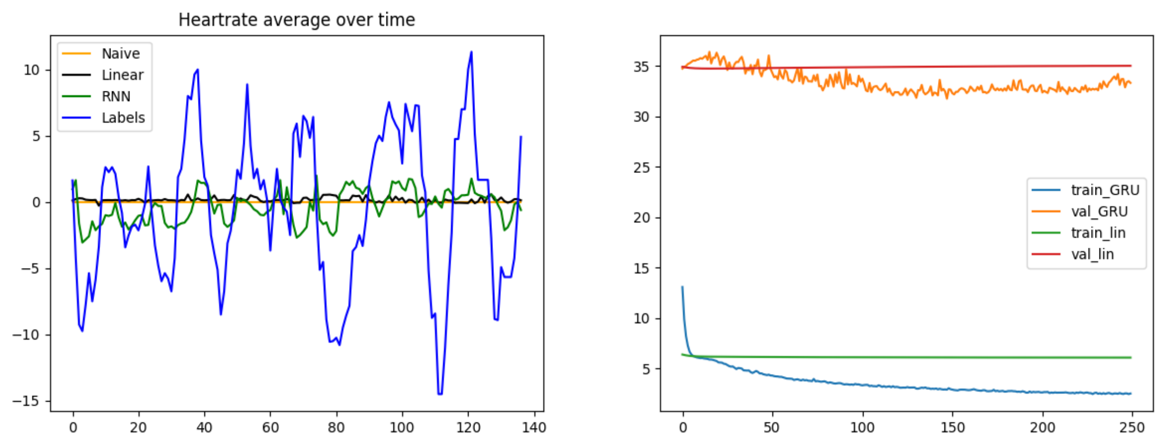


Figure 23 Patient 8 prediction (left) and training results (right)

TRAINING – VALIDATION PARTLY OF ALL PATIENTS

Finally, we repeat the training from the previous section, but now select the validation dataset as the first month of data of each patient. The idea behind this is that during the pilot study more data will be accumulated, and when a new training round is started, the first month of data could be used to continuously validate the local model updates. Note that now we only have to train a single model as the validation data will be the same. Below you can find the MSE loss on the validation data. Note that again the Naïve model is performing the best, although again all models are very close to each other.

Validation #	Naïve	GRU	Linear
-	**1.9492**	2.0494	1.9609

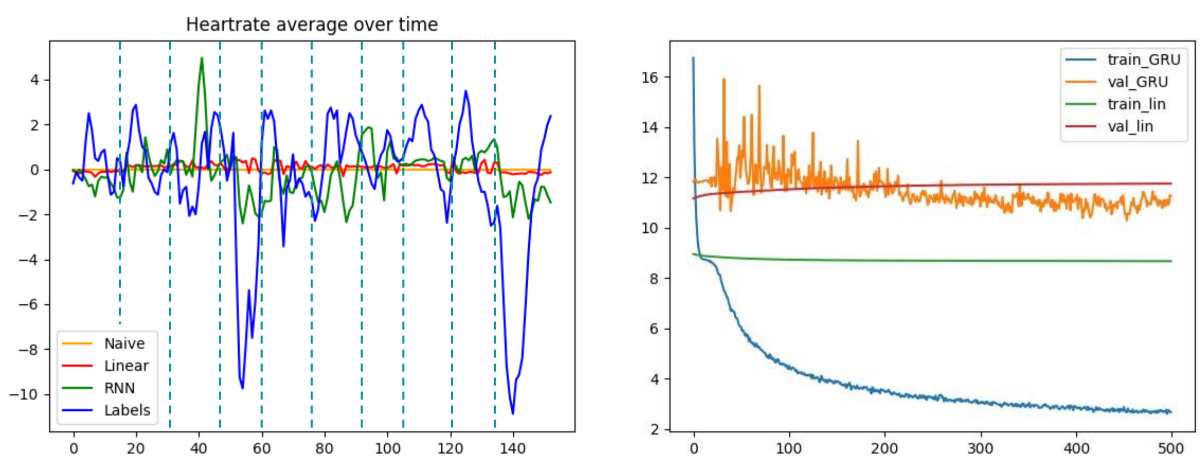


Figure 24 The validation predictions (left) and training results (right), where the validation data is a portion of data from all the patients. Note that the different portions are depicted by the vertical lines in the left image.

