Ref. Ares(2026)754347 - 22/01/2026

AWARE Exploratory Research Plan - Final

Deliverable ID:	D2.4
Project acronym:	AWARE
Grant:	101167442
Call:	HORIZON-SESAR-2023-DES-ER-02
Topic:	HORIZON-SESAR-2024-DES-ER2-WA2-4
Consortium coordinator:	FTTS
Edition date:	22 January 2026
Template edition:	02.00.01
Edition:	00.06
Status:	Official
Classification:	PU

Abstract

The AWARE project's Exploratory Research Plan (ERP) aims to develop an Artificial Situational Awareness (ASA) system to enhance collaboration between Air Traffic Control Officers (ATCOs) and automated systems in air traffic management (ATM). Building on prior research, AWARE seeks to overcome limitations in human-machine interaction by enabling the ASA system to predict ATCO intentions and provide timely, context-sensitive support. The ERP outlines a structured methodology and validation exercises to measure the ASA system's impact on ATCO workload, situational awareness, and operational safety. Through simulations under varied traffic conditions, the project aims to improve ATM efficiency, capacity, safety, and environmental sustainability.

Authoring & Approval

Author(s) of the document

Name/Beneficiary short name	Position/Title	Date
Sebastian Soldat/FTTS	FTTS contributor	01/09/2025
Ivor Vrhovec/FTTS	FTTS Contributor	01/09/2025
Pavle Dujmović/FTTS	FTTS Contributor	01/09/2025
Tomislav Radišić/FTTS	FTTS Contribution leader	15/10/2025
Kristina Samardžić/FTTS	FTTS Contributor	18/11/2025

Reviewed by

Name/Beneficiary short name	Position/Title	Date
Tomislav Radišić/FTTS	FTTS contribution leader	27 November 2025
Roland Guraly/SLOT	SLOT contribution leader	20 November 2025
Hólmfríður Elvarsdóttir/TERN	TERN contribution leader	20 November 2025
Javier Alberto Pérez Castán/UPM	UPM contribution leader	20 November 2025
Celina Vetter/ZHAW	ZHAW contribution leader	20 November 2025

Approved for submission to the SESAR 3 JU by

Name/Beneficiary short name	Position/Title	Date
Tomislav Radišić/FTTS	FTTS contribution leader	22 January 2026

Rejected by[1]

Name/Beneficiary short name	Position/Title	Date

Document History

Edition	Date	Status	Company Author	Justification
00.01	01/09/2025	Initial Draft	S. Soldat, I. Vrhovec	Update from intermediate version
00.02	15/09/2025	Draft	S. Soldat, P. Dujmović, T. Radišić	New sections added

[1] Representatives of the beneficiaries involved in the project

EUROPEAN PARTNERSHIP
Co-funded by the European Union

00.03	15/10/2025	Complete Draft	T. Radišić	Sent for internal review
00.04	25/11/2025	Second Draft	K. Samardžić	Update
00.05	27/11/2025	Final	T. Radišić	Final
00.06	22/01/2026	Updated Final	T. Radišić	Updated according to SJU comments

Copyright Statement

AWARE

Achieving Human-Machine Collaboration with Artificial Situational Awareness

AWARE

This document is part of a project that has received funding from the SESAR 3 Joint Undertaking under grant agreement No 101167442 under European Union’s Horizon Europe research and innovation programme.

Table of contents

List of figures

List of tables

1 Executive summary

The AWARE Exploratory Research Plan (ERP) defines a structured approach for developing and validating the Artificial Situational Awareness (ASA) system, intended to support Air Traffic Control Officers (ATCOs) in managing complex and high-traffic airspaces through improved human-machine collaboration for Air Traffic Management (ATM). The primary purpose of this ERP is to provide a clear and methodical framework for conducting validation exercises to test ASA capabilities, addressing ATCOs' needs for enhanced situational awareness and real-time operational support.

The ERP outlines a series of experiments that will progress the ASA system from an initial maturity level (TRL1) to an exit maturity level (TRL2). These experiments are designed to test ASA's ability to predict ATCO intentions accurately, reduce ATCO workload, and assist in decision-making across different traffic scenarios. The validation exercises range from baseline scenarios—where ATCOs manage traffic without ASA support—to scenarios where ASA offers partial and full assistance, providing ATCOs with insights, real-time recommendations, and automated support for specific tasks.

The planned experiments and validation exercises, which are set to occur in simulated high and projected future traffic volumes, cover key objectives including ATCO intent recognition, system performance under variable conditions, and ASA's ability to reduce human workload. Each exercise is tailored to measure ASA's impact on ATM efficiency, capacity, environmental sustainability, and overall safety. These objectives are critical for the ERP, as they address the need for innovative solutions to meet the demands of increasingly complex airspaces within the European air traffic management system.

Validation will proceed through controlled simulations that mimic real-world operational demands, assessing ASA's compatibility with ATCO workflows and its potential to improve air traffic flow, reduce delays, and optimise environmental outcomes. The validation objectives focus on demonstrating that ASA can achieve high levels of accuracy in task execution and effectively support ATCO decision-making without inducing additional workload.

The experiments under the ERP form a comprehensive roadmap to ensure ASA's readiness for future operational use. By validating ASA's capabilities within the SESAR framework, the ERP aligns with SESAR 3 Joint Undertaking's vision of enhancing safety, efficiency, and automation in European ATM, establishing a new benchmark for human-machine collaboration in air traffic control.

EUROPEAN PARTNERSHIP
Co-funded by the European Union

2 Introduction

2.1 Purpose of the document

The purpose of this Exploratory Research Plan is to outline the initial research objectives, methodologies, and anticipated outcomes for developing the ASA system in air traffic management. This document provides a framework for exploring innovative solutions aimed at enhancing collaboration between ATCOs and automated systems. It also identifies key areas of uncertainty and research questions that will guide future validation efforts. The ERP serves as a foundational guide to ensure that research activities align with the project's long-term goals of improving safety, efficiency, and automation in ATM operations.

2.2 Intended readership

The main audience of this Exploratory Research Plan document is:

- SESAR Joint Undertaking,
- Air Navigation Service Providers (ANSPs),
- Air Traffic Controllers,
- Airspace Users,
- UAV Users,
- Air Traffic Management research community,
- Air Traffic Management system developers, and
- regulators in the field of aviation

2.3 Background

The AWARE solution builds on the SESAR project AISA, which developed an artificial situational awareness system using aeronautical data exchange models like Aeronautical Information Exchange Model (AIXM) and Flight Information Exchange Model (FIXM). This system created a knowledge graph of the traffic situation, enabling a reasoning engine to generate insights and automate monitoring tasks. Enhanced with machine learning modules for predicting future states, the proof-of-concept was tested in human-in-the-loop simulations. Results showed the system's situational awareness was comparable to that of human operators in en-route air traffic control (ATC), providing an additional safety net by automating 47 tasks and suggesting actions for ATCOs.

However, the system faced challenges in collaborating with ATCOs due to:

- Lack of awareness of ATCOs' intents and goals, resulting in untimely support.
- Ineffective human-machine interface (HMI) to convey information.
- Inflexibility in adjusting the level of support.

A major barrier to effective human-machine collaboration in ATM is the machine's inability to understand ATCOs' needs, providing appropriate information or actions at the right time. AISA highlighted this issue by exploring shared situational awareness but failed to address the challenge of inputting to ATCOs without disrupting their workflow. This disconnect indicates that while humans understand machine capabilities, machines lack reciprocal understanding of human operators. In simple interactions, this is manageable, but in complex environments with diverse traffic, it is crucial for machines to understand humans to achieve effective collaboration.

The ambition is to expand the artificial situational awareness system's functionality to provide timely and relevant support by assessing ATCOs' intents and goals. This support can assist ATCOs in tasks they intend to perform and autonomously execute tasks outside their immediate attention. Developing a combination of machine learning (ML) and non-ML tools will determine the feasibility of such supportive actions. The goal is for ATCOs to set the desired level of support from the artificial intelligence (AI), maintaining situational awareness while avoiding complacency and the out-of-the-loop effect.

AWARE's approach focuses on developing a proof-of-concept AI Assistant Application to demonstrate the Artificial Situational Awareness solution, which supports air traffic control operators by identifying and aiding in their goals (Figure 1). The ASA system (green block in Figure 1), developed during the AISA project, uses aeronautical information exchange models to create a knowledge graph representing the traffic situation, converted to the Resource Description Framework. This graph, enhanced with rule-based knowledge and a reasoning engine, builds a semantic understanding of the traffic situation but cannot assess ATCOs' goals and actions. To enable effective human-machine collaboration (blue block in Figure 1), the system will track ATCOs' attention and HMI. By contextualizing this information within the traffic situation, the system will infer ATCOs' intent and determine supportive actions.

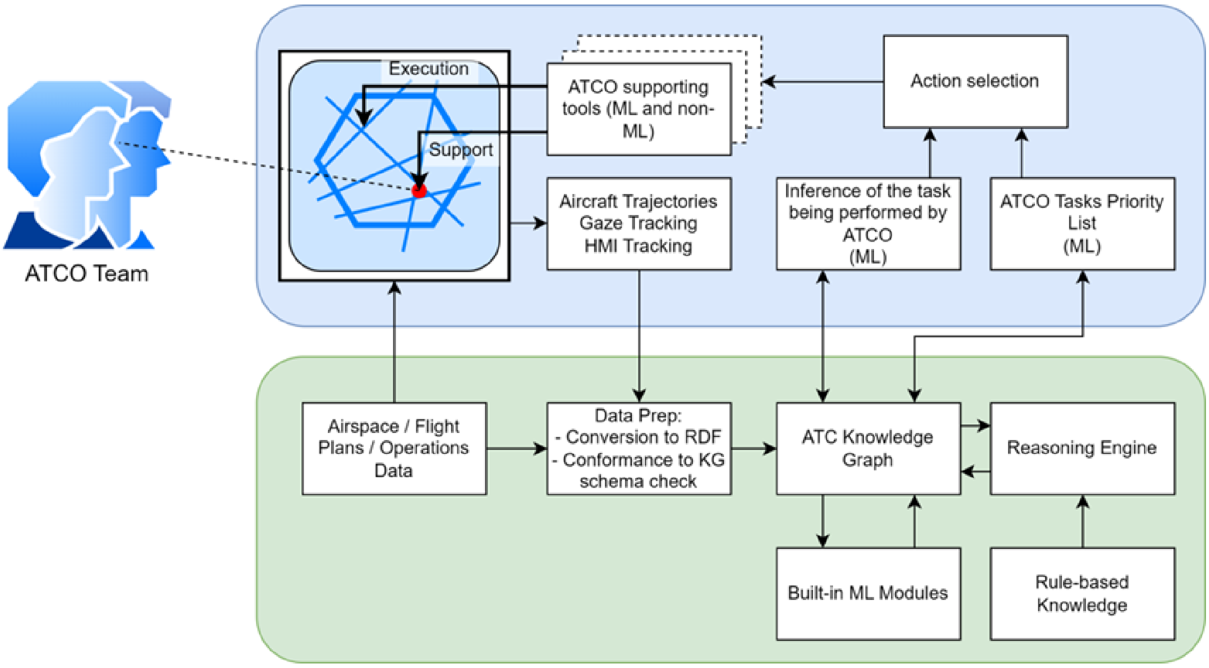

Figure 1. AI Assistant Application

The system will explore two ways of assisting ATCOs: i) helping with tasks they are currently focused on and ii) managing tasks outside their immediate attention. Adaptability is ensured as ATCOs can control the process by dynamically selecting the number of actions the system provides per unit of time. The following sections detail the concepts and methodologies involved in this approach. Enhancing the ASA system involves expanding its capabilities with new knowledge-graph schemas, vocabulary, and rule-based knowledge. Implementing an effective HMI at the controller's working position will facilitate information sharing and enable ATCOs to select the required level of automation support, ensuring the human remains central to the system.

2.4 Structure of the document

This Exploratory Research Plan document is structured as follows:

1. Executive Summary of the document,
2. Introduction of the document,
3. Context of the exploratory research plan. This chapter describes the context of the exploratory research plan for the AWARE solution. This chapter is an introduction to the following chapters, where the exploratory research plan is detailed.
4. Exploratory research plan. This chapter describes the exploratory research plan approach designed to progress the AWARE solution from the initial maturity level to the exit maturity level.
5. Validation exercises. This chapter lists and describes the exercise validation plans.
6. References. This chapter lists the main references the Concept of Operations document shall comply to or on which it has been built.

2.5 Glossary of terms

Term	Definition	Source of the definition
AI ASSISTANT APPLICATION (AAA)	The AI Assistant Application (AAA) supports ATCOs by providing real-time insights, automating tasks, and enhancing decision-making through data integration and machine learning.	Part B of the AWARE project proposal document ([2])
ARTIFICIAL SITUATIONAL AWARENESS (ASA) SYSTEM	Artificial Situational Awareness system that uses AI to enhance ATCOs' situational awareness by automating monitoring, predicting traffic, and assisting decision-making.	Part B of the AWARE project proposal document ([2])
OUT-OF-THE-LOOP EFFECT	The out-of-the-loop effect occurs when automation reduces human engagement, leading to decreased situational awareness and impaired decision-making during critical tasks.	
RDFox	High-performance knowledge graph reasoner	
SPARQL	SPARQL Protocol and RDF Query Language	

Table 1. Glossary of terms

2.6 List of acronyms

Term	Definition
AI	Artificial Intelligence
AIM	Aeronautical Information Model
AISA	Artificial Situational Awareness system
AIXM	Aeronautical Information Exchange Model
ANSP	Air Navigation Service Providers
AOI	Area of Interest
ASA	Artificial Situational Awareness
ATC	Air Traffic Control
ATCO	Air Traffic Control Officer
ATM	Air Traffic Management
ATSEP	Air Traffic Safety Electronics Personnel
AWARE	Achieving Human-Machine Collaboration with Artificial Situational Awareness

EUROPEAN PARTNERSHIP
Co-funded by the European Union

Term	Definition
C/AP	Capacity
CEF	Cost Efficiency
CPDLC	Controller-Pilot Data Link Communications
EASA	European Union Aviation Safety Agency
ECAC	European Civil Aviation Conference
ECG	Electrocardiogram
ENV	Environment
ERP	Exploratory Research Plan
FEFF	Fuel Efficiency
FIXM	Flight Information Exchange Model
FRD	Functional Requirements for Development
HEG	Hemoencephalography
HITL	Human-In-The-Loop
HMI	Human-Machine Interface
HP	Human Performance
HR	Heart Rate
HRV	Heart Rate Variability
JU	Joint Undertaking
KG	Knowledge Graph
KPA	Key Performance Area
KPI	Key Performance Indicator
ML	Machine Learning
NLP	Natural Language Processing
OSED	Operational Service and Environment Description
R&I	Research and Innovation
SA	Situational Awareness

Term	Definition
SAF	Safety
SESAR	Single European Sky ATM Research
SHACL	Shapes Constraint Language
SME	Subject Matter Expert
SUP	ATCO Supervisors
SWIM	System Wide Information Management
TA	Technical Area
TRL	Technology Readiness Level
UAV	Unmanned Aerial Vehicle
UC	Use Case

Table 2. List of acronyms

EUROPEAN PARTNERSHIP
Co-funded by the European Union

3 Context of the exploratory research plan

3.1 Exploratory research plan context

The Exploratory Research Plan is framed within the context of advancing SESAR Solutions that aim to improve air traffic management through enhanced automation and collaboration between human operators and machine systems. This ERP focuses on the AWARE solution, designed to support ATCOs in maintaining situational awareness (SA) during complex air traffic scenarios. The research targets research areas such as Modes of Human-Machine Collaboration, Attention Tracking and Intent Analysis, ATCO Supporting Tools and the Enhanced Artificial Situational Awareness System. The concept will be assessed by running HITL (human-in-the-loop) simulations where ATCOs will be supported by the AI assistant application.

The SESAR vision for a fully digitalized European sky is the driving force behind this research. Increasing air traffic demand, the introduction of unmanned aerial vehicles (UAVs), and the integration of new entrants into the airspace necessitate advanced automation to maintain safety and efficiency, [4]. The ERP outlines the challenges addressed by the AWARE solution, which are posed by the growing complexity of air traffic operations, as well as the need for more adaptable and resilient systems to manage airspace dynamically.

A central aspect of this ERP is to explore how the ASA system can bridge the gap between human operators and highly automated systems. The key research questions revolve around improving the interpretability, accuracy, and timeliness of machine-generated support provided to ATCOs. Furthermore, the plan investigates how automation can be introduced incrementally into air traffic control environments without disrupting human operators' workflows, while ensuring that human-machine collaboration is optimised for enhanced decision-making.

This exploratory phase is crucial for identifying potential breakthroughs and technological uncertainties, which will serve as the basis for the development, validation, and eventual deployment of SESAR Solutions.