In Figure 24 the model results are plotted. Note for the RNN, the internal state is reset for each patient, i.e. the left image is not a continuous prediction from the RNN.

MODEL TRAINING CONCLUSIONS

Based on the results above, we can conclude that the models have similar results (if not worse results) compared to the naïve predictors. This indicates that the data and labels we have, have no deeper relationship to them as more advanced predictors (i.e. RNN and Linear models) do not consistently outperform the naïve predictors. Hence, we conclude that the dataset provided and the selected labels of interest are not featureful enough to show proper machine learning training.

APPENDIX B: QUESTIONNAIRES TO PATIENTS (USE CASE 2 SPECIFIC)

The following tables provide the questionnaires that were presented to the patients enrolled in the UC2 pilot study, along with the results collected.

Table 8. Questionnaire provided to patients enrolled in the Use Case 2.

	Question	Possible answers with the respective grading
1	How long did it take you to familiarize yourself with the use of our e-health system?	1: More than 2 months 2: 1-2 months 3: 2-4 weeks 4: 7-14 days 5: Less than 1 week
2	How complex did you find the use of our e-health system?	1: Far too complex to use 2: Quite complex to use – complexity limited my engagement 3: Somewhat complex – complexity did not limit my engagement 4: Slightly complex – I faced only minor difficulties 5: Not complex at all – I faced no difficulties
3	Please grade your satisfaction with respect to our e-health system's responsiveness:	1: Not responsive at all – unresponsiveness led me to quit using it 2: Quite unresponsive – unresponsiveness posed difficulties / consumed time, but did not lead me to quit use 3: Neutral 4: Quite responsive – occasionally I faced some difficulties, but I was mostly satisfied 5: Perfectly responsive
4	How secure did you feel with respect to personal data protection while using our e-health system?	1: Very insecure 2: Somewhat insecure 3: Neutral 4: Quite secure 5: Perfectly secure
5	How much do you believe the e-health system helped you in improving management of your health?	1: Not at all 2: Slightly 3: Moderately 4: Much 5: Very much
6	How do you believe that the use of our e-health system affected interaction / relationship with your attending physician, as compared to standard-of-care?	1: Very negatively 2: Negatively 3: Neutrally 4: Positively 5: Very positively

7	Do you believe that the use of the e-health system could prevent some unscheduled visits to healthcare facilities	1: Not at all – could actually lead to an increase in unscheduled visits 2: Probably not 3: Perhaps 4: Probably 5: Surely
8	Did the use of the e-health system make you feel safer?	1: Surely not 2: Probably not 3: Perhaps 4: Probably 5: Surely
9	Do you believe that an e-health system like this should be increasingly used in the future?	1: Surely not 2: Probably not 3: Perhaps 4: Probably 5: Surely
10	Would you recommend this e-health system to other patients?	1: Surely not 2: Probably not 3: Perhaps 4: Probably 5: Surely

Subject ID	Question a)	Question b)	Question c)	Question d)	Question e)	Question f)	Question g)	Question h)	Question i)	Question j)	Total score
1	4	4	1	4	4	5	4	4	5	5	**40**
2	5	5	4	5	4	5	4	4	5	5	**46**
3	5	4	4	4	5	5	4	5	5	5	**46**
4	4	3	4	4	4	5	3	5	4	4	**40**
5	4	4	5	5	5	5	5	5	5	5	**48**
6	5	5	4	5	4	4	3	4	5	4	**43**
7	3	2	2	3	3	4	2	2	4	4	**29**
8	5	3	2	5	4	4	4	4	5	4	**40**
9	5	5	4	5	3	5	3	5	5	5	**40**
10	4	4	4	4	4	4	4	5	5	5	**43**
11	5	4	5	5	4	5	4	5	3	5	**45**
12	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	**N/A**
14	5	5	4	4	4	4	2	4	4	4	**40**
15	3	3	5	5	5	5	5	5	3	5	**44**
16	5	5	4	5	4	4	4	5	4	5	**45**
17	5	5	4	4	4	3	4	4	5	5	**43**
18	5	3	2	5	1	4	4	4	5	5	**38**
19	4	4	4	5	5	4	5	5	5	5	**46**
20	5	5	5	5	5	5	3	5	4	5	**47**

Figure 25. Patients' questionnaire recorded answers.