The AWARE experiments and validation activities will take place in Sweden, focusing on the en-route phase of air traffic control. The technical environment emphasizes human-AI teaming, where artificial intelligence supports air traffic controllers by enhancing situational awareness and decision-making capabilities.

The key assumptions relevant to the scope of experiments or validation activities in the AWARE exploratory research plan are drawn from both technological, social, legal and scope assumptions.

3.2 Scope

The AWARE project's Exploratory Research Plan focuses on the development and validation of an ASA system designed to enhance the decision-making capabilities of ATCOs. By integrating AI into the ATC environment, the system aims to provide real-time assistance to ATCOs, improving safety and operational efficiency in complex and high-traffic airspace situations.

The AWARE project addresses numerous research questions aimed at improving human-AI collaboration within the air traffic management domain. Research questions can be divided into 5 categories as follows.

Intent

- Is AI Assistant Application able to recognise intention from categorization of gaze point sequences and interactions with HMI in connection with simulation data and by comparing it with a comprehensive list of en-route air traffic tasks with adequate level of accuracy?
- May temporary problems related to low eye-tracking reliability such as misidentification of gaze focus or calibration problems be compensated by data from interaction with HMI and vice versa?
- How much data input from eye-tracking and interaction with HMI is necessary for adequate prediction from modelling ATCOs' cognitive processes to recognise ATCO intention?

Situational awareness and workload

- Is it possible to assess ATCO's situational awareness by assessing their intent compared to the list of priority tasks?
- Is AI Assistant Application able to correctly identify increased and excessive mental workload compared to subjective workload assessment?

Supporting ATCOs

- Is support provided by the application acceptable to the ATCO?
- How do ATCOs choose to be supported? Are higher levels of automation (task execution) acceptable in situations with increased traffic volumes and high complexity?
- How can design solutions support intuitive assessment of appropriateness of assistance to prevent costs in terms of raising ATCO workload for processing information.
- What is the effect of the support provided by the application on ATCO's workload?
- Is the information that Will be presented to the ATCO as part of the supporting actions enough to have positive impact on any of the performance KPIs?

System performance

- Can the developed AISA system be improved to work sufficiently fast in real-time in interaction with ATM system?
- Can all monitoring tasks from AISA ConOps be automated?
- Is the accuracy acceptable for all monitoring tasks?

Safety and security

- Which hazards arise from the integration Of the Al Assistant Application?
- Is the safety improved in the proposed human-machine environment by detection and compensation for selective and focused attention in ATCO SA and support in situation awareness safety-critical tasks?
- Can the mitigation measures reduce the potential risk and increase safety to satisfy potential safety requirements?
- How can artificial situational awareness be used to mitigate cyber-security risks? Can it model the type of awareness required from the human operators?
- What are the sources of cyber-security risks in artificial situational awareness system?

The following objectives define the overarching goals of the AWARE project and align closely with the research questions:

- General Objective of the project - developing an AI Assistant Application (AAA) focused on improving ATC operations by leveraging Enhanced Artificial Situational Awareness to provide context-aware support in real time.
- Specific Objective 1 - implementing methods to assess ATCOs' intent and goals, allowing the AI system to better interpret and align with their decision-making processes during operations.
- Specific Objective 2 - creating a flexible approach that adapts AI assistance based on the current situation, using both Machine Learning (ML) and non-ML tools to support ATCOs effectively.
- Supporting Technical Objective 1 - enabling the system to track and interpret ATCO visual attention in combination with other data, providing actionable insights that reflect the current traffic situation.
- Supporting Technical Objective 2 - researching ways to detect and address any loss of situational awareness, ensuring ATCOs remain fully engaged and in control, even in complex or automation-heavy environments.
- Supporting Technical Objective 3 - defining interoperability specifications for the ASA system, enabling its integration with other ATM roles and systems, and exploring its potential benefits beyond ATC, including improvements in cybersecurity and certification processes.

The following table provides a concise overview that links each project objective to its relevant research questions, ensuring that each objective's scope and goals align with the questions driving the research and validation activities.

Project objective	Related research questions
General Objective	- Can the system developed in AISA be improved to work sufficiently fast in real-time in interaction with ATM system? - Can all monitoring tasks from AISA ConOps be automated? - Is the accuracy acceptable for all monitoring tasks? - Which hazards arise from the integration of the AI Assistant Application? - How can artificial situational awareness be used to mitigate cyber-security risks?
SO1	- Is the AI Assistant Application able to recognise intention from categorization of gaze point sequences and interactions with HMI in connection with simulation data? - How much data input from eye-tracking and interaction with HMI is necessary for adequate prediction of ATCO intentions? - May problems like low eye-tracking reliability be compensated by data from interaction with HMI?
SO2	- Is support provided by the application acceptable to the ATCO? - Are higher levels of automation acceptable in high-traffic situations? - How do ATCOs choose to be supported, and what is the effect on their workload? - How can design solutions prevent higher ATCO workload due to the support application?
STO1	- How much data input from eye-tracking and interaction with HMI is necessary for adequate prediction of ATCO intention? - May eye-tracking reliability problems be compensated by interaction with HMI data? - Is it possible to assess ATCO's situational awareness by assessing their intent compared to the list of priority tasks?
STO2	- Is AI Assistant Application able to identify increased and excessive mental workload compared to subjective workload assessment? - What is the effect of the application on ATCO's situational awareness? - Is the AI capable of detecting and compensating for loss of situational awareness in ATCOs?
STO3	- What are the sources of cyber-security risks in artificial situational awareness systems? - How can artificial situational awareness be used to mitigate cyber-security risks? - Can the system model the type of awareness required from human operators to mitigate these risks? - What are the potential benefits of ASA beyond ATC operations, such as in certification and pre-tactical operations?

Table 3. General objectives of the project and the relevant research questions, [2]

The AWARE project employs a phased research methodology that consists of:

1. Knowledge engineering - creating a knowledge representation by gathering and organising data relevant to ATC operations by using existing data models such as AIXM and FIXM to construct a knowledge graph that describes the current traffic situation.
2. Reasoning - applying AI and machine learning techniques to infer additional insights from the knowledge graph. By querying the knowledge graph, the AI assistant can predict future traffic situations and automate certain monitoring tasks, enhancing the situational awareness of the system.
3. Intent assessment - exploring how to align ATCO tasks identified by the artificial situational awareness system with the ATCO's predicted intent based on attention and HMI tracking. By

mapping tasks that need to be solved with those the ATCO intends to address, the system can assess the ATCO's goals.

4. Task extraction - identifying specific ATC tasks from the data interactions and situational context.
5. Task prioritisation - evaluating the relative importance of various tasks and suggesting a sequence based on urgency, workload, and safety.
6. Action selection - based on the prioritized tasks and reasoning outcomes, the system recommends or executes specific actions.
7. ATC tools - developing a set of ATCO support tools to demonstrate the system's ability to interact safely and effectively with the traffic situation, including tasks such as sector coordination, conflict detection and resolution, and flight handovers. These tools, leveraging both machine learning and non-ML approaches, aim to enhance the quality of service and support ATCOs in various operational areas.
8. Novel HMI components - the solution incorporates innovative HMI elements, such as eye-tracking and mouse-click monitoring, to enhance the interaction between the AI assistant and ATCOs. These components improve the system's ability to understand and support ATCOs' cognitive processes.

This approach ensures both rigorous testing of the ASA system and iterative development based on real-world feedback. The solution validation follows an approach that ensured the system meets its operational objectives. The process is structured as follows:

1. Creating airspace - the necessary airspace configurations are created for validation exercises. These configurations are tailored to the specific requirements of the project's scenarios and simulations.
2. Creating exercises - developing detailed exercises designed to ensure that the system can be validated effectively later, with clear scenarios that will test critical functionalities.
3. Preliminary simulation for data gathering - initial simulations are run to gather control data.
4. ATCO training – preparing the controllers for the Polaris system which includes manuals, videos and simulations.
5. Exercise run – ATCOs will perform four traffic scenarios, each representing a specific combination of traffic volume and complexity and designed to highlight a specific operational feature. They are distributed into four groups, with each group receiving a different level of ASA system support. One of these groups will serve as a validation control group.

The following table illustrates the clear connection between the research questions and the methodology, ensuring that the experimental and validation processes are targeted to answer the key research questions posed by the project.

Co-funded by the European Union

Research Questions		Methodology (Experimental and Validation)	Connection Between Methodology and Research Questions
Intent	Is the AI Assistant Application able to recognise intent from categorization of gaze point sequences and interactions with HMI in connection with simulation data?	- Knowledge engineering - Intent assessment - Task extraction - Task prioritization - Action selection - Preliminary simulation for data gathering	The AI's predicted intention (what the AI expects the ATCO to do next) is compared to the actual action taken by the ATCO, allowing the system to assess the accuracy of its intent recognition. This is done by analysing gaze point sequences and HMI interactions in simulated scenarios.
	May temporary problems like low eye-tracking reliability be compensated by interaction with HMI data? And vice versa?	- Knowledge engineering - Reasoning (rules) - ATC tools (support and execution) - Creating airspace - Preliminary simulation for data gathering	Interaction with HMI data is combined with rule-based reasoning to compensate for any eye-tracking errors, ensuring reliable intent recognition.
	How much data input from eye-tracking and interaction with HMI is necessary for adequate prediction of ATCO cognitive processes?	- Task extraction - Task prioritization - Intent assessment - Creating exercises - Preliminary simulation for data gathering	Task extraction methods analyse data input from eye-tracking and HMI interactions to predict ATCO cognitive processes, validated through exercises designed to assess accuracy.
Situational Awareness and Workload	Is it possible to assess ATCO's situational awareness by assessing their intent compared to the list of priority tasks?	- Intent assessment Task prioritization - Reasoning (rules) - Creating exercises	Intent assessment and task prioritization techniques compare ATCOs' actions to predefined priority tasks, allowing validation through exercises that measure situational awareness under different conditions that will be tested while running HITL experiment.
	Is AI Assistant Application able to correctly identify increased and excessive mental workload compared to subjective workload assessment?	- Task prioritization - Action selection - ATCO training (preparing for new system) - Creating exercises	Task prioritization and action selection methods assess mental workload, validated by comparing AI-identified workload levels with subjective workload assessments in HITL training exercises.

Research Questions		Methodology (Experimental and Validation)	Connection Between Methodology and Research Questions
Supporting ATCOs	Is support provided by the application acceptable to the ATCO?	- ATC tools (support and execution) - Action selection - First exercise run (baseline scenario) - Second exercise run (human-AI teaming)	The acceptability of the AI support will be evaluated through questionnaires administered to ATCOs after each exercise. These questionnaires will assess their satisfaction with the AI's assistance, ease of use, and overall impact on their performance. Feedback will provide direct insights into whether the AI support meets ATCO expectations and operational needs.
	How do ATCOs choose to be supported? Are higher levels of automation acceptable in situations with high traffic volumes and complexity?	- Task extraction - Task prioritization - ATCO training - Second exercise run (human-AI teaming)	ATCOs’ preference for support and automation levels is explored through training and human-AI teaming exercises in high-complexity scenarios that we will test in HITL experiments, ensuring usability in real-world operations.
System Performance	Can the system, developed in AISA, be improved to work sufficiently fast in real-time in interaction with ATM system?	- Knowledge engineering - Reasoning (rules) - Preliminary simulation - ATC tools (support and execution)	The system's real-time performance will be evaluated by measuring the time it takes to process incoming data. Specifically, the AI must process the data within a defined time frame. For example, if data enters the system every X seconds, the AI's processing time must be less than X seconds. This ensures the system can handle continuous data input without causing delays in ATC decision-making, maintaining real-time interaction with the ATM system.
	Can all monitoring tasks from AISA ConOps be automated?	- Task extraction - ATC tools (support and execution) - Preliminary simulation for data gathering	Task extraction methods are used to automate monitoring tasks, with results validated through simulations to determine the feasibility of automation for all AISA ConOps tasks.

Research Questions		Methodology (Experimental and Validation)	Connection Between Methodology and Research Questions
Safety and Security	Which hazards arise from the integration of the AI Assistant Application?	- Reasoning (rules) Creating airspace - Preliminary simulation - ATC tools (support and execution)	Safety hazards are identified through rule-based reasoning and simulations, using ATC tools to test AI integration and ensure risk mitigation strategies.
	How can artificial situational awareness be used to mitigate cyber-security risks? Can it model the type of awareness required from the human operators?	- Cyber-security risk assessment of AI components - Describing applicable methods and defence mechanisms against attacks - Developing guidelines for secure AI development in ATC systems - Conducting literature reviews and workshops regarding cybersecurity - Examining encryption protocols for secure data transmission/storage	The methodology examines how the system can support users in identifying and understanding cybersecurity threats, ensuring awareness of potential risks in AI-assisted ATC operations.

Table 4. Connection between research questions and the project methodology

After the first phase of the project, which includes the AI Assistant application development, HITL experiments will be conducted in a simulated environment. Each experiment will consist of a traffic scenario specifically designed to address different operational feature. Furthermore, each scenario will include one group of ATCOs performing the scenario without ASA support, along with three additional groups operating with varying levels of ASA support. The group working without support will provide baseline results, establishing a performance benchmark against which the ASA supported conditions can be compared. The following aspects will be measured: task efficiency, decision-making speed, and relevant safety indicators (e.g., conflict detection and resolution times). This experiment will assess the system’s impact on:

- Capacity
- Cost efficiency
- Environment
- Human performance
- Operational efficiency
- Safety

The methodology will incorporate both qualitative and quantitative performance metrics. Quantitative metrics include task completion times, accuracy of AI predictions, and relevant key performance indicators. Qualitative performance metrics refer to ATCOs feedback that will provide valuable information regarding system usability, workload impacts and level of trust in the ASA system.

The final phase of the research will compare the outcomes of the baseline and AI-supported experiments. This analysis will assess the ASA system's effectiveness in reducing ATCO workload, improvement in decision-making under complex traffic conditions and ability to enhance operational safety and efficiency.

The AWARE project's methodology is carefully designed to answer the key research questions and achieve the project's objectives. Each phase of the experimental approach is aligned with specific research questions. Research Questions define the areas of focus (e.g., automation acceptance, workload reduction). Objectives translate these into practical goals (e.g., enhancing decision-making, reducing workload). The methodology ensures that each research question is systematically addressed through simulations and trials, leading to measurable outcomes.

3.3 Key R&I needs

The research and innovation needs addressed by this exploratory research plan focus on enhancing the capabilities of the Artificial Situational Awareness system, improving Human-Machine Interface design, and deepening our understanding of Air Traffic Control Officers' cognitive processes. These research questions explore essential areas for development, aiming to provide practical solutions for ATCOs and improve overall safety, efficiency, and collaboration between human operators and AI systems.

ATCO Intent Analysis

Understanding ATCO intent is a crucial component of effective AI support. The research questions in this area investigate whether the AI Assistant Application can accurately recognise ATCOs' intentions by analysing gaze point sequences and HMI interactions in real-time. The AI's ability to compare this data with a comprehensive list of en-route air traffic tasks is key to ensuring an adequate level of accuracy. Furthermore, there is a focus on whether the AI system can correctly identify when an ATCO is experiencing increased or excessive mental workload, particularly when compared to traditional subjective workload assessments and implicit performance measures.

ASA System

The ASA system's ability to perform in real-time within air traffic control environments is a central focus of this research. Key questions explore whether the system developed in the AISA project can be improved by leveraging advanced technologies such as SPARQL, SHACL-rules, or RDFox (a high-performance knowledge graph reasoner) to enable faster processing speeds in simulations. Additionally, there is a need to determine whether all monitoring tasks outlined in the AISA Concept of Operations can be fully automated while maintaining acceptable accuracy levels. The integration of the ASA system with simulation software is also scrutinized to assess whether it can maintain adequate processing speeds to ensure real-time performance.

HMI and Attention Tracking

For the HMI, research is focused on improving the AI Assistant Application's ability to track ATCOs' visual attention and interactions with the HMI. A critical question is whether the AI Assistant can reliably identify the ATCOs' "object of preoccupation" by combining visual attention tracking with real-time data from HMI interactions. The reliability of eye-tracking data is another area of concern, particularly whether temporary disruptions, such as misidentification of gaze or calibration problems, can be compensated for by HMI interaction data, and vice versa. These research needs are essential to ensure that the AI system can accurately track ATCOs' focus and attention, a key aspect of situational awareness.

ATCO Supporting Tools

Another critical research and innovation (R&I) need revolves around the support provided by the AI Assistant Application and its impact on ATCOs. The research will evaluate whether the support offered by the AI is acceptable to ATCOs in practice and assess its effects on ATCO workload. Furthermore, the research will explore how the AI system influences operational metrics such as flight efficiency, cost efficiency, and, importantly, safety. These insights will help tailor the AI system's functionalities to provide meaningful assistance without overwhelming ATCOs or introducing new operational risks.

Modes of Human-Machine Collaboration

The exploration of different modes of collaboration between humans and AI systems is central to this research. Key questions address how ATCOs prefer to be supported, especially under conditions of increased traffic volume and high complexity. Research will determine if higher levels of automation are acceptable in these scenarios and investigate how design solutions can intuitively support ATCOs in assessing the appropriateness of AI assistance without adding to their cognitive workload. Additionally, the amount of data required from eye-tracking and HMI interactions to model ATCO cognitive processes effectively will be assessed, helping to refine the AI's ability to predict ATCO intentions accurately.

3.4 Estimated performance contributions

The AWARE project is expected to deliver significant performance improvements across several key areas of ATM. The ASA system being developed will support ATCOs by enhancing situational awareness, improving safety, and optimising operations.

In terms of environmental impact (ENV), the ASA system will contribute moderately by reducing fuel consumption and emissions through better traffic management and minimized delays. Although not a primary focus of the project, this contribution is important for achieving greener operations, with expected improvements in fuel efficiency and reduced CO_2 emissions. Therefore, a reduction in actual average CO_2 emission per flight (ENV1) of >5% is expected.

Capacity (CAP) will see a major boost from the ASA system, as it enables ATCOs to handle more flights per hour without compromising safety. By automating routine tasks and providing real-time traffic predictions, the system allows controllers to manage higher traffic volumes and adapt to dynamic airspace configurations, including mixed-traffic scenarios with unmanned aerial systems (UAS). Therefore, once the Solution is implemented, a significant increase (>25%) in en-route throughput, in challenging airspace, per unit time (CAP2) is expected.

The ASA system will have a high impact on operational efficiency (OPS), enhancing decision-making by reducing delays and optimising traffic flow. The ability to predict traffic conditions and offer automated suggestions will help ATCOs manage airspace more efficiently, resulting in reduced delays, improved adherence to flight schedules, and more effective use of airspace resources. Thus, a decrease in actual average fuel burn per flight (FEFF1) is expected.

In terms of security (SEC), the ASA system does not directly target security improvements but indirectly contributes by enhancing the overall reliability and robustness of ATM operations, ensuring that air traffic is managed safely and predictably.

The contribution to safety (SAF) is expected to be high, as the ASA system automates monitoring tasks, reduces human errors, and enhances ATCOs' situational awareness, thus providing the ATCO with an additional safety net. By providing early conflict detection and decision support, the system is anticipated to significantly reduce safety incidents and improve conflict resolution times. Even though there is no performance indicator related to activation of safety nets, we think that this performance indicator should be introduced and measured. Keeping the ATCO in the loop would prevent activation of safety nets (e.g. STCA) so a reduction of those alerts by 20% expected. It is anticipated that the frequency of events per flight hour associated with the following Key Performance Indicators (KPIs) will decrease: SAF1.3, SAF1.4.3, SAF1.5.1, SAF1.5.2, SAF1.5.3, SAF1.5.4, SAF1.8.1, SAF1.8.2, SAF1.10.1, SAF1.10.2, SAF1.10.3, and SAF1.10.4. This reduction aligns with the performance expectations outlined in the SESAR Performance Framework, which targets improvements in safety and operational efficiency through enhanced situational awareness and improved air traffic management technologies.

3.5 Initial and exit maturity levels

Project/ Proposed SESAR solution(s) ID	Proposed SESAR solution(s) title	Initial maturity level	Exit maturity level	Reused validation material from past R&I Initiatives
ASA	Artificial Situational Awareness System	TRL1	TRL2	AISA

Table 5. Maturity levels table

4 Exploratory research plan

4.1 Exploratory research plan approach

The exploratory research plan is designed to progress the ASA system from its current Technology Readiness Level (TRL1) to the exit maturity level (TRL2), addressing R&I needs. The approach focuses on four experiments with different traffic load and complexity. ATCOs performing the experiment will be divided into four groups, one group performing without ASA support and three groups receiving varying level of ASA system support. These experiments are essential for validating the ASA system's maturity. This project builds on the work done in the AISA project, where the ASA system was initially developed, and the approach is based on the entire project scope, not just these two experiments. The plan consists of two primary phases: development and validation.

In AISA, human-in-the-loop simulations demonstrated that the system's situational awareness was comparable to that of a human operator, although the project revealed limitations in human-machine interaction, particularly in understanding ATCO intent and the lack of a flexible HMI for real-time collaboration. In AWARE, the ASA system will be expanded with new methodologies to overcome these limitations. The system will integrate advanced intent recognition techniques and a more sophisticated HMI to improve real-time collaboration between ATCOs and the AI system. The development will involve several phases of simulation and validation, focusing on expanding automation capabilities, enhancing human-AI teaming, and refining decision-making processes. Additionally, the ASA system will evolve to better predict ATCO intentions and provide more intuitive support, all while ensuring that the system meets the operational and safety standards required for higher levels of maturity in real-world environments.

The development of the ASA system builds on the foundation established by the AISA project, where the initial solution achieved TRL1. In the AISA project, the ASA system underwent a series of HITL simulations to validate its core capabilities. The primary methodology involved leveraging a knowledge graph, which was built using aeronautical data models such as AIXM and FIXM, to automate certain monitoring tasks typically handled by ATCOs. The system's ability to generate insights about air traffic situations and automate routine tasks was tested against real-time scenarios in the en-route phase of ATM. While the system proved capable of maintaining situational awareness similar to a human operator, it faced significant challenges in interacting effectively with ATCOs. Specifically, the AISA system struggled to understand ATCO intent, provide timely support, and adjust its level of assistance dynamically. The maturity gate assessment at the conclusion of AISA validated the system's potential, allowing it to reach TRL1. However, the assessment also highlighted key areas needing improvement, particularly in enhancing human-machine collaboration and refining the HMI.

Validation of the ASA system is crucial to ensure that it meets the operational and safety requirements needed for deployment in real-world environments. This phase includes four experiments designed to test the system under medium to high traffic loads and increasing levels of complexity (Figure 2). Each experiment represents the complete set of methodologies used to collect the data required to address its associated research questions. Under this definition, an experiment encompasses the human-in-the-loop (HITL) exercises as well as all objective and subjective data collection methods. HITL exercises are characterized by their operational context, including the working environment, airspace, procedures, and traffic scenarios. These exercises incorporate specific traffic scenarios that are derived from real-world operations but may be intentionally modified to ensure adequate representation of tasks relevant to the evaluation of human-AI teaming methods.

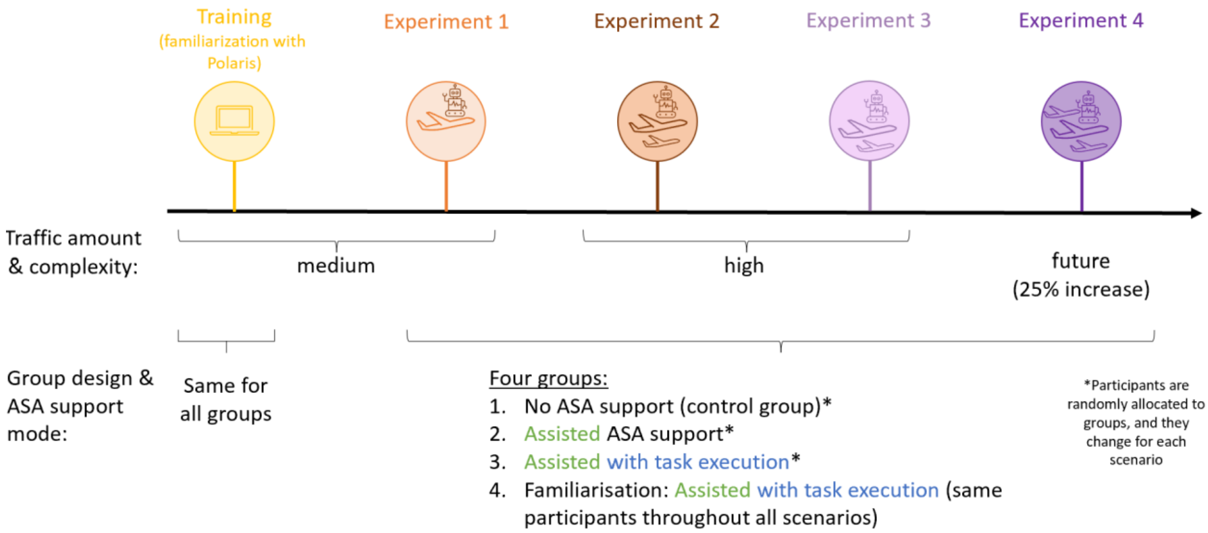

Figure 2 Validation Experiments Set-up

The first experiment, conducted under medium traffic load and complexity, will evaluate how effectively the ASA system monitors traffic, supports ATCOs and automates selected tasks to reduce workload while maintaining safety. This scenario will help determine whether the proposed level of support is necessary and effective under medium complexity conditions.

The second and third experiments will focus on Human-AI teaming under conditions of high traffic load and complexity. They will assess how well the system collaborates with ATCOs, enhances situational awareness, supports decision-making, and manages workload. The second experiment will be based on traffic scenario designed to test the conflict detection and resolution workflow while the third experiment will be based on traffic scenario focused on traffic interacting with the military restricted areas.

The fourth experiment will simulate a future operational environment with a 25% increase in a peak sector occupancy. The final experiment will test if the ASA system can operate effectively and ensure robustness under extreme traffic conditions.

For these experiments, ATCOs will be randomly assigned to one of four groups and:

- Group 1: No ASA support (baseline).
- Group 2: Assisted ASA support for decision-making. The ASA system will provide limited support, assisting with tasks the ATCO is actively performing (2A level of support).
- Group 3: Assisted ASA support with task execution. The ASA system assists with active tasks and autonomously handles certain tasks in the background (2B level of support).
- Group 4 (Familiarisation group) Assisted ASA support provided throughout all experiments. Unlike the other groups, participants in this group will maintain consistent assisted ASA support with task execution during Experiments 2, 3, and 4.

EUROPEAN PARTNERSHIP
Co-funded by the European Union

Group 1 performs each scenario without ASA support, providing a consistent control benchmark for evaluating changes in performance, workload, and situational awareness. Groups 2–4 will validate the ASA system's ability to monitor operations, support active tasks, and automate selected functions safely and accurately. ATCOs in Groups 1–3 are reassigned randomly for each experiment, while Group 4 remains fixed to assess how increasing experience with the system influences performance.

This structure enables meaningful comparisons while ensuring that findings are robust and representative of diverse operational contexts. Separating participants into groups allows for a systematic examination of the ASA system's impact under distinct conditions. By isolating specific levels of support and varying the degree of task execution, it is possible to identify how different modes of ASA assistance influence performance, workload, and situational awareness.

The results of the four experiments will be compared, allowing for the identification of performance improvements and any remaining challenges in system integration. Feedback from these experiments will inform further refinements of the ASA system, ensuring that the solution is adapted to meet the operational needs of ATCOs and the performance expectations of the SESAR program. The interdependencies between these experiments are key to ensuring a thorough validation of the ASA system.

Insights from visual attention tracking and HMI interactions will be crucial in refining the ASA system's real-time decision-making capabilities and enhancing human-AI collaboration. The ASA system's use of SPARQL to obtain data from the knowledge graph, and RDFox for knowledge graph reasoning will be key to enhancing its real-time decision-making capabilities. Simulations will measure the system's response time and its ability to process and reason about complex air traffic scenarios in near real-time.

Throughout the research plan, feedback from earlier project phases and from the content integration process will be integrated to refine the experiments and address any unresolved challenges. For example, any gaps identified in the system's performance in earlier simulations, such as delays in automation or challenges in interpreting ATCO intent, will be addressed in subsequent validation cycles. This iterative process ensures that the system evolves in response to real-world operational feedback.

In terms of performance measurement, each validation exercise will contribute to achieving KPAs and KPIs outlined in the SESAR framework. The key KPAs targeted in the validation experiments include Capacity, Cost-Efficiency, Operational Efficiency, Safety, Environment and Human Performance, with corresponding Key Performance Indicators (KPIs) designed to measure the system's impact in these areas.

The following table lists the human factor variables that will be measured in the experiments. Aside from these, several key variables of the simulations and traffic will be noted. These include, for example, the number of aircraft in the sector, trajectories, horizontal and vertical speed etc.

Variables	HF Topic	Apparatus	Time of measurement
Pupil diameter	workload	eye tracking glasses or fusion bar	during simulation
Scanning pattern	attention, SA	eye tracking glasses or fusion bar	during simulation

EUROPEAN PARTNERSHIP
Co-funded by the European Union

Variables	HF Topic	Apparatus	Time of measurement
Fixation frequency	attention, SA	eye tracking glasses or fusion bar	during simulation
Fixation duration	attention, SA	eye tracking glasses or fusion bar	during simulation
Time for detection of relevant information	attention, SA	eye tracking glasses or fusion bar	during simulation
Number of revisits of AOIs	attention, SA	eye tracking glasses or fusion bar	during simulation
HR	workload	Biometrical sensor (Meditech/Life)	during simulation
HRV	workload	Biometrical sensor (Meditech/Life)	during simulation
Respiration rate	workload	Biometrical sensor (Meditech/Life)	during simulation
Skin conductance	workload	Biometrical sensor (Meditech/Life)	during simulation
Skin temperature	workload	Biometrical sensor (Meditech/Life)	during simulation
ECG	workload	Biometrical sensor (Meditech/Life)	during simulation
HEG	workload	Biometrical sensor (Meditech/Life)	during simulation
Gender	-	Questionnaire	before simulation
Experience	performance	Questionnaire	before simulation
Age	-	Questionnaire	before simulation
Job function	-	Questionnaire	before simulation
Amount of sleep / fatigue	fatigue	Questionnaire	before simulation
Level of stress	fatigue	Questionnaire	before & after simulation
Level of sleep/fatigue (how they feel at the moment)	fatigue	Questionnaire (Karolinska Sleepiness Scale)	before & after simulation
Sleep quality	fatigue	Questionnaire	before simulation
Familiarity with airspace, simulation software	performance	Questionnaire	before simulation
Time of day	fatigue	Questionnaire	before simulation
Attitude towards AI	-	Questionnaire	after simulation
Usability of AI support	-	Questionnaire	after simulation

EUROPEAN PARTNERSHIP
Co-funded by the European Union

Variables	HF Topic	Apparatus	Time of measurement
Trust in AI support	-	Questionnaire	after simulation
Interface rating	-	Questionnaire	after simulation
HAT (Human Automation teaming) rating	-	Questionnaire	after simulation
Task load	workload	Simulation data	during simulation
Traffic load	workload	Simulation data	during simulation
Traffic complexity	workload	Simulation data	during simulation
Traffic handling strategy	performance	activity coding	during simulation

Table 6. HF variables measured in the experiments

For the Capacity KPA, the experiment will focus on en-route throughput (CAP1, [5]) as the key KPI. This KPI will measure the number of aircraft processed per unit time in en-route sectors, capturing how effectively the ASA system supports ATCOs in managing air traffic, especially under high-demand conditions. To calculate this KPI, we will measure % and the total number of movements, per volume of En-Route airspace per hour for specific traffic mix and density (Very High, High and Medium Complexity) at peak demand hours during both the baseline and Human-AI teaming scenarios.

In terms of Cost-Efficiency, the relevant KPI is flights per ATCO-hour on duty (CEF2, [5]). This will evaluate the efficiency gains provided by the ASA system, particularly its ability to reduce the workload of ATCOs by automating certain tasks. To calculate this KPI, data will be collected on variables such as the count of flights handled divided by the number of ATCO-Hours applied by ATCOs on duty. The difference in ATCO workload between the baseline and Human-AI teaming scenarios will be analysed to demonstrate improvements in cost-efficiency.

For Operational Efficiency, the experiment will focus on actual average fuel burn per flight (FEFF1, [5]) as the key KPI. The ASA system's ability to optimise flight routes and improve fuel efficiency will be evaluated using objective data on the total amount of actual fuel burnt divided by the number of flights. These variables will be measured during simulations to compare the fuel consumption between the baseline and Human-AI teaming scenarios, providing clear evidence of the system's impact on operational efficiency.

The Safety KPA will focus on reducing the total number of estimated accidents with ATM contribution per year (SAF1, [5]). This KPI will measure how effectively the ASA system enhances situational awareness and reduces the risk of safety-critical incidents. According to the SESAR Performance Framework document, this KPI is not measurable directly. Instead, it is determined by Safety practitioners via the AIM Models and validated with the support of numerous Safety PIs. Some of the relevant groups of PIs included are mid-air collisions – en-route.