APPENDIX C: QUESTIONNAIRES TO EXTERNAL STAKEHOLDERS AND RESULTS

The IntellIoT consortium members used the opportunity of end-user workshops and other events (as well as contacts from other EU projects and/or partnerships) to circulate questionnaires in order to assess the perception of industry stakeholders and researchers associated with the field of NG-IoT applications towards specific issues on the field and solutions that IntellIoT proposes. In this Appendix, the specific questions presented to the external stakeholders are presented along with the results that have been received. An analysis of the answers, especially concerning free-form answers (i.e. answers that the correspondents freely expressed their opinions instead of providing a Yes/No or degree of agreement/disagreement reply) that have been gathered.

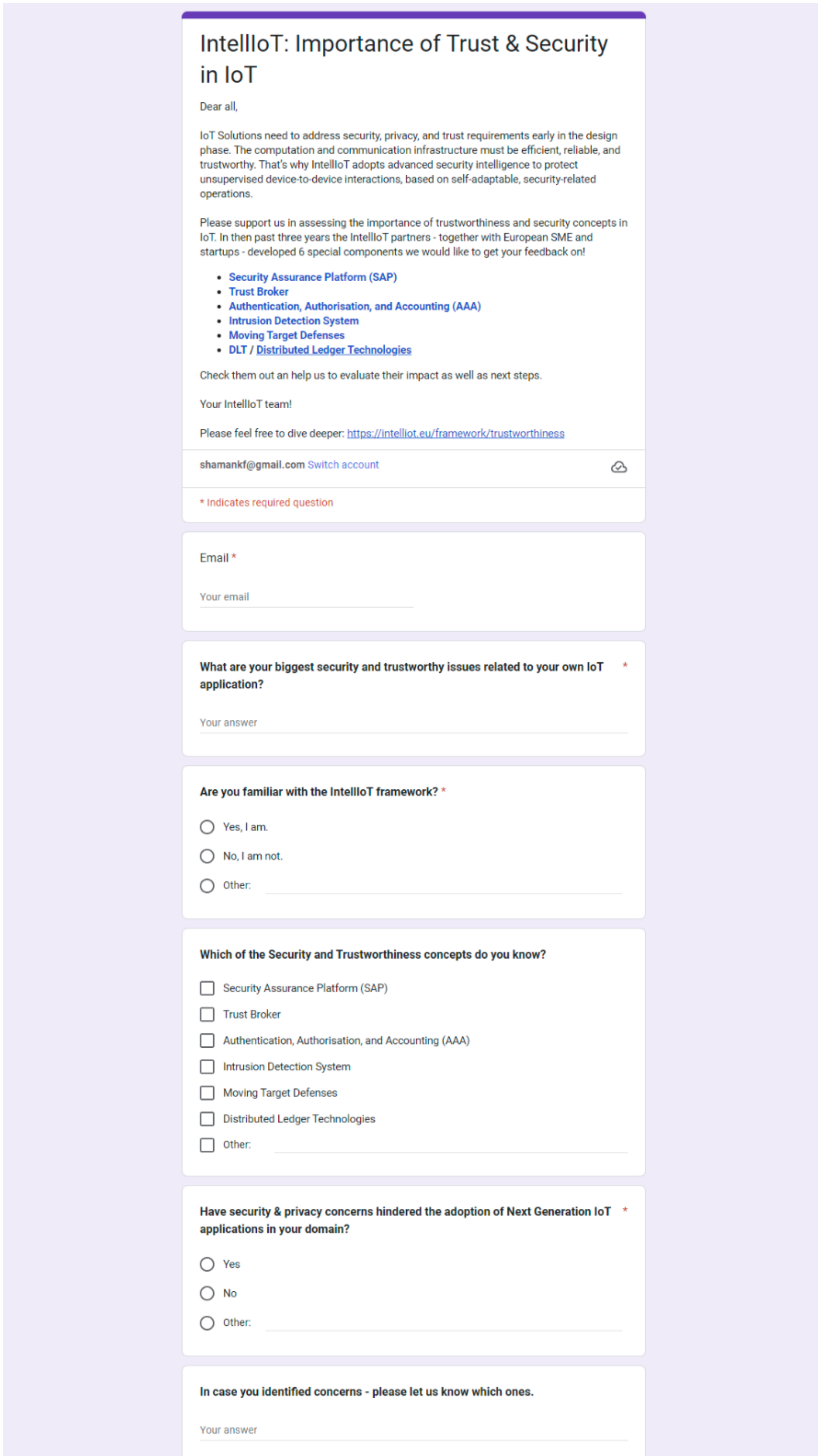
Questionnaires

The following questionnaire has been provided to external stakeholders through an online Google Form (a screenshot is also provided after Table 9). It should be noted that in the beginning of the questionnaire an introduction to the IntellIoT components is provided by including links to information material of all components as well as the overall IntellIoT website.

Table 9. Questionnaire provided to external stakeholders

	Question	Reply Format
1	What are your biggest security and trustworthy issues related to your own IoT application?	Free text
2	Are you familiar with the IntellIoT framework?	**Choices**: Yes, I am / No I am not / Other
3	Which of the Security and Trustworthiness concepts do you know?	**Multiple Choices**: Security Assurance Platform (SAP) / Trust Broker / Authentication, Authorisation, and Accounting (AAA) / Intrusion Detection System / Moving Target Defenses / Distributed Ledger Technologies / Other
4	Have security & privacy concerns hindered the adoption of Next Generation IoT applications in your domain?	**Choices**: Yes / No / Other:
5	In case you identified concerns - please let us know which ones.	Free text
6	Do you believe that the inclusion of trustworthy IoT capabilities, as proposed by IntellIoT, would facilitate the deployment of Next Generation IoT applications in verticals domains (e.g., agriculture, healthcare, manufacturing)?	**Choices**: Yes / No / Other:
7	How could such an inclusion look like?	Free text
8	The Trust-related technologies provided by IntellIoT help cover the security & privacy needs of NG-IoT applications.	**Scale**: 1: Strongly Disagree up to 5: Strongly Agree