For the Environment KPA, the experiment will focus on the actual average CO_2 emission per flight (ENV1, [5]) as the key KPI. This KPI measures the environmental impact of flights by calculating the total CO_2 emissions generated per flight and assessing how effectively the ASA system contributes to reducing this footprint. To calculate this KPI, we will measure the total amount of fuel burnt and apply a CO_2 emission index of 3.15 (CO_2 per kilogram of fuel) to convert the fuel burn into CO_2 emissions. The resulting value will then be divided by the number of flights to determine the average CO_2 emission per flight.

Co-funded by the European Union

Finally, for Human Performance, the relevant KPI is the suitability of the technical system in supporting the tasks of human actors (HP2, [5]). This will be assessed by evaluating the adequacy of task allocation between humans and the system, ensuring that the division of tasks minimizes human workload while optimising system support. It also assesses the timeliness and accuracy of system responses, determining whether the system provides the necessary information at the right time and with the required precision to help human actors perform their tasks effectively. Additionally, the evaluation considers the design of the HMI, ensuring it follows human factors principles to reduce errors, enhance usability, and improve situational awareness. Data will be collected through subjective feedback on the usability of AI support, attitude towards AI, and interface rating. This data will help assess how effectively the AI system integrates with the human operator's workflow and reduces cognitive load.

This data is shown in the following table, which presents the way each KPAs' relevant KPI is calculated, its unit of measurement and the joint PIs.

KPA	KPI	UNIT	CALCULATION	PIs
CAPACITY	En-route throughput, in challenging airspace, per unit time (CAP2)	Relative change of movements (% and number of movements)	% and the total number of movements, per volume of En-Route airspace per hour for specific traffic mix and density at peak demand hours	KPIs are calculated directly without intermediate PIs
COST-EFFICIENCY	Flights per ATCO-Hour on duty (CEF2)	/	Count of Flights handled divided by the number of ATCO-Hours applied by ATCOs on duty	KPIs are calculated directly without intermediate PIs
ENVIRONMENT	/	Kg CO_2 per flight	Amount of fuel burnt x 3.15 (CO_2 emission index) divided by the number of flights	**ENV1** Actual Average CO_2 Emission per flight
HUMAN PERFORMANCE	/	/	For the purposes of this solution, ATCOs qualitative assessment questionnaire answers will be used	Suitability of the technical system in supporting human tasks (HP2)
OPERATIONAL EFFICIENCY	Actual Average fuel burnt per flight (FEFF1)	Kg fuel per flight	Total amount of actual fuel burnt divided by the number of flights	KPIs are calculated directly without intermediate PIs

KPA	KPI	UNIT	CALCULATION	PIs
SAFETY	Total number of estimated accidents with ATM contribution (SAF1)	Yearly count of events	Usually not directly measurable. Usually determined by Safety practitioners via the AIM, validated by Safety PIs. For the purposes of this solution, a count of each specific event will be made.	**SAF1.3** Imminent Collision **SAF1.4.3** ATC induced Imminent Infringement **SAF1.5.1** Conflict due to Level Bust (Crew/Aircraft induced) **SAF1.5.2** Conflict due to Lateral Deviation (Crew/Aircraft induced) **SAF1.5.3** Conflict due to Speed Deviation (Crew/Aircraft induced) **SAF1.5.4** Conflict due to V. rate Deviation (Crew/Aircraft induced) **SAF1.8.1** Unit/Sector (ATC) Induced Tactical conflict **SAF1.10.1** Level Bust **SAF1.10.2** Aircraft Lateral Deviation **SAF1.10.3** Aircraft Speed Deviation **SAF1.10.4** Aircraft V. rate Deviation

Table 7. KPAs and relevant KPIs/PIs, [5]

In conclusion, the interdependencies between the baseline group and Human-AI teaming groups are essential for a comprehensive validation of the ASA system. Validation experiments will ensure that the system evolves to meet operational needs, contributing to enhanced ATCO performance and alignment with SESAR's overall objectives.

4.2 Stakeholders’ expectations and involvement

The table below outlines the roles, responsibilities, and expected contributions of each stakeholder group in the AWARE project. Each group provides support for the development, validation, and integration of the ASA system.

Stakeholder	Involvement	Why it matters to the stakeholder
ATCOs	ATCOs are crucial to the AWARE project, contributing to the development and validation of the ASA system. They provide insights into human-machine interaction, helping to assess workload, intent recognition, and safety. IFATCA, as a project partner, ensures that ATCOs' perspectives and operational needs are integrated into the system design.	They will be supported in their work, enabling them to avoid extremely stressful situations which occur when complexity rises unexpectedly. They will not suffer from complacency or out-of-the-loop effect because they will be actively involved in their work at all times. Their role will not change to supervisory, and their skills will be honed and appreciated.
SUPs	ATCO Supervisors (SUPs) are involved in overseeing the integration of AI-based tools such as the ASA system to ensure it aligns with operational safety, effectiveness, and efficiency. Their role includes supervising ATCOs and managing system use during high-complexity scenarios, ensuring smooth coordination between human operators and automated systems.	The AWARE project directly impacts SUPs by introducing advanced tools that assist in maintaining situational awareness and decision-making under pressure. The ASA system aims to support human-machine collaboration, offering SUPs a comprehensive overview of air traffic operations and improving ATCOs’ ability to manage high-stress situations. This matters to SUPs because it can enhance their oversight role, reduce workload, and improve safety by automating routine monitoring tasks and detecting potential issues early. Furthermore, as air traffic continues to grow, SUPs will benefit from the improved ATCO capacity to handle more complex traffic scenarios while ensuring safety and efficiency.
ATSEPs	Air Traffic Safety Electronics Personnel (ATSEPs) are involved in the AWARE project by contributing to the integration and maintenance of the ASA system. They ensure the seamless interoperability of the system with existing ATM infrastructure, focusing on its technical reliability, cybersecurity, and operational safety.	The AWARE project is crucial for ATSEPs because it introduces advanced automation tools that enhance the safety and efficiency of ATM systems. As automation increases, ATSEPs are responsible for maintaining the technical infrastructure supporting these systems. The ASA system's interoperability with other roles and systems directly impacts ATSEPs by requiring them to ensure system reliability and cybersecurity while addressing new technical challenges associated with high automation. The project offers opportunities for ATSEPs to improve their expertise in managing complex, interconnected ATM systems and staying ahead of emerging threats like cyber-attacks.

Stakeholder	Involvement	Why it matters to the stakeholder
ANSPs	ANSPs are key stakeholders in AWARE (LFV and UkSATSE are partners), contributing to the operational validation of the ASA system using real-time data. They assess its impact on workload and safety, supporting human-machine collaboration. In the SESAR framework, ANSPs help ensure seamless integration of ASA into existing air traffic management systems.	Supporting ATCOs in performing their tasks, even with system being able to provide action support to only the bottom 50% of tasks in terms of complexity, will allow a 100% increase in capacity. Additional safety net provided by the automation of all monitoring tasks (80% of which were already automated in the previous project) will improve safety. Previous work showed that in over 1400 instances tested, artificial situational awareness did not have instances of degraded situational awareness (compared to dozens for humans). Keeping humans as part of the highly complex sociotechnical system such as ATM/ATC also provides safety benefits due to unmatched resourcefulness and resilience of the human.
Airspace Users	Airspace Users are involved in the AWARE project by providing operational feedback to ensure the ASA system meets their needs for efficient flight operations. They contribute to the validation of trajectory-based operations, helping assess the system's impact on performance, fuel efficiency, and safety in real-world conditions.	Higher capacity enables lower costs of the ANSP's operations which should translate to lower route charges. We estimate that the AWARE solution will generate a 5% - 25% capacity increase in en-route capacity by enabling ATCOs to accommodate more traffic. This would mean a 112 – 940 million euro benefit for ANSPs annually, starting gradually from 2031 and onwards [7].
Society	Society is involved in the AWARE project through its focus on improving air traffic safety, efficiency, and environmental sustainability. By enhancing situational awareness and reducing air traffic delays, the project aims to benefit the public by contributing to safer, greener, and more reliable air transportation systems.	Direct supporting actions, even the so-called 'housekeeping' tasks taken autonomously by the system, could provide environmental benefit by providing aircraft with direct routes which might not be approved otherwise in the instances of ATCO's high workload. Also, increased capacity would allow aircraft to fly their optimal routes instead of avoiding congested airspace.
ATM Research Community	The ATM research community plays a key role in the AWARE project by contributing expertise in artificial intelligence, human-machine interaction, and operational performance. They collaborate on system development, validation, and performance assessments, ensuring the ASA system aligns with cutting-edge research and supports future air traffic management needs.	This method could benefit all ATM researchers by allowing them to explore different facets of automation issues that were previously left unexplored. Also, our work will provide guidelines (best practices) for researchers using non-invasive gaze tracking for real-time decision-making in ATM. Benefits, or lack thereof, of using ATCO intent to enable human-machine collaboration will be available for all researchers to evaluate and use (or avoid using if benefits do not materialise).

Table 8. Stakeholders' expectations and involvement

4.3 Validation objectives

The following validation objectives are defined to support the exploratory research plan, focusing on addressing key research and innovation needs aligned with the DES performance framework. These objectives are structured to ensure traceability to other project deliverables and provide clear success criteria, aligned with SESAR guidelines and validation requirements. The success of each objective will be measured through KPIs and metrics in line with the latest DES performance framework.

The following table describes relevant KPAs and their Detailed Validation Objective, each KPA has a defined Success Criteria.

KPA	DETAILED VALIDATION OBJECTIVE	SUCCESS CRITERIA
CAPACITY	Assess the increase in capacity due to the AWARE solution.	AWARE leads to increased Capacity.
COST EFFICIENCY	Assess the improvement in Cost Efficiency due to the AWARE solution.	AWARE leads to increased Cost Efficiency.
ENVIRONMENT	Assess the improvement in fuel efficiency due to AWARE solution.	AWARE leads to reduction in fuel consumption.
HUMAN PERFORMANCE	Assess the impact on the required HP levels due to AWARE solution.	AWARE leads to lower workload levels.
OPERATIONAL EFFICIENCY	Assess the improvement on the required Operational Efficiency due to the AWARE solution.	AWARE leads to increased Operational Efficiency.
SAFETY	Assess the improvement of Safety due to the AWARE solution.	AWARE leads to fewer conflicts, resulting in increased safety.

Table 9. Relevant KPAs and their Detailed Validation Objective, [1]

Key Performance Indicators, Performance Indicators, and data collection methods to meet each success criteria and detailed validation objectives are listed in the table below.

DETAILED VALIDATION OBJECITVE	KPIs	PIs	DATA COLLECTION METHOD
		CAPACITY	
Assess the increase in capacity due to the AWARE solution.	**CAP2**-En-route throughput, in challenging airspace, per unit time		Counting flights in HITL simulation task scenarios.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

DETAILED VALIDATION OBJECITVE	KPIs	PIs	DATA COLLECTION METHOD
			HF variables measured in the experiments and questionnaire (table 6)
COST EFFICIENCY			
Assess the improvement in Cost Efficiency due to the AWARE solution.	**CEF2**-Flights per ATCO-Hour on duty		Data collected from the HITL scenarios HF variables measured in the experiments and questionnaire (table 6)
ENVIRONMENT			
Assess the improvement in fuel efficiency due to AWARE solution.	**ENV1**- Actual Average CO_2 Emission per flight		Data collected from the HITL scenarios
HUMAN PERFORMANCE			
Assess the impact on the required HP levels due to AWARE solution	**HP2**- Suitability of the technical system in supporting human tasks		Data collected from the HITL scenarios HF variables measured in the experiments and questionnaire (table 6)
OPERATIONAL EFFICIENCY			
Assess the improvement on the required Operational Efficiency due to the AWARE solution.	**FEFF1**-Actual Average fuel burnt per flight		Data collected from the HITL scenarios HF variables measured in the experiments and questionnaire (table 6)
SAFETY			
Assess the improvement of Safety due to the AWARE solution.	**SAF1**-Total number of estimated accidents with ATM contribution	**SAF1.3** Imminent Collision	Data collected from the HITL scenarios HF variables measured in the experiments and questionnaire (**Table 6**)

EUROPEAN PARTNERSHIP
Co-funded by the European Union

DETAILED VALIDATION OBJECITVE	KPIs	PIs	DATA COLLECTION METHOD
		SAF1.4.3 ATC induced Imminent Infringement	Data collected from the HITL scenarios
		SAF1.5.1 Conflict due to Level Bust (Crew/Aircraft induced)	Data collected from the HITL scenarios
		SAF1.5.2 Conflict due to Lateral Deviation (Crew/Aircraft induced)	Data collected from the HITL scenarios
		SAF1.5.3 Conflict due to Speed Deviation (Crew/Aircraft induced)	Data collected from the HITL scenarios
		SAF1.5.4 Conflict due to V. rate Deviation (Crew/Aircraft induced)	Data collected from the HITL scenarios
		SAF1.8.1 Unit/Sector (ATC) Induced Tactical conflict	Data collected from the HITL scenarios
		SAF1.10.1 Level Bust	Data collected from the HITL scenarios
		SAF1.10.2 Aircraft Lateral Deviation	Data collected from the HITL scenarios
		SAF1.10.3 Aircraft Speed Deviation	Data collected from the HITL scenarios
		SAF1.10.4 Aircraft V. rate Deviation	Data collected from the HITL scenarios

Table 10. KPIs, PIs and data collection methods, [1], [5]

To assess KPIs and meet success criteria, variables for each KPA and specific validation objectives are presented in the table below. These variables may be refined as the project develops.

KPA	DETAILED VALIDATION OBJECTIVE	SUCCESS CRITERIA	VARIABLES		
			DEPENDENT	INDEPENDENT	CONTROL
CAPACITY	Assess the increase in capacity due to the AWARE solution.	AWARE leads to increased Capacity.	Traffic movement	Use of AWARE (without AWARE; with AWARE)	Airspace structure Number of aircraft in the sector Complexity

Co-funded by the European Union

KPA	DETAILED VALIDATION OBJECTIVE	SUCCESS CRITERIA	VARIABLES		
			DEPENDENT	INDEPENDENT	CONTROL
COST-EFFICIENCY	Assess the improvement in Cost Efficiency due to the AWARE solution.	AWARE leads to increased Cost Efficiency.	ATCO hours	Use of AWARE (without AWARE; with AWARE)	Airspace structure
ENVIRONMENT	Assess the improvement in fuel efficiency due to AWARE solution.	AWARE leads to reduction in fuel consumption.	Amount of fuel burnt	Use of AWARE (without AWARE; with AWARE)	Airspace structure
HUMAN PERFORMANCE	Assess the impact on the required HP levels due to AWARE solution.	AWARE leads to lower workload levels.	ATCO Workload	Use of AWARE (without AWARE; with AWARE)	Years of experience Type of Scenarios Airspace structure
OPERATIONAL EFFICIENCY	Assess the improvement on the required Operational Efficiency due to the AWARE solution.	AWARE leads to increased Operational Efficiency.	Fuel consumptions	Use of AWARE (without AWARE; with AWARE)	Airspace structure
SAFETY	Assess the improvement of Safety due to the AWARE solution.	AWARE leads to fewer conflicts, resulting in increased safety.	Yearly count of events Near Mid-Air Collisions Imminent Collisions	Use of AWARE (without AWARE; with AWARE)	Airspace structure Complexity Type of Scenario

Table 11. Variables for each KPA and specific validation objectives, [1], [5]

To ensure consistency, section 5 will further detail how these validation objectives are applied across various validation exercises, including defining specific success criteria for each sub-operating environment. Where different TRLs are addressed, the exploratory research plan identifies which validation objectives apply to each TRL, ensuring a coherent validation approach across all phases of the research.

4.4 Validation assumptions

Assumption ID	Assumption title	Assumption description	Justification	Impact Assessment
F-0005	En-route fuel consumption contribution	The en-route phase of flight contributes approximately 64.61% of total fuel consumption.	While the general assumption values en-route fuel consumption at 64.61%, our project will assume a higher contribution since we are primarily considering en-route flights. Given our small sample size, we will use actual data without extensive averaging.	A direct impact is predicted
M-0002	Peak traffic per day, very high and high complexity airspace	Peak traffic in very high and high complexity en-route airspace occurs for approximately 5 hours per day.	Although peak traffic typically lasts around 5 hours daily, our simulations focus on shorter durations. For the scope of en-route operations and a smaller sample size, we will use actual data captured during peak periods without averaging values.	A direct impact on workload and performance evaluation is predicted.
M-0004	Peak traffic per medium complexity en-route airspace	Peak traffic in medium complexity en-route airspace occurs for approximately 2 hours per day.	While medium complexity airspace experiences peak traffic for around 2 hours daily, our simulations will focus on shorter periods. We will rely on real-time data from these simulations rather than averaging values.	A direct impact on workload and performance evaluation is predicted.

Table 12 Validation assumptions overview, [6]

4.5 Validation exercises list

To answer high level research questions, five validation exercises are planned:

1. Workshop for gathering feedback on ASA system functionality.
2. HITL simulation with medium traffic levels and varying ASA support (no support, 2A or 2B level of support).
3. HITL simulation with high traffic levels and varying ASA support (no support, 2A or 2B level of support). This simulation is based on scenario designed to test the conflict detection and resolution workflow.

4. HITL simulation with high traffic levels and varying ASA support (no support, 2A or 2B level of support). This simulation is based on scenario designed to focus on traffic interacting with the military restricted areas.

5. HITL simulation with future traffic levels and varying ASA support (no support, 2A or 2B level of support).

Table 13 lists the HITL simulation exercises that will be performed and their traffic and support levels. For exercises TVAL.02.1, TVAL.02.2, TVAL.02.3, and TVAL.02.4, all participating ATCOs will be divided into four groups, each with a different level of ASA support:

- Group 1: No ASA support
- Group 2: 2A support level
- Group 3: 2B support level
- Group 4: 2B support level

At level 2A ASA system will support ATCOs in executing their tasks and will not execute any tasks autonomously. At level 2B, ASA system will support ATCOs in executing their tasks while also executing tasks autonomously.

The difference between Groups 3 and 4 is that ATCOs in Group 4 remain in the same group for all four exercises (TVAL.02.1, TVAL.02.2, TVAL.02.3, and TVAL.02.4), while ATCOs in the other groups rotate between groups for each exercise.

Table 14, Table 15, Table 16, Table 17, and Table 18 contain information about the validation exercises.

Exercise / Traffic level	TVAL.02.1	TVAL.02.2	TVAL.02.3	TVAL.02.4
Medium	Baseline + 3 groups of varying ASA support	-	-	-
High	-	Baseline + 3 groups of varying ASA support	Baseline + 3 groups of varying ASA support	-
Future	-	-	-	Baseline + 3 groups of varying ASA support

Table 13. List of HITL simulation validation exercises

Identifier	TVAL.01.1-AWARE-ASA-TRL2
Title	Workshop for gathering feedback on ASA system functionality
Description	This workshop represents the first test of both the system's functionality and its HMI. A limited number of ATCOs from various ANSPs will receive a guided interactive demonstration of ASA system and will provide feedback through interviews, questionnaires and plenary discussion. The objective is to give ATCOs first-hand experience with how the ASA system assist them during en-route operations. In addition, ATCO feedback will be used to refine the final ASA system functionalities, establish general rules and procedures, and further develop the HMI features.
KPA/TA addressed	Human performance
Addressed expected performance contribution(s)	Improvement in all HP indicators.
Maturity level	TRL2
Use cases	No use case identified. This is an exercise is a workshop for gathering feedback on ASA system functionality. ATCOs will not use the ASA system in simulation exercises.
Validation technique	Workshop with ATCOs
Validation platform	Polaris augmented with ASA system
Validation location	LFV
Start date	03/11/2025
End date	7/11/2025
Validation coordinator	ZHAW
Status	in progress
Dependencies	None
Linked Element Type	N/A
<SESAR Solution>	Artificial Situational Awareness System (ASA system)
<Project>	AWARE
<Sub-Operating Environment>	N/A
<Validation Objective>	Allow ATCOs to experience first-hand the way ASA system can support them during nominal en-route operations and gather feedback.

Table 14. Validation exercise TVAL.01.1-AWARE-ASA-TRL2 layout, [8]

Identifier	TVAL.02.1-AWARE-ASA-TRL2
Title	HITL simulation with medium traffic levels and varying ASA levels of support.
Description	This exercise will be done on Polaris ATM system with medium traffic levels (compared to peak time of peak day) and varying ASA levels of support. The planned duration is 25 minutes. The objective of this exercise is to gather data for the assessment of benefits of the ASA system. By using four groups with different levels of ASA support, the exercise aims to identify how each support level influences ATCO performance and other relevant KPAs under medium traffic complexity.
KPA/TA addressed	Capacity, cost-efficiency, environment, safety, human performance, operational efficiency
Addressed expected performance contribution(s)	No performance contribution is expected, this is a baseline exercise.
Maturity level	TRL2
Use cases	U.C.1, U.C.2, U.C.3, U.C.4 [8]
Validation technique	HITL simulations
Validation platform	Polaris ATM system
Validation location	LFV
Start date	01/02/2026
End date	15/02/2026
Validation coordinator	ZHAW
Status	Planned
Dependencies	None
Linked Element Type	N/A
<SESAR Solution>	Artificial Situational Awareness System (ASA system)
<Project>	AWARE
<Sub-Operating Environment>	N/A
<Validation Objective>	Gather data for the later assessment of the benefits of the ASA system.

Table 15. Validation exercise TVAL.02.1-AWARE-ASA-TRL2 layout, [8]

Identifier	TVAL.02.2-AWARE-ASA-TRL2
Title	HITL simulation with high traffic levels and varying ASA levels of support.
Description	This exercise will be done on Polaris ATM system augmented with ASA system, with high traffic levels (compared to peak time of peak day) and varying levels of ASA support. The planned duration is 20 minutes, with a primary focus on conflict detection and resolution. The objective of this exercise is to gather data for the assessment of benefits of the ASA system. By using four groups with different levels of ASA support, the exercise aims to identify how each support level influences ATCO performance and other relevant KPAs.
KPA/TA addressed	Capacity, cost-efficiency, environment, safety, human performance, operational efficiency
Addressed expected performance contribution(s)	Improvement in all KPAs is expected.
Maturity level	TRL2
Use cases	U.C.1, U.C.2, U.C.3, U.C.4 [8]
Validation technique	HITL simulations
Validation platform	Polaris ATM system augmented with ASA system
Validation location	LFV
Start date	01/02/2026
End date	15/02/2026
Validation coordinator	ZHAW
Status	Planned
Dependencies	None
Linked Element Type	N/A
<SESAR Solution>	Artificial Situational Awareness System (ASA system)
<Project>	AWARE
<Sub-Operating Environment>	N/A
<Validation Objective>	Gather data for the assessment of benefits of the ASA system.

Table 16. Validation exercise TVAL.02.2-AWARE-ASA-TRL2 layout, [8]

EUROPEAN PARTNERSHIP

Co-funded by the European Union

Identifier	TVAL.02.3-AWARE-ASA-TRL2
Title	HITL simulation with high traffic levels and varying ASA levels of support.
Description	This exercise will be done on Polaris ATM system augmented with ASA system, with high traffic levels (compared to peak time of peak day) and varying levels of ASA support. The planned duration is 20 minutes, with a primary focus on traffic interaction with military restricted areas. The objective of this exercise is to gather data for the assessment of benefits of the ASA system. By using four groups with different levels of ASA support, the exercise aims to identify how each support level influences ATCO performance and other relevant KPAs.
KPA/TA addressed	Capacity, cost-efficiency, environment, safety, human performance, operational efficiency
Addressed expected performance contribution(s)	Improvement in all KPAs is expected.
Maturity level	TRL2
Use cases	U.C.1, U.C.2, U.C.3, U.C.4 [8]
Validation technique	HITL simulations
Validation platform	Polaris ATM system augmented with ASA system
Validation location	LFV
Start date	01/02/2026
End date	15/02/2026
Validation coordinator	ZHAW
Status	Planned
Dependencies	None
Linked Element Type	N/A
<SESAR Solution>	Artificial Situational Awareness System (ASA system)
<Project>	AWARE
<Sub-Operating Environment>	N/A
<Validation Objective>	Gather data for the assessment of benefits of the ASA system.

Table 17. Validation exercise TVAL.02.3-AWARE-ASA-TRL2 layout, [8]

Identifier	TVAL.02.4-AWARE-ASA-TRL2
Title	HITL simulation with future traffic levels and varying ASA levels of support.
Description	This exercise will be done on Polaris ATM system augmented with ASA system, with future traffic levels and varying levels of ASA support. The planned duration is 20 minutes. Future traffic will be defined as peak traffic increased by 25%. The objective of this exercise is to gather data for the assessment of benefits of the ASA system in future traffic scenarios. By using four groups with different levels of ASA support, the exercise aims to identify how each support level influences ATCO performance and whether the ASA system can effectively assist ATCOs under future high complexity conditions.
KPA/TA addressed	Capacity, cost-efficiency, environment, safety, human performance, operational efficiency
Addressed expected performance contribution(s)	Improvement in all KPAs is expected.
Maturity level	TRL2
Use cases	U.C.1, U.C.2, U.C.3, U.C.4 [8]
Validation technique	HITL simulations
Validation platform	Polaris ATM system augmented with ASA system
Validation location	LFV
Start date	01/02/2026
End date	15/02/2026
Validation coordinator	ZHAW
Status	Planned
Dependencies	None
Linked Element Type	N/A
<SESAR Solution>	Artificial Situational Awareness System (ASA system)
<Project>	AWARE
<Sub-Operating Environment>	N/A
<Validation Objective>	Gather data for the assessment of benefits of the ASA system in presumed future scenarios with increased traffic levels.

Table 18. Validation exercise TVAL.02.4-AWARE-ASA-TRL2 layout, [8]

4.6 Validation exercises planning

The validation exercise TVAL.01.1-AWARE-ASA-TRL2 is planned to be performed as a workshop at LFV during November 2025 (M16), coordinated by ZHAW.

The validation exercise TVAL.02.1-AWARE-ASA-TRL2 is planned to be performed as simulation exercises at LFV from 1st of February 2026 to 15th of February 2026 (M20), also coordinated by ZHAW.

The validation exercise TVAL.02.2-AWARE-ASA-TRL2 is planned to be performed simultaneously with validation exercise #1 as simulation exercises at LFV from 1st of February 2026 to 15th of February 2026 (M21), also coordinated by ZHAW.

The validation exercise TVAL.02.3-AWARE-ASA-TRL2 is planned to be performed simultaneously with validation exercise #1 and #2 as simulation exercises at LFV from 1st of February 2026 to 15th of February 2026 (M21), also coordinated by ZHAW.

The validation exercise TVAL.02.4-AWARE-ASA-TRL2 is planned to be performed simultaneously with validation exercise #1, #2 and #3 as simulation exercises at LFV from 1st of February 2026 to 15th of February 2026 (M21), also coordinated by ZHAW.

4.7 Deviations with respect to the SESAR 3 JU project handbook

The AWARE project fully adheres to the SESAR 3 JU guidelines, with no noted deviations from the SESAR 3 Project Handbook, [3]. All deliverables, milestones, and processes are updated and aligned with the handbook's requirements, ensuring compliance and consistency with SESAR’s project execution framework and objectives.

5 Validation exercises

5.1 Validation exercise TVAL.01.1-AWARE-ASA-TRL2 plan

5.1.1 Validation exercise description and scope

The validation exercise.01.1-AWARE-ASA- TRL2 will be conducted as part of a workshop for gathering feedback on the ASA system functionality. The primary objective is to perform an early-stage assessment of the ASA system (TRL2) through hands-on demonstrations and user feedback on system outputs. A total of 10 ATCOs will participate in the workshop: 2 from LFV, 4 from UkSATSE, and 4 from IFATCA.

This workshop addresses human performance, aiming to improve related KPAs by focusing on ATCO feedback via interviews and questionnaires regarding the system's usability and functionality.

In the introduction, each ATCO will first receive a system briefing, an explanation of the validation exercise environment and guidance on how to approach the hands-on demonstrations. After the introductory segment, short traffic sample demonstrations will be used to showcase all ASA system functionalities. The emphasis will be on demonstrating the system's functionalities rather than evaluating ATCO task performance. Following the demonstrations, participants will complete questionnaires and participate in an interview to provide detailed feedback.

The key feedback areas include:

- Evaluation of the ASA system across both Advisory (Support) Mode and Execution Mode.
- Insights into system functionality, usability, and ATCO receptivity.
- Assessment of system consistency, as well as accuracy and timeliness.
- Usability and potential impact of the ASA system on ATCOs situational awareness during typical en-route operations.
- Feedback on the HMI.

The validation exercise targets TRL2, representing an early-stage assessment focusing on ASA system functionality and ATCO interaction, where user feedback is essential for further development phases.

5.1.2 Stakeholder's expectations and benefit mechanisms addressed by the exercise

Stakeholder	Involvement	Why it matters to the stakeholder
ATCOs	ATCOs actively interact with the ASA system in a guided workshop, experiencing its functionality firsthand. Through interviews and questionnaires, they provide early feedback on the system's usability and potential impact on their situational awareness during typical en-route operations.	The exercise centres on ATCOs operational needs, enabling them to shape ASA system's development to their needs by offering early feedback. Their input is crucial to ensure ASA system's design aligns with practical requirements.

Table 19. Stakeholders' involvement and expectations for exercise TVAL.01.1-AWARE-ASA-TRL2

5.1.3 Validation objectives

SESAR solution validation objective	SESAR solution success criteria	Coverage and comments on the coverage of SESAR solution validation objective in exercise #01	Exercise validation objective	Exercise success criteria
Assess the impact on the required Human Performance levels due to AWARE solution	AWARE leads to lower workload levels	Partially, ATCOs will only provide feedback on demonstration and limited interaction material.	Gathering feedback on ASA system functionality	Positive feedback from ATCOs

Table 20. Validation objectives addressed in validation exercise TVAL.01.1-AWARE-ASA-TRL2

5.1.4 Validation scenarios

N/A, validation exercise TVAL.01.1-AWARE-ASA-TRL2 will be performed as a workshop for gathering feedback on ASA system, so validation scenarios and use cases are not applicable to this validation exercise.

5.1.4.1 Reference scenario(s)

N/A.

5.1.4.2 Solution scenario(s)

N/A.

5.1.5 Exercise validation assumptions

N/A. No additional assumptions compared to the ones presented in section 4.4.

5.1.6 Limitations and impact on the level of significance

In terms of expected limitations and impact on the level of significance, validation exercise TVAL.01.1-AWARE-ASA-TRL2:

- may be limited by the small sample size of ATCO participants, which restricts the representativeness of findings and limits the generalizability of results to other operational environments.
- may incur the risk of variability in ATCO expertise levels, potentially impacting feedback consistency and the perceived reliability of the ASA system's effectiveness across different experience levels.
- may be influenced by workshop-specific conditions that do not fully replicate the complexities of real-time en-route operations, thus impacting the external validity of results.

5.1.7 Validation exercise platform / tool and validation technique

5.1.7.1 Validation exercise platform / tool characteristics

The validation exercise TVAL.01.1-AWARE-ASA-TRL2 will be held at LFV on the Polaris platform, enhanced with the ASA system to support an interactive demonstration of ASA system's functionalities tailored to en-route operations.

Outcomes from this exercise will be validated through ATCO feedback on ASA system usability, gathered via interviews and questionnaires. This qualitative approach will provide early insights into ASA system's operational relevance, aiding system development and improvements.

5.1.7.2 Validation exercise technique

For validation exercise TVAL.01.1-AWARE-ASA-TRL2, the chosen validation technique is a workshop-based approach involving guided interaction with the ASA system. This technique enables ATCOs to experience the system's functionality in a controlled, simulated environment, allowing for in-depth qualitative feedback through direct interaction, interviews, and structured questionnaires.

5.1.8 Data collection and analysis

5.1.8.1 Data and data collection methods

For validation exercise TVAL.01.1-AWARE-ASA-TRL2, primarily qualitative and limited quantitative data will be collected to assess the ASA system's functionality in a simulated environment. Qualitative data will come from ATCO feedback via interviews and questionnaire responses, capturing insights on usability. ATCO feedback will inform about needed refinements and enhancements to ASA system functionalities, ensuring it evolves in alignment with ATCO operational needs.

Data will be recorded as text and stored digitally in secure file logs, with video recording (with consent) of each session for additional qualitative review. Data processing will involve organising and coding

Co-funded by the European Union

responses in text analysis software and using basic descriptive statistics in Excel for quantitative responses. Data will be reported in D2.24 Exploratory Research Report.

5.1.8.2 Analysis methods

The analysis of validation exercise TVAL.01.1-AWARE-ASA-TRL2 results will employ qualitative data analysis and descriptive statistics to assess the feedback and interactions from ATCOs with the ASA system. Data collected through structured interviews and questionnaires will be coded to identify recurring themes, sentiments, and usability issues, helping in future system development.

5.1.9 Exercise planning and management

5.1.9.1 Activities

<u>Preparatory phase (PREP):</u>

- **Preparation of questionnaires:** Design and finalize structured questionnaires to capture ATCO feedback on ASA system usability and effectiveness.
- **Development of interview protocol:** Create a detailed interview guide for gathering in-depth qualitative insights from ATCOs.
- **Demonstration presentation setup:** Develop a guided, interactive presentation of ASA system features within the Polaris platform.
- **Platform integration and testing:** Integrate the ASA system into Polaris and conduct preliminary tests to ensure functionality and stability.
- **Selection of demonstration traffic sample:** Identify which system functionalities should the demonstration address and prepare traffic scenario sample that would showcase all functionalities equally.
- **Participant recruitment and scheduling:** Identify and schedule a limited number of ATCOs for participation in the workshop.