9	If you already operate an IoT-enabled application with trustworthiness provisions, do you agree that IntellIoT's security & trustworthiness technologies offer a specific advantage in comparison to the relevant solutions that you already employ?	**Scale**: 1: Strongly Disagree up to 5: Strongly Agree
10	From a business (added value) perspective, the increased security, privacy & overall trustworthiness emphasised by IntellIoT can be considered an important competitive advantage for NG IoT market offerings.	**Scale**: 1: Strongly Disagree up to 5: Strongly Agree
11	If you agree that increased security, privacy & overall trustworthiness can be considered an important competitive advantage for NG IoT market offerings, can you elaborate on the specific competitive advantage that the NG IoT market offerings you are associated with, stand to gain?	Free text
12	What is your background?	**Choices**: I am a technical expert / I am a business expert. / I am a user of IoT Solutions, e.g. in manufacturing, smart city, agriculture etc. / Other:
13	Which part of the world are you from?	**Choices**: Europe / Asia / Africa / North America / South America / Other:



IntellIoT: Importance of Trust & Security in IoT

Dear all,

IoT Solutions need to address security, privacy, and trust requirements early in the design phase. The computation and communication infrastructure must be efficient, reliable, and trustworthy. That's why IntellIoT adopts advanced security intelligence to protect unsupervised device-to-device interactions, based on self-adaptable, security-related operations.

Please support us in assessing the importance of trustworthiness and security concepts in IoT. In then past three years the IntellIoT partners - together with European SME and startups - developed 6 special components we would like to get your feedback on!

- Security Assurance Platform (SAP)
- Trust Broker
- Authentication, Authorisation, and Accounting (AAA)
- Intrusion Detection System
- Moving Target Defenses
- DLT / Distributed Ledger Technologies

Check them out an help us to evaluate their impact as well as next steps.

Your IntellIoT team!

Please feel free to dive deeper: https://intelliot.eu/framework/trustworthiness

shamankf@gmail.com Switch account

* Indicates required question

Email *

Your email

What are your biggest security and trustworthy issues related to your own IoT application? *

Your answer

Are you familiar with the IntellIoT framework? *

- Yes, I am.
- No, I am not.
- Other:

Which of the Security and Trustworthiness concepts do you know?