<u>Execution phase (EXEC):</u>

- **Conducting the workshop:** Conduct the workshop, where ATCOs experience ASA system functionalities in a simulated en-route context.
- **Collection of feedback:** Record ATCO feedback via structured interviews, questionnaires, and video recordings.

<u>Post-execution phase (POST-EXEC):</u>

- **Data processing and coding:** Organise and code qualitative data for analysis and input quantitative responses into a statistical tool.
- **Preliminary data analysis:** Perform an initial analysis of qualitative feedback and quantitative data.
- **Feedback integration:** Document findings and identify potential ASA system improvement areas for further system development.

5.1.9.2 Roles and responsibilities in the exercise

In validation exercise TVAL.01.1-AWARE-ASA-TRL2, several consortium members are involved and hold distinct responsibilities essential for the preparation, execution, and analysis of the workshop.

ZHAW will oversee the development, implementation, and analysis of the interviews and questionnaires.

LFV will provide both the ATCO participants, and the physical space required for conducting the workshop.

UkSATSE and IFATCA will contribute additional ATCOs, enhancing the exercise's diversity by ensuring that participants bring a range of operational expertise.

TERN and FTTS will supply technical support, focusing on the integration of the ASA system into the Polaris platform.

5.1.9.3 Time planning

Activity	Week			
	1	2	3	4
Preparation of questionnaires (PREP)	■	■		
Development of interview protocol (PREP)	■	■		
Demonstration presentation setup (PREP)	■	■		
Platform integration and testing (PREP)	■	■		
Selection of demonstration traffic sample (PREP)	■	■		
Participant recruitment and scheduling (PREP)	■	■		
Conducting the workshop and collection of feedback (EXEC)			■	
Data processing and coding (POST-EXEC)			■	■
Preliminary data analysis (POST-EXEC)			■	■
Feedback integration (POST-EXEC)			■	■

Table 21. Detailed exercise TVAL.01.1-AWARE-ASA-TRL2 time planning

Preparatory phase of validation exercise TVAL.01.1-AWARE-ASA-TRL2 will be executed during the 3rd and 4th week of October, starting with the preparation of questionnaires and development of interview protocol.

The execution phase will follow the preparatory phase will be performed during the 1st week of November 2025.

The post-execution phase will start after the conducted workshop in the 2nd week of November 2025 and will continue until the end of the month. The results of the post-execution phase will be reported in D2.24 Exploratory Research Report.

5.1.9.4 Identified risks and mitigation actions

Risks	Impact (1-low, 2-medium, 3-high)	Likelihood (1-low, 2-medium, 3-high)	Criticality (calculated based on likelihood and impact)	Mitigation actions
Ukrainian ATCOs cannot leave Ukraine or UkSATSE becomes unavailable	2 - medium	1 - low	1 -low	LFV and IFATCA provide more controllers and higher input than previously planned. In extreme case, simulation campaign is halved.

Table 22. Exercise TVAL.01.1-AWARE-ASA-TRL2 risks and mitigation actions, [2]

5.2 Validation exercise TVAL.02.1-AWARE-ASA-TRL2 plan

5.2.1 Validation exercise description and scope

This exercise is designed to validate the ASA system's performance under medium traffic load and complexity using the Polaris ATM platform, with varying levels of ASA support. Depending on the group ATCOs are assigned to, they will perform exercise either without any system support, with system-assisted task execution, or with system support complemented by autonomous operations running in the background. A total of 20 ATCOs will participate. The reference airspace is the Malmö ACC Sector 23 (ESMM23), which controls en-route traffic. The scenario is based on actual traffic from 4th of August 2024.

Its objective is to collect operational data from medium traffic load and complexity that will be used to assess the system's impact on key performance areas. HITL simulations will serve as the validation technique, providing a realistic environment for ATCOs to interact with the ASA system. The applicable TRL for this exercise is TRL2. Validation will occur at LFV using the Polaris ATM platform, encompassing use cases U.C.1, U.C.2, U.C.3 and U.C.4 from the SPR-INTEROP/OSED framework.

5.2.2 Stakeholder's expectations and benefit mechanisms addressed by the exercise

Stakeholder	Involvement	Why it matters to the stakeholder
ATCOs	ATCOs are essential participants in this HITL simulation, actively managing high-traffic scenarios. They will interact directly with the ASA system, using it as a support tool to manage medium traffic scenarios in the Polaris system. Their interaction and feedback on the system's effectiveness will be gathered through HITL simulations.	This exercise matters to ATCOs as it informs ASA system development tailored to their needs. By highlighting medium workload and operational challenges, ATCOs' input ensures that ASA enhancements will be practical and supportive, ultimately improving their situational awareness, reducing workload and fostering a safer, more efficient work environment.
SUPs	SUPs are involved in overseeing ATCOs during the HITL simulation, ensuring procedural adherence and providing operational context to support accurate data collection. Their role includes monitoring ATCO workload and decision-making processes, which helps validate the ASA system's design to align with supervisory needs in traffic management.	For SUPs, this exercise is vital as it helps shape ASA functionalities that support effective oversight and workload management. By establishing medium level operational challenges and workload, SUPs ensure that ASA enhancements will improve team coordination, streamline operations and better address supervisory needs in high-traffic conditions.
ATSEPs	ATSEPs support the HITL simulation by ensuring the functionality and reliability of the Polaris ATM system used in the exercise. They monitor technical performance, address system issues and provide insights into equipment limitations, helping validate the ASA system's compatibility with existing infrastructure and technology.	This exercise is crucial for ATSEPs as it highlights system performance under realistic operational loads, informing ASA integration with current infrastructure. Their involvement ensures that ASA enhancements will be compatible, reliable and supportable within existing technical systems, ultimately aiding ATSEPs in maintaining system safety and operational efficiency.
ANSPs	ANSPs will observe performance by simulating medium traffic levels on the Polaris ATM system with varying level of ASA support. The objective is to gather data for subsequent ASA evaluations, assessing key performance areas such as capacity, cost-efficiency, safety, human performance and operational efficiency.	This exercise provides ANSPs with crucial data to assess ASA's impact on traffic management efficiency, safety, and cost-effectiveness. It helps identify areas for operational improvement, ensuring ANSPs can make data-driven decisions on ASA integration for enhanced airspace management.

Stakeholder	Involvement	Why it matters to the stakeholder
Airspace Users	Airspace users will indirectly benefit from this exercise as it gathers performance metrics for the ASA system's potential benefits. By evaluating traffic management without ASA, airspace users gain insight into improvements ASA might bring in efficiency, safety, and cost-effectiveness once implemented.	For airspace users, this exercise is essential as it assess the potential ASA improvements in flight efficiency, safety, and cost savings in medium traffic loaf. By understanding current performance without ASA, users can anticipate tangible benefits ASA may provide, impacting scheduling, fuel use, and overall operational costs.
ATM Research Community	The ATM research community will analyse data from this exercise to assess ASA's potential in enhancing air traffic management. This involvement supports research on human-AI collaboration, validates methodologies in high-traffic scenarios and aligns ASA advancements with SESAR goals for improved operational efficiency and safety.	This exercise is significant for the ATM research community as it provides essential data for validating ASA's impact on operational efficiency and safety. Insights gained will guide future research, support SESAR-aligned innovation and help refine AI-human collaboration frameworks critical for advancing air traffic management solutions.

Table 23. Stakeholders' involvement and expectations for exercise TVAL.02.1-AWARE-ASA-TRL2

5.2.3 Validation objectives

SESAR solution validation objective	SESAR solution success criteria	Coverage and comments on the coverage of SESAR solution validation objective in exercise TVAL.02.1-AWARE-ASA-TRL2	Exercise validation objective	Exercise success criteria
Assess the increase in Capacity due to the AWARE solution.	AWARE leads to increased Capacity.	Fully covered	Measure ASA system's impact on en-route throughput and traffic handling capability during peak demand hours	ASA system assists ATCOs in maintaining or increasing throughput without performance degradation.
Assess the improvement in Cost Efficiency due to the AWARE solution.	AWARE leads to increased Cost Efficiency.	Fully covered	Assess ASA system's effect on flights handled per ATCO-hour, improving resource allocation.	ASA system results in a higher ratio of flights per ATCO-hour, optimising operational costs.

SESAR solution validation objective	SESAR solution success criteria	Coverage and comments on the coverage of SESAR solution validation objective in exercise TVAL.02.1-AWARE-ASA-TRL2	Exercise validation objective	Exercise success criteria
Assess the improvement in Fuel Efficiency due to AWARE solution.	AWARE leads to reduction in fuel consumption.	Fully covered	Evaluate ASA system's role in reducing fuel consumption and CO_2 emissions by optimising flight management.	ASA system reduces average CO_2 emissions per flight compared to baseline operations.
Assess the improvement of Safety due to the AWARE solution.	AWARE leads to fewer conflicts, resulting in increased safety.	Fully covered	Analyse the reduction in conflict events and overall safety risks with ASA system support during peak traffic.	ASA system support lowers the number of conflict-induced safety incidents.
Assess the impact on the required Human Performance levels due to AWARE solution.	AWARE leads to lower workload levels.	Fully covered	Assess the reduction in ATCO workload and stress levels with ASA system assistance during high traffic.	ATCOs report lower workload and improved task performance with ASA system support.
Assess the improvement on the required Operational Efficiency due to the AWARE solution.	AWARE leads to increased Operational Efficiency.	Fully covered	Evaluate ASA system's effect on operational fuel efficiency during high-traffic scenarios.	ASA system improves fuel efficiency, reducing the average fuel burn per flight.

Table 24. Validation objectives addressed in validation exercise TVAL.02.1-AWARE-ASA-TRL2

5.2.4 Validation scenarios

The validation scenario identified as TVAL.02.1-AWARE-ASA-TRL2 aims to gather data for the evaluation of the ASA system's impact on ATCO operations in medium traffic load and complexity. Conducted on the Polaris ATM system, this exercise will gather data to establishing metrics for capacity, safety, cost-efficiency, environmental impact, human performance, and operational efficiency. Each Group with varying level of ASA support will be executed at least five times. By benchmarking ATCO performance in Group 1 without ASA support, essential insights into the potential advantages that ASA may offer are provided.

5.2.4.1 Reference scenario(s)

The reference scenario for this validation exercise is the Group 1 of the exercise TVAL.02.1-AWARE-ASA-TRL2 in which ATCOs manage medium traffic load without any ASA system support. This scenario is in line with typical operational conditions, allowing ATCOs to perform their tasks using standard procedures. The reference scenario will provide a direct comparison of traffic management, efficiency, and safety outcomes. Since the ASA system is not included in reference scenarios, they serve as the baseline for evaluating the improvements provided by the ASA-supported solution scenario.

5.2.4.2 Solution scenario(s)

The solution scenario for Validation Exercise TVAL.02.1-AWARE-ASA-TRL2 mirrors the reference scenarios in terms of airspace configuration, traffic load and operational parameters. The only difference is the inclusion of the ASA system which provides support to ATCOs. This scenario is designed to assess the impact of ASA assistance on ATCO performance, traffic management, and safety outcomes during medium traffic conditions, allowing for direct comparison with the reference scenario.

5.2.5 Exercise validation assumptions

Assumption ID	Assumption title	Assumption description	Justification	Impact Assessment
VA2.1	ASA System Intermediate Support	The ASA system will provide intermediate support to ATCOs but will not perform any autonomous actions.	This assumption ensures that the focus remains on how ASA assists ATCOs without taking over any tasks autonomously, aligning with the experiment’s objectives.	A direct impact on workload and situational awareness results is predicted
VA2.2	ASA System Autonomous Support	The ASA system will provide fully autonomous support to ATCOs.	The focus will be on how ASA performs autonomous tasks, and how it will choose which tasks to take over from the ATCO.	A direct impact on workload and situational awareness results is predicted

VA2.3	Stable Simulation Environment	The simulation environment will remain stable without introducing unexpected disruptions (e.g., technical issues).	A stable environment is necessary to ensure accurate data collection and isolate the impact of the ASA system's intermediate support.	High: Any disruptions may compromise the reliability of results
VA2.4	ATCO Familiarity with ASA System	ATCOs participating in the exercise have received adequate training on the ASA system's intermediate functions.	Ensures ATCOs are comfortable using the ASA system, minimizing variability in results due to unfamiliarity.	Moderate: Insufficient training could affect performance outcomes.

Table 25. Validation exercise TVAL.02.1-AWARE-ASA-TRL2 assumptions

5.2.6 Limitations and impact on the level of significance

The following limitations may impact the level of significance and the generalizability of the validation results:

- The controlled simulation environment may not fully capture the complexity of real-world traffic conditions, limiting generalization to operational settings.
- A limited number of ATCO participants may impact the representativeness of the findings, making it difficult to extrapolate to a broader population.
- The new operational environment (Polaris and the ASA system), combined with the absence of a planner position, may affect ATCOs' situational awareness, operational efficiency, and workload.
- the ASA system's performance at EASA level 2B, while advanced, is still in its early stages (TRL2), which restricts the breadth of conclusions that can be drawn about its effectiveness in operational settings.
- External factors, such as varying traffic mixes and weather are not fully represented in the simulation, which may affect the applicability of results to different environments.
- May introduce potential sequence effects, such as fatigue or learning, affecting feedback reliability over the course of the experiment, despite efforts to control these factors.
- The findings will provide valuable insights, but further real-world validation is necessary to generalize the results to larger scales, such as the ECAC level.

5.2.7 Validation exercise platform / tool and validation technique

5.2.7.1 Validation exercise platform / tool characteristics

The Polaris ATM system integrated with the ASA system providing both intermediate and execution support to ATCOs. is employed as the primary validation platform for the TVAL.02.1-AWARE-ASA-TRL2

Co-funded by the European Union

exercise. Polaris is a robust and adaptable ATM tool used for simulating complex traffic scenarios, providing a comprehensive environment for analysing ATCO performance under varying traffic densities. This system allows for precise replication of operational conditions observed during peak times, supporting a high level of data fidelity essential for effective baseline measurements. Its capacity for handling medium to high traffic loads ensures that the collected data accurately reflects real-world operational challenges that ATCOs face, facilitating an objective assessment of human performance without ASA support.

5.2.7.2 Validation exercise technique

The chosen validation technique for this exercise is HITL simulations. This approach involves ATCOs participating in controlled simulations where they manage medium traffic levels with our without the assistance of ASA functionalities. This technique allows real-time interaction between ATCOs and the ASA system, making it ideal for assessing the system's support in medium traffic scenarios. HITL simulations are particularly advantageous for collecting detailed insights into human behaviour, decision-making, workload, and situational awareness under various operational conditions. These simulations capture both qualitative and quantitative performance data, offering an in-depth view of how ATCOs adapt and respond to dynamic air traffic scenarios.

5.2.8 Data collection and analysis

5.2.8.1 Data and data collection methods

The TVAL.02.1-AWARE-ASA-TRL2 validation exercise will employ comprehensive data collection strategies to capture ATCO performance and situational awareness under medium traffic levels with varying levels of ASA support. The core data collection methods will include simulation logs, screen recording, eye tracking recordings, post-exercise questionnaires, and real-time observations by subject matter experts (SMEs). The Polaris ATM system's built-in simulation logging capabilities will provide an extensive record of ATCO interactions, encompassing response times and task management actions. Screen and eye tracking recordings will give insight into ATCO intent and task prioritisation. This quantitative data will form the foundation for evaluating operational efficiency, workload distribution, situational awareness and safety metrics.

Post-simulation debriefing through structured questionnaires will complement these logs, allowing ATCOs to offer subjective insights into their workload perception and encountered challenges. This qualitative approach will be essential for understanding human factors such as situational awareness, decision-making under pressure, and overall stress levels during high traffic scenarios. Additionally, Subject Matter Expert (SME) observations recorded in real time will focus on ATCO behaviour, communication practices, and procedural adherence, providing a nuanced layer of context to the simulation data.

5.2.8.2 Analysis methods

The analysis of the collected data will follow a methodical approach designed to yield a clear understanding of ATCO performance in a medium traffic scenario with varying level of ASA support. Quantitative analysis will involve detailed processing of simulation log data, screen and eye tracking recordings to identify key performance outcomes related to capacity, safety, and operational throughput. Statistical analysis, including descriptive and inferential techniques, will be used to establish baseline metrics (referred to Group 1), which will serve as comparison points for ASA-supported scenarios (referred to Groups 2 -4).

Simultaneously, qualitative analysis of ATCO feedback from questionnaires and SME notes will be conducted using thematic analysis to detect recurring insights into workload and situational management. This step will enrich the understanding of human performance by linking quantitative data with subjective experiences. The cross-referencing of quantitative results with qualitative observations will ensure the robustness and reliability of the findings, facilitating a comprehensive assessment of ATCO capabilities in both non-AI supported and AI supported traffic management. The results will help in assessing the impact of ASA in enhancing ATCO performance and mitigating workload.

5.2.9 Exercise planning and management

5.2.9.1 Activities

The activities for Validation Exercise TVAL.02.1-AWARE-ASA-TRL2 are divided into three phases as follows:

Preparatory phase (PREP):

- Setting up the validation infrastructure.
- Integrating the ASA system with the POLARIS ATM platform.
- Conducting preparatory sessions for ATCOs.

Execution phase (EXEC):

- Conducting HITL simulations under medium traffic loads with varying level of ASA support.
- Monitoring ATCO performance and system functionality in real time.

Post-execution phase (POST-EXEC):

- Evaluating simulation outcomes.
- Preparing the exercise report.

5.2.9.2 Roles and responsibilities in the exercise

In validation exercise TVAL.02.1-AWARE-ASA-TRL2, several consortium members are involved and hold distinct responsibilities essential for the preparation, execution, and analysis of the workshop.

ZHAW will oversee the development, implementation, and analysis of the interviews and questionnaires.

LFV will provide both the ATCO participants, and the physical space required for conducting the workshop.

UkSATSE and IFATCA will contribute additional ATCOs, enhancing the exercise's diversity by ensuring that participants bring a range of operational expertise.

TERN and FTTS will supply technical support, focusing on the integration of the ASA system into the Polaris platform.

5.2.9.3 Time planning

Activity	Week 1	Week 2
Setting up validation infrastructure (PREP)		
Integrating the ASA system with the POLARIS ATM platform (PREP)		
Conducting preparatory sessions for ATCOs (PREP)		
Running HITL simulations with medium traffic load and data collection (EXEC)		
Evaluating simulation outcomes (POST-EXEC)		
Preparing the exercise report (POST-EXEC)		

Table 26. Detailed exercise TVAL.02.1-AWARE-ASA-TRL2 time planning

5.2.9.4 Identified risks and mitigation actions

Risks	Impact (1-low, 2-medium, 3-high)	Likelihood (1-low, 2-medium, 3-high)	Criticality (calculated based on likelihood and impact)	Mitigation actions
Ukrainian ATCOs cannot leave Ukraine or UkSATSE becomes unavailable.	2	2	2	LFV and IFATCA provide more controllers and higher input than previously planned. In extreme case, simulation campaign is halved.
Inability to assess ATCO intent from traffic situation semantics, attention tracking, and HMI tracking	3	1	1	ATCO intent will be either partially mapped or mapped to a smaller subset of parameters to enable the continuation of research. An additional deliverable should be prepared to provide detailed descriptions of methodology, encountered problems, and possible avenues of future research.

Table 27. Exercise TVAL.02.1-AWARE-ASA-TRL2 risks and mitigation actions, [2]

5.3 Validation exercise TVAL.02.2-AWARE-ASA-TRL2 plan

5.3.1 Validation exercise description and scope

This validation exercise aims to assess the benefits of the ASA system integrated into Polaris under high traffic levels with specific focus on the conflict detection and resolution. The operational scope includes ATCOs using the ASA system with varying levels of support. Depending on their assigned group, ATCOs will perform the exercise either without any system support, with system-assisted task execution, or with system support complemented by autonomous operations running in the background. These varying support levels will enable an evaluation of how each level influences ATCO performance, workload, situational awareness, capacity, and overall operational efficiency. The exercise addresses use cases U.C.1, U.C.2, U.C.3 and U.C.4, focusing on scenarios where ATCOs interact with the ASA system to manage air traffic effectively. In total, 20 ATCOs divided into four groups will participate in this exercise. The reference airspace is the Malmö ACC Sector 23 (ESMM23), which controls en-route traffic. The scenario is based on actual traffic from 4th of August 2024.

The key validation objectives are to gather data on the ASA system's impact across KPAs, including capacity, cost-efficiency, environmental sustainability, safety, human performance, and operational efficiency. HITL simulations will serve as the validation technique, providing a realistic environment for ATCOs to interact with the augmented Polaris ATM system. The applicable TRL for this exercise is TRL2, indicating that the technology concept and application have been formulated but not yet validated in an operational setting.

5.3.2 Stakeholder's expectations and benefit mechanisms addressed by the exercise

Stakeholder	Involvement	Why it matters to the stakeholder
ATCOs	ATCOs are the primary actors in this validation exercise. They will interact directly with the ASA system, using it as a support tool to manage high traffic scenarios in the Polaris system. Their interaction and feedback on the system's effectiveness will be gathered through HITL simulations.	The exercise evaluates the ASA system's capacity to enhance ATCO performance with varying level of ASA support. Areas that matter are improving situational awareness, reducing workload during peak periods, and ensuring that the system aids in maintaining high safety and efficiency standards. Success in this exercise could lead to operational improvements, better handling of air traffic, and increased safety.

Stakeholder	Involvement	Why it matters to the stakeholder
SUPs	SUPs play a critical role in overseeing ATCOs' interaction with the Polaris system during the validation exercise. Their focus is on ensuring that ATCOs are managing traffic effectively under high traffic levels with varying level of ASA support, and they will provide feedback on operational safety, effectiveness, and human-machine collaboration.	For SUPs, this exercise offers a detailed understanding of ATCO performance under high-stress conditions with or without advanced AI support. SUPs are interested in assessing how the use of ASA impacts operational safety and workload management. The insights gathered will guide future decisions on integrating AI-based tools like ASA.
ATSEPs	ATSEPs will be involved in ensuring the correct integration of the ASA system with the existing Polaris ATM infrastructure. They will support technical maintenance and ensure system reliability during the validation exercise.	For ATSEPs, this exercise provides critical insights into the operational readiness and maintenance demands of integrating advanced AI support systems like ASA.
ANSPs	ANSPs will oversee the execution of the validation exercise, focusing on how the ASA system can be used to enhance airspace capacity and manage traffic flow efficiently during peak periods.	ANSPs are directly interested in how the ASA system can improve capacity, operational efficiency, and cost-effectiveness in managing air traffic. This validation could demonstrate how ASA improves decision-making in real-time, reduces delays, and supports higher levels of traffic throughput without compromising safety.
Airspace Users	Airspace users, such as airlines, will be indirectly involved through the outcomes of the exercise, as they will benefit from improved traffic management and reduced delays.	For airspace users, the results of this exercise could lead to more efficient traffic management, reducing delays and operational costs. Improved operational efficiency, as evidenced by the exercise, would result in lower fuel consumption, enhanced punctuality, and reduced costs for airlines. This aligns with environmental sustainability and operational efficiency objectives.
ATM Research Community	The ATM research community will analyse the data gathered during the exercise, focusing on the human factors and technical performance of the system under high traffic conditions with varying ASA support.	For the research community, this exercise provides valuable data on human performance and system interactions in high-complexity scenarios. These insights will inform future developments of ASA systems, ensuring that the technology meets the operational needs of ATCOs and ANSPs and supports safer, more efficient traffic management.

Table 28. Stakeholders' involvement and expectations for exercise TVAL.02.2-AWARE-ASA-TRL2

5.3.3 Validation objectives

SESAR solution validation objective	SESAR solution success criteria	Coverage and comments on the coverage of SESAR solution validation objective in exercise TVAL.02.2-AWARE-ASA-TRL2	Exercise validation objective	Exercise success criteria
Assess the increase in Capacity due to the AWARE solution	AWARE leads to increased Capacity.	Fully covered	Measure ASA system's impact on en-route throughput and traffic handling capability during peak demand hours	ASA system assists ATCOs in maintaining or increasing throughput without performance degradation.
Assess the improvement in Cost Efficiency due to the AWARE solution.	AWARE leads to increased Cost Efficiency.	Fully covered	Assess ASA system's effect on flights handled per ATCO-hour, improving resource allocation.	ASA system results in a higher ratio of flights per ATCO-hour, optimising operational costs.
Assess the improvement in Fuel Efficiency due to AWARE solution.	AWARE leads to reduction in fuel consumption.	Fully covered	Evaluate ASA system's role in reducing fuel consumption and CO_2 emissions by optimising flight management.	ASA system reduces average CO_2 emissions per flight compared to baseline operations.
Assess the improvement of Safety due to the AWARE solution.	AWARE leads to fewer conflicts, resulting in increased safety.	Fully covered	Analyse the reduction in conflict events and overall safety risks with ASA system support during peak traffic.	ASA system support lowers the number of conflict-induced safety incidents.
Assess the impact on the required Human Performance levels due to AWARE solution.	AWARE leads to lower workload levels.	Fully covered	Assess the reduction in ATCO workload and stress levels with ASA system assistance during high traffic.	ATCOs report lower workload and improved task performance with ASA system support.

SESAR solution validation objective	SESAR solution success criteria	Coverage and comments on the coverage of SESAR solution validation objective in exercise TVAL.02.2-AWARE-ASA-TRL2	Exercise validation objective	Exercise success criteria
Assess the improvement on the required Operational Efficiency due to the AWARE solution.	AWARE leads to increased Operational Efficiency.	Fully covered	Evaluate ASA system's effect on operational fuel efficiency during high-traffic scenarios.	ASA system improves fuel efficiency, reducing the average fuel burn per flight.

Table 29. Validation objectives addressed in validation exercise TVAL.02.2-AWARE-ASA-TRL2

5.3.4 Validation scenarios

The validation scenarios for Exercise TVAL.02.2-AWARE-ASA-TRL2 involve high level traffic conditions with the focus on conflict detection and resolution. Exercise will be run with the ASA system providing varying support to ATCOs. Each group will include five ATCOs, ensuring that every support level is performed at least five times. This allows for comparison of performance, efficiency, and safety between baseline operations and ASA-supported operations.

5.3.4.1 Reference scenario(s)

The reference scenario for this validation exercise is Group 1 of the exercise TVAL.02.2-AWARE-ASA-TRL2 in which ATCOs manage high traffic loads without any assistance from the ASA system. This scenario is in line with typical operational conditions, allowing ATCOs to perform their tasks using standard procedures. The reference scenarios will provide a direct comparison of traffic management, efficiency, and safety outcomes. Since the ASA system is not included in reference scenarios, they serve as the baseline for evaluating the improvements provided by the ASA-supported solution scenario.

5.3.4.2 Solution scenario(s)

The solution scenario for Validation Exercise TVAL.02.2-AWARE-ASA-TRL2 mirrors the reference scenarios in terms of airspace configuration, traffic load and operational parameters. The only difference is the inclusion of the ASA system which provides support to ATCOs. This scenario is designed to assess the impact of ASA assistance on ATCO performance, traffic management, and safety outcomes during high traffic conditions, allowing for direct comparison with the reference scenario.

5.3.5 Exercise validation assumptions

Assumption ID	Assumption title	Assumption description	Justification	Impact Assessment
VA2.1	ASA System Intermediate Support	The ASA system will provide intermediate support to ATCOs but will not perform any autonomous actions.	This assumption ensures that the focus remains on how ASA assists ATCOs without taking over any tasks autonomously, aligning with the experiment's objectives.	A direct impact on workload and situational awareness results is predicted
VA2.2	ASA System Autonomous Support	The ASA system will provide fully autonomous support to ATCOs.	The focus will be on how ASA performs autonomous tasks, and how it will choose which tasks to take over from the ATCO.	A direct impact on workload and situational awareness results is predicted
VA2.3	Stable Simulation Environment	The simulation environment will remain stable without introducing unexpected disruptions (e.g., technical issues).	A stable environment is necessary to ensure accurate data collection and isolate the impact of the ASA system's intermediate support.	High: Any disruptions may compromise the reliability of results
VA2.4	ATCO Familiarity with ASA System	ATCOs participating in the exercise have received adequate training on the ASA system's intermediate functions.	Ensures ATCOs are comfortable using the ASA system, minimizing variability in results due to unfamiliarity.	Moderate: Insufficient training could affect performance outcomes.

Table 30. Validation exercise TVAL.02.2-AWARE-ASA-TRL2 assumptions

5.3.6 Limitations and impact on the level of significance

The following limitations may impact the level of significance and the generalizability of the validation results:

- The controlled simulation environment may not fully capture the complexity of real-world traffic conditions, limiting generalization to operational settings.
- A limited number of ATCO participants may impact the representativeness of the findings, making it difficult to extrapolate to a broader population.
- The new operational environment (Polaris and the ASA system), combined with the absence of a planner position, may affect ATCOs' situational awareness, operational efficiency, and workload.

Co-funded by the European Union

- the ASA system's performance at EASA level 2B, while advanced, is still in its early stages (TRL2), which restricts the breadth of conclusions that can be drawn about its effectiveness in operational settings.
- External factors, such as varying traffic mixes and weather are not fully represented in the simulation, which may affect the applicability of results to different environments.
- May introduce potential sequence effects, such as fatigue or learning, affecting feedback reliability over the course of the experiment, despite efforts to control these factors.
- The findings will provide valuable insights, but further real-world validation is necessary to generalize the results to larger scales, such as the ECAC level.

5.3.7 Validation exercise platform / tool and validation technique

5.3.7.1 Validation exercise platform / tool characteristics

The tools and platform required for this validation exercise include the POLARIS ATM system, integrated with the ASA system providing both intermediate and execution support to ATCOs. The exercise will be conducted in a simulation environment at the designated validation site, with HITL simulations enabling real-time interaction between ATCOs and the ASA system. The POLARIS system is equipped to simulate high traffic conditions with and without autonomous task execution by the ASA system, ensuring the validation of its supportive and execution role in scenario focused on conflict detection and resolution.

5.3.7.2 Validation exercise technique

The chosen validation technique for this validation exercise is Human-In-The-Loop simulations. This technique allows real-time interaction between ATCOs and the ASA system, making it ideal for assessing the system's support in high traffic scenarios. HITL simulations are particularly suitable for capturing detailed human performance data and evaluating the ASA system's impact on ATCO workload, safety, and operational efficiency, meeting the exercise's validation objectives.

5.3.8 Data collection and analysis

5.3.8.1 Data and data collection methods

The validation exercise will collect both quantitative and qualitative data. Quantitative data will include metrics such as eye tracking data, traffic throughput, conflict resolution, fuel consumption, and ATCO workload levels, while qualitative data will focus on ATCO feedback through post-simulation questionnaires assessing workload, situational awareness, and system usability. Data collection will be performed through system logs, real-time monitoring, and surveys. Multiple simulation iterations are planned to ensure robust data collection, and all data sets (including traffic data, system performance logs, and ATCO inputs) will be processed and cleaned to ensure consistency and accuracy. Output formats will include system log files, video recordings, and questionnaire data. Financial resources are sufficient for data collection, and proprietary and validated data sources will be used to meet research needs.

5.3.8.2 Analysis methods

The collected data will be analysed using both quantitative and qualitative methods to assess the impact of ASA system support on ATCO performance and operational efficiency. Quantitative data will be compared between the reference (referred to Group 1) and solution scenarios (referred to Groups 2-4) to identify any significant improvements in key performance indicators such as traffic throughput and safety. Qualitative data, such as ATCO feedback, will be analysed to identify trends in workload and system usability. Post-processing will involve cleaning and organising the data for consistency, ensuring accurate analysis. These methods will help determine whether the ASA system meets the success criteria and addresses the research objectives.

5.3.9 Exercise planning and management

5.3.9.1 Activities

The activities for Validation Exercise TVAL.02.2-AWARE-ASA-TRL2 are divided into three phases as follows:

Preparatory phase (PREP):

- Setting up the validation infrastructure.
- Integrating the ASA system with the POLARIS ATM platform.
- Conducting training sessions for ATCOs on using the ASA system.

Execution phase (EXEC):

- Running the HITL simulations under high traffic loads with varying level of ASA support.
- Monitoring ATCO performance and system functionality in real time.

Post-execution phase (POST-EXEC):

- Evaluating simulation outcomes.
- Preparing the exercise report.

5.3.9.2 Roles and responsibilities in the exercise

In validation exercise TVAL.02.2-AWARE-ASA-TRL2, several consortium members are involved and hold distinct responsibilities essential for the preparation, execution, and analysis of the exercise.

ZHAW will oversee the development, implementation, and analysis of the interviews and questionnaires.

LFV will provide both the ATCO participants, and the physical space required for conducting the experiments.

UkSATSE and IFATCA will contribute additional ATCOs, enhancing the exercise's diversity by ensuring that participants bring a range of operational expertise.

TERN and FTTS will supply technical support, focusing on the integration of the ASA system into the Polaris platform.