- Security Assurance Platform (SAP)
- Trust Broker
- Authentication, Authorisation, and Accounting (AAA)
- Intrusion Detection System
- Moving Target Defenses
- Distributed Ledger Technologies
- Other:

Have security & privacy concerns hindered the adoption of Next Generation IoT applications in your domain? *

- Yes
- No
- Other:

In case you identified concerns - please let us know which ones.

Your answer

Figure 26. Snippet of the Questionnaire circulated

Results

Demographics

We received 25 responses representing individuals from industry (19) and academia (6). It should be noted that all correspondents are **not** affiliated with any of the IntellIoT partners. The majority of the responders' background is reported as *Technical* (72% - 18/25), while 24% (6/25) are working more on the business development side. A single responder identified himself/herself as R&D project manager with Technical Background. A significant number of responders had good (14/25) or partial (3/25) familiarity with the IntellIoT project, while the rest of them were only informed about the project using the publicly available material. It should be noted that the respondents that expressed familiarity with the project (good or partial) were affiliated with organizations that participated in one or more events (calls, workshops, etc) that IntellIoT organized. Lastly, 24 out of the 25 respondents are located in Europe, while a single respondent is located in Africa.

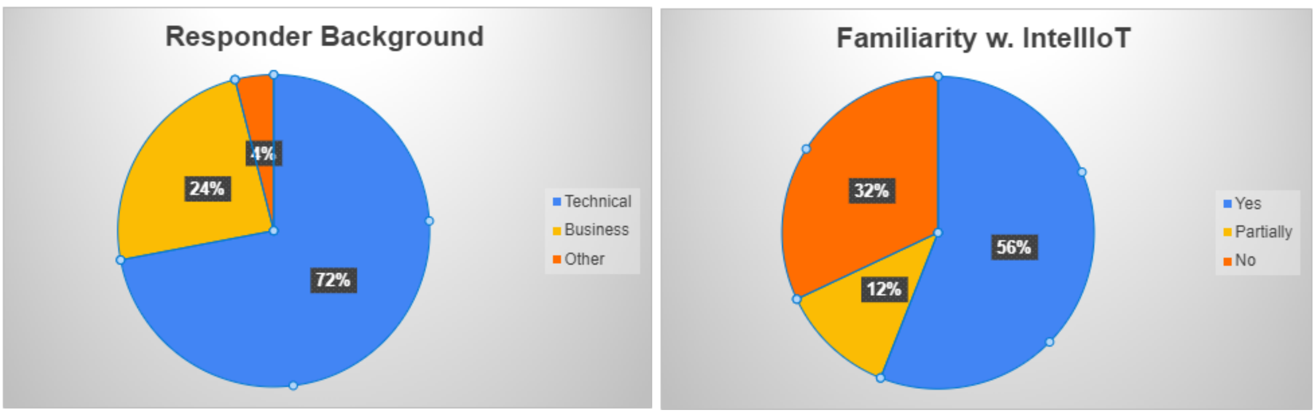


Figure 27. Basic Demographics of the Respondents

Familiarity with the concepts presented

When presented with the Security and Trustworthiness concepts, all the respondents were familiar with at least a number of them. The graph below focuses on the concepts that IntellIoT directly addresses. As indicated in the *Other* field of the acceptable reply, additional concepts that the respondents were familiar with are related to Data Loss Prevention, Mobile/IoT Management as well as security technologies marked as Proprietary by them.

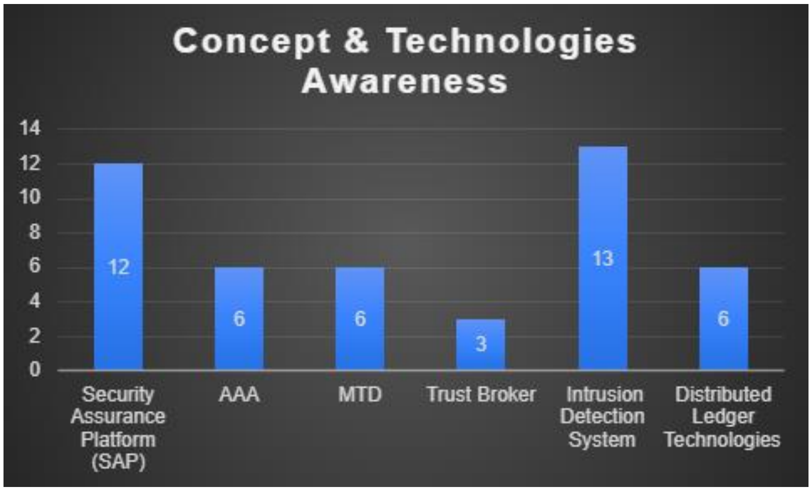
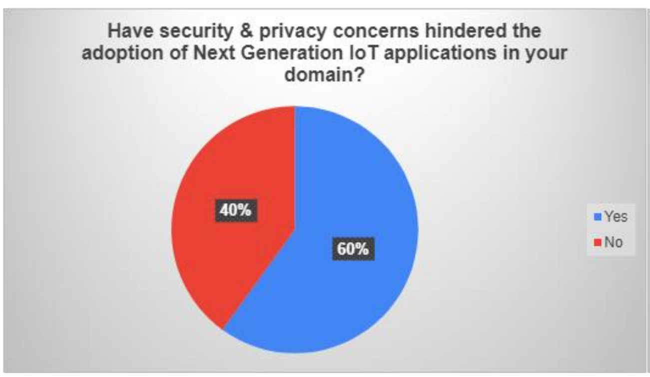
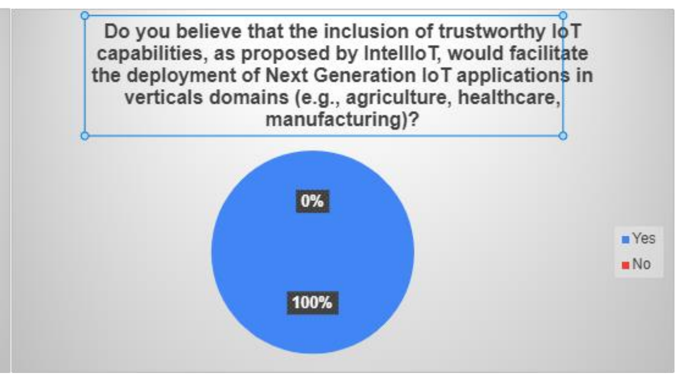


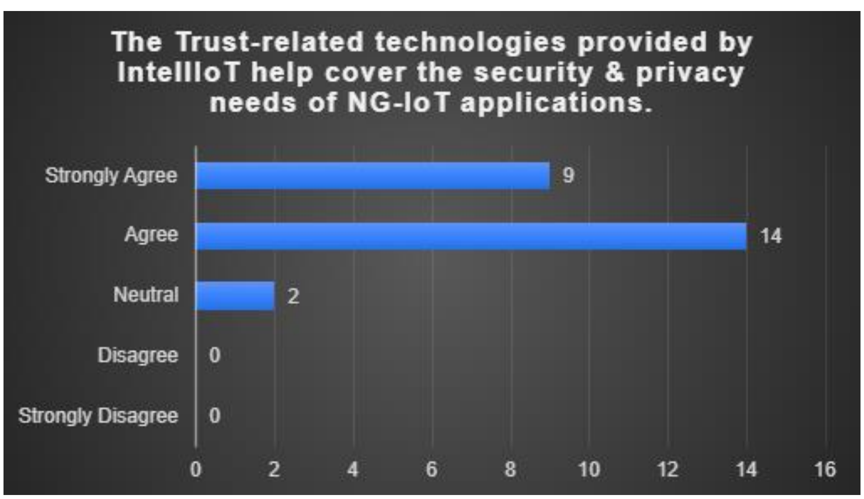
Figure 28. Concept and Technology Awareness of the Questionnaire Respondents

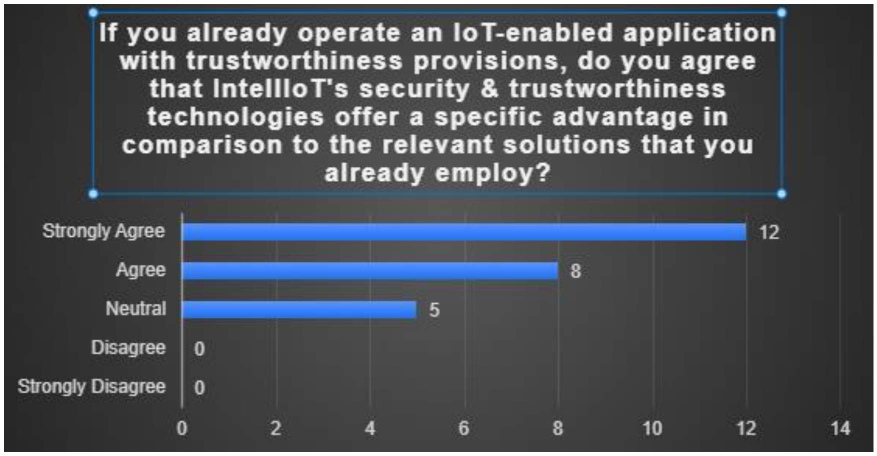
Quantifiable Responses

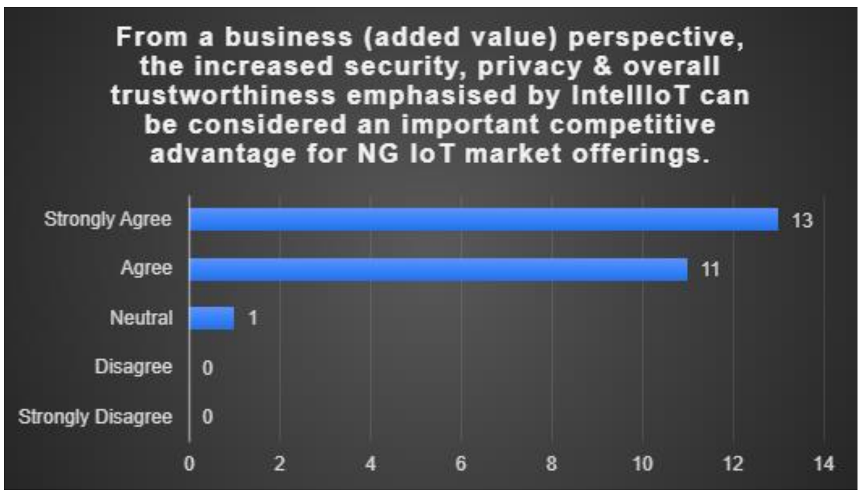
The following graphs present the results gathered on questions that provided prespecified answers (i.e. questions that allowed a Yes/No/Other reply or a scaled reply between Strongly Disagree and Strongly Agree).











Non-Quantifiable Responses

In several questions (more specifically questions 1, 5, 7, and 11), respondents were asked to provide their input in free text form. In the following table, we try to summarize the replies that we have received (we provide the general context and not the details related to the specific applications of the responders).

Table 10. Summary of the responses received when respondents were asked to express their opinion

	Question	Reply Summary
1	What are your biggest security and trustworthy issues related to your own IoT application?	Most of the responders express concerns about data privacy and protection both for stored data or data in transit. Concerning the later, securing communications seems to be a significant issue, however it is also related to communication reliability so that data can be exchanged without corruptions and in timely manner. Also, the issue of proper authentication mechanisms for both users and devices (especially third party devices participating in an IoT application) seems to raise a lot of concern. Physical tampering of IoT devices is reported as a major issue.

		Lastly, responders stress regulatory compliance both in terms of security and privacy are very important for user adoption, as well as the existence of security mechanisms to prevent any incidents that will cause mistrust and reputation loss from the users.
5	In case you identified concerns - please let us know which ones.	Privacy and related legal and ethical issues are the top concerns, especially for applications that interact with humans and gather personal data. Also the complexities and labor-intensive setup and maintenance of security mechanisms remains a significant concern.
7	How could such an inclusion look like?	Responders stress the improvement of the security and privacy of their applications, especially when user trust needs to be gained or regulatory approval is mandatory.
11	If you agree that increased security, privacy & overall trustworthiness can be considered an important competitive advantage for NG IoT market offerings, can you elaborate on the specific competitive advantage that the NG IoT market offerings you are associated with, stand to gain?	Responders mention that the adoption of a trusted platform is both a way to address market requirements (e.g. regulations) and a way to stand-out among competitors that have not fully tackled security and privacy issues. It is also pointed out that securing the IoT applications is a means to provide uninterrupted services to customers. Lastly, the high impact of cybersecurity incidents is stressed (both related to service interruption and costs as well as reputation damage).