5.3.9.3 Time planning

Activity	Week 1	Week 2
Setting up validation infrastructure (PREP)		
Integrating the ASA system with the POLARIS ATM platform (PREP)		
Conducting training sessions for ATCOs (PREP)		
Running HITL simulations with high traffic load and data collection (EXEC)		
Evaluating simulation outcomes (POST-EXEC)		
Preparing the exercise report (POST-EXEC)		

Table 31. Detailed exercise TVAL.02.2-AWARE-ASA-TRL2 planning

5.3.9.4 Identified risks and mitigation actions

Risks	Impact (1-low, 2-medium, 3-high)	Likelihood (1-low, 2-medium, 3-high)	Criticality (calculated based on likelihood and impact)	Mitigation actions
Ukrainian ATCOs cannot leave Ukraine or UkSATSE becomes unavailable.	2	2	2	LFV and IFATCA provide more controllers and higher input than previously planned. In extreme case, simulation campaign is halved.
Inability to assess ATCO intent from traffic situation semantics, attention tracking, and HMI tracking	3	1	1	ATCO intent will be either partially mapped or mapped to a smaller subset of parameters to enable the continuation of research. An additional deliverable should be prepared to provide detailed descriptions of methodology, encountered problems, and possible avenues of future research.

Table 32. Exercise TVAL.02.2-AWARE-ASA-TRL2 risks and mitigation actions, [2]

5.4 Validation exercise TVAL.02.3-AWARE-ASA-TRL2 plan

5.4.1 Validation exercise description and scope

The validation exercise TVAL.02.3-AWARE-ASA-TRL2 is designed to assess the ASA system's impact on ATCO performance under simulated high traffic levels on the Polaris ATM system, augmented with ASA functionality. During the exercise, ATCOs will work with different levels of ASA support depending on their assigned group, with particular focus on interactions involving restricted military airspace. Participants will complete the scenario either with no system support, with system-assisted task execution, or with system support complemented by autonomous background operations. The exercise addresses use cases U.C.1, U.C.2, U.C.3 and U.C.4. In total, 20 ATCOs, divided into four groups, will take part in the exercise. The reference airspace is the Malmö ACC Sector 23 (ESMM23), which controls en-route traffic. The scenario is based on actual traffic from 4th of August 2024.

The key validation objectives are to gather data on the ASA system's impact across KPAs, including capacity, cost-efficiency, environmental sustainability, safety, human performance, and operational efficiency. HITL simulations will serve as the validation technique, providing a realistic environment for ATCOs to interact with the augmented Polaris ATM system. This exercise is aligned with TRL 2.

5.4.2 Stakeholder's expectations and benefit mechanisms addressed by the exercise

Stakeholder	Involvement	Why it matters to the stakeholder
ATCOs	ATCOs are the primary actors in this validation exercise. They will interact directly with the ASA system, using it as a support tool to manage high traffic scenarios in the Polaris system. Their interaction and feedback on the system's effectiveness will be gathered through HITL simulations.	The exercise evaluates the ASA system's capacity to enhance ATCO performance with varying level of ASA support. Areas that matter are improving situational awareness, reducing workload during peak periods, and ensuring that the system aids in maintaining high safety and efficiency standards. Success in this exercise could lead to operational improvements, better handling of air traffic, and increased safety.
SUPs	SUPs play a critical role in overseeing ATCOs' interaction with the Polaris system during the validation exercise. Their focus is on ensuring that ATCOs are managing traffic effectively under high traffic levels with ASA support, and they will provide feedback on operational safety, effectiveness, and human-machine collaboration.	For SUPs, this exercise offers a detailed understanding of ATCO performance under high-stress conditions while using AI support. SUPs are interested in assessing how the use of ASA impacts operational safety and workload management. The insights gathered will guide future decisions on integrating AI-based tools like ASA.
ATSEPs	ATSEPs will be involved in ensuring the correct integration of the ASA system with the existing Polaris ATM infrastructure. They will support technical maintenance and ensure system reliability during the validation exercise.	For ATSEPs, this exercise provides critical insights into the operational readiness and maintenance demands of integrating advanced AI support systems like ASA.

Stakeholder	Involvement	Why it matters to the stakeholder
ANSPs	ANSPs will oversee the execution of the validation exercise, focusing on how the ASA system can be used to select which task will be performed by AI, enhance airspace capacity and manage traffic flow efficiently during peak periods.	ANSPs are directly interested in how the ASA system can improve capacity, operational efficiency, and cost-effectiveness in managing air traffic. This validation could demonstrate how ASA improves decision-making in real-time, reduces delays, and supports higher levels of traffic throughput without compromising safety.
Airspace Users	Airspace users, such as airlines, will be indirectly involved through the outcomes of the exercise, as they will benefit from improved traffic management and reduced delays.	For airspace users, the results of this exercise could lead to more efficient traffic management, reducing delays and operational costs. Improved operational efficiency, as evidenced by the exercise, would result in lower fuel consumption, enhanced punctuality, and reduced costs for airlines. This aligns with environmental sustainability and operational efficiency objectives.
ATM Research Community	The ATM research community will analyse the data gathered during the exercise, focusing on the human factors and technical performance of the system under high traffic conditions with ASA support.	For the research community, this exercise provides valuable data on human performance and system interactions in high-complexity scenarios. These insights will inform future developments of ASA systems, ensuring that the technology meets the operational needs of ATCOs and ANSPs and supports safer, more efficient traffic management.

Table 33. Stakeholders' involvement and expectations for exercise TVAL.02.3-AWARE-ASA-TRL2

5.4.3 Validation objectives

SESAR solution validation objective	SESAR solution success criteria	Coverage and comments on the coverage of SESAR solution validation objective in exercise TVAL.02.3-AWARE-ASA-TRL2	Exercise validation objective	Exercise success criteria
Assess the increase in Capacity due to the AWARE solution.	AWARE leads to increased Capacity.	Fully covered	Measure ASA system's impact on en-route throughput and traffic handling capability during peak demand hours	ASA system assists ATCOs in maintaining or increasing throughput without performance degradation.
Assess the improvement in Cost Efficiency due to the AWARE solution.	AWARE leads to increased Cost Efficiency.	Fully covered	Assess ASA system's effect on flights handled per ATCO-hour, improving resource allocation.	ASA system results in a higher ratio of flights per ATCO-hour, optimising operational costs.
Assess the improvement in Fuel Efficiency due to AWARE solution.	AWARE leads to reduction in fuel consumption	Fully covered	Evaluate ASA system's role in reducing fuel consumption and CO_2 emissions by optimising flight management.	ASA system reduces average CO_2 emissions per flight compared to baseline operations.
Assess the improvement of Safety due to the AWARE solution.	AWARE leads to fewer conflicts, resulting in increased safety.	Fully covered	Analyse the reduction in conflict events and overall safety risks with ASA system support during peak traffic.	ASA system support lowers the number of conflict-induced safety incidents.
Assess the impact on the required Human Performance levels due to AWARE solution.	AWARE leads to lower workload levels.	Fully covered	Assess the reduction in ATCO workload and stress levels with ASA system assistance during high traffic.	ATCOs report lower workload and improved task performance with ASA system support.

EUROPEAN PARTNERSHIP
Co-funded by the European Union

SESAR solution validation objective	SESAR solution success criteria	Coverage and comments on the coverage of SESAR solution validation objective in exercise TVAL.02.3-AWARE-ASA-TRL2	Exercise validation objective	Exercise success criteria
Assess the improvement on the required Operational Efficiency due to the AWARE solution.	AWARE leads to increased Operational Efficiency.	Fully covered	Evaluate ASA system's effect on operational fuel efficiency during high-traffic scenarios.	ASA system improves fuel efficiency, reducing the average fuel burn per flight.

Table 34. Validation objectives addressed in validation exercise TVAL.02.3-AWARE-ASA-TRL2

5.4.4 Validation scenarios

The validation scenario TVAL.02.3-AWARE-ASA-TRL2 focuses on assessing the ASA system's support capabilities in a controlled air traffic management environment. Conducted on the Polaris ATM platform with embedded ASA functionalities, this scenario simulates high traffic levels, reflective of peak operational conditions. Exercise will be run with the ASA system providing varying support to ATCOs. Each experimental group will include five ATCOs, ensuring that every support level is performed at least five times. This allows for comparison of performance, efficiency, and safety between baseline operations and ASA-supported operations. The primary aim is to gather quantitative and qualitative data on KPAs such as capacity, cost-efficiency, environmental impact, safety, human performance, and operational efficiency, aligning with the SESAR 3 framework's targeted performance improvements. Data collected through HITL simulations will inform the assessment of ASA's contributions to performance metrics, aiding in future implementations and developments across ATC environments.

5.4.4.1 Reference scenario(s)

Group 1 of the exercise TVAL.02.3-AWARE-ASA-TRL2 sets a baseline conditions, enabling a comparative analysis against the ASA-supported solution scenario. This simulation leverages existing ATC processes under high traffic demands, capturing performance across KPAs without additional technological support. The aim is to understand how standard procedures handle capacity constraints, cost-efficiency, environmental impact, and human performance metrics, specifically focusing on ATCO workload and situational awareness in unassisted settings. The findings from this scenario establish the benchmark for evaluating the ASA's impact, enabling a structured comparison that highlights the extent of ASA's performance enhancements. This approach aligns with SESAR guidelines for validation, ensuring a controlled and repeatable baseline.

EUROPEAN PARTNERSHIP
Co-funded by the European Union

5.4.4.2 Solution scenario(s)

The solution scenario incorporates the ASA system at varying level of ASA support into the Polaris ATM environment, supporting ATCOs with task management and decision-making under high traffic conditions with specific focus on traffic interaction with restricted military areas. Based on the level of support, the ASA system can actively assists ATCOs in high-stress environments by automating routine tasks and augmenting situational awareness, thereby reducing cognitive load. Key metrics, including workload reduction, improved response times, and enhanced safety outcomes, will be measured and compared against the reference scenario. Expected benefits in operational efficiency, safety, and environmental performance will also be analysed, with the results contributing to TRL advancement and SESAR objectives. The data from this scenario are intended to validate ASA's readiness for higher TRL stages, aligning with the SESAR Performance Framework's objectives for future ATM system evolution

5.4.5 Exercise validation assumptions

Assumption ID	Assumption title	Assumption description	Justification	Impact Assessment
VA2.1	ASA System Intermediate Support	The ASA system will provide intermediate support to ATCOs but will not perform any autonomous actions.	This assumption ensures that the focus remains on how ASA assists ATCOs without taking over any tasks autonomously, aligning with the experiment's objectives.	A direct impact on workload and situational awareness results is predicted
VA2.2	ASA System Autonomous Support	The ASA system will provide fully autonomous support to ATCOs.	The focus will be on how ASA performs autonomous tasks, and how it will choose which tasks to take over from the ATCO.	A direct impact on workload and situational awareness results is predicted

VA2.3	Stable Simulation Environment	The simulation environment will remain stable without introducing unexpected disruptions (e.g., technical issues).	A stable environment is necessary to ensure accurate data collection and isolate the impact of the ASA system's intermediate support.	High: Any disruptions may compromise the reliability of results
VA2.4	ATCO Familiarity with ASA System	ATCOs participating in the exercise have received adequate training on the ASA system's intermediate functions.	Ensures ATCOs are comfortable using the ASA system, minimizing variability in results due to unfamiliarity.	Moderate: Insufficient training could affect performance outcomes.

Table 35. Validation exercise TVAL.02.3-AWARE-ASA-TRL2 assumptions

5.4.6 Limitations and impact on the level of significance

The following limitations may impact the level of significance of the validation results:

- The simulation environment cannot fully replicate the variability and unpredictability of real-world air traffic operations.
- HITL simulations may exclude certain dynamic factors that shape ATCO decision-making during peak traffic, such as sudden weather disruptions or complex traffic flow deviations.
- Although the ASA system operates at EASA level 2B, it remains at an early development stage (TRL2), which limits the extent to which conclusions about its operational effectiveness can be drawn.
- A limited number of ATCO participants may impact the representativeness of the findings, making it difficult to extrapolate to a broader population.
- The introduction of a new operational environment (Polaris and the ASA system), combined with the absence of a planner position, may influence ATCO situational awareness, operational efficiency, and workload.
- Sequence effects, such as learning or fatigue, may occur over the course of the experiment, despite mitigation efforts, potentially affecting the reliability of participant feedback.

While the results will offer valuable insights, additional real-world validation will be required before the findings can be generalized to larger operational contexts, such as the ECAC area.

5.4.7 Validation exercise platform / tool and validation technique

5.4.7.1 Validation exercise platform / tool characteristics

The validation exercise will be conducted on the POLARIS ATM system, enhanced with ASA functionality to support varying level of ASA support. The exercise will be conducted in a simulation

EUROPEAN PARTNERSHIP
Co-funded by the European Union

environment at a designated validation site, with HITL simulations enabling real-time interaction between ATCOs and the ASA system. The POLARIS system is equipped to simulate high traffic conditions with augmented ASA system, ensuring the validation fucuses.

5.4.7.2 Validation exercise technique

The validation exercise utilises a HITL simulation technique, enabling real-time interactions between ATCOs and the ASA-augmented Polaris ATM system under medium to high traffic conditions. HITL simulations are selected to closely replicate operational environments, allowing for controlled assessments of ATCO performance, situational awareness, and workload with ASA support. This approach provides high-fidelity insights into human-machine collaboration, as ATCOs are exposed to varied traffic scenarios while receiving ASA guidance and support. Data from these sessions, including qualitative feedback and quantitative performance metrics, will be analysed to evaluate ASA's effectiveness across KPAs, ensuring rigorous assessment against SESAR 3 performance objectives.

5.4.8 Data collection and analysis

5.4.8.1 Data and data collection methods

The validation exercise will collect both quantitative and qualitative data. Quantitative data will include metrics such as eye tracking data, traffic throughput, conflict resolution, fuel consumption, and ATCO workload levels, while qualitative data will focus on ATCO feedback through post-simulation questionnaires assessing workload, situational awareness, and system usability. Data collection will be performed through system logs, real-time monitoring, and surveys. Multiple simulation iterations are planned to ensure robust data collection, and all data sets (including traffic data, system performance logs, and ATCO inputs) will be processed and cleaned to ensure consistency and accuracy. Output formats will include system log files, video recordings, and questionnaire data. Financial resources are sufficient for data collection, and proprietary and validated data sources will be used to meet research needs.

5.4.8.2 Analysis methods

The collected data will be analysed using both quantitative and qualitative methods to assess the impact of ASA system support on ATCO performance and operational efficiency. Quantitative data will be compared between the baseline (referred to Group 1) and solution scenarios (referred to Group 2-4) to identify any significant improvements in key performance indicators such as traffic throughput and safety. Qualitative data, such as ATCO feedback, will be analysed to identify trends in workload and system usability. Post-processing will involve cleaning and organising the data for consistency, ensuring accurate analysis. These methods will help determine whether the ASA system meets the success criteria and addresses the research objectives.

5.4.9 Exercise planning and management

5.4.9.1 Activities

Preparatory phase (PREP):

- Setting up the validation infrastructure.
- Integrating the ASA system with the POLARIS ATM platform.

- Conducting training sessions for ATCOs for working with the AI.

<u>**Execution phase (EXEC):**</u>

- Running the HITL simulations under high traffic loads with varying level of ASA support.
- Monitoring ATCO performance and system functionality in real time.

<u>**Post-execution phase (POST-EXEC):**</u>

- Data processing: All of the data collected need to be cleaned and analysed.
- Preparing the exercise report.

5.4.9.2 Roles and responsibilities in the exercise

In validation exercise TVAL.02.3-AWARE-ASA-TRL2, several consortium members are involved and hold distinct responsibilities essential for the preparation, execution, and analysis of the exercise.

ZHAW will oversee the development, implementation, and analysis of the interviews and questionnaires.

LFV will provide both the ATCO participants, and the physical space required for conducting the workshop.

UkSATSE and IFATCA will contribute additional ATCOs, enhancing the exercise's diversity by ensuring that participants bring a range of operational expertise.

TERN and FTTS will supply technical support, focusing on the integration of the ASA system into the Polaris platform.

5.4.9.3 Time planning

Activity	Week	
	1	2
Setting up validation infrastructure (PREP)	■	
Integrating the ASA system with the POLARIS ATM platform (PREP)	■	
Conducting training sessions for ATCOs (PREP)	■	■
Running HITL simulations with high traffic load and data collection (EXEC)	■	■
Evaluating simulation outcomes (POST-EXEC)		■
Preparing the exercise report (POST-EXEC)		■

Table 36. Detailed exercise TVAL.02.3-AWARE-ASA-TRL2 time planning

5.4.9.4 Identified risks and mitigation actions

Risks	Impact (1-low, 2-medium, 3-high)	Likelihood (1-low, 2-medium, 3-high)	Criticality (calculated based on likelihood and impact)	Mitigation actions
Ukrainian ATCOs cannot leave Ukraine or UkSATSE becomes unavailable.	2	2	2	LFV and IFATCA provide more controllers and higher input than previously planned. In extreme case, simulation campaign is halved.
Inability to assess ATCO intent from traffic situation semantics, attention tracking, and HMI tracking	3	1	1	ATCO intent will be either partially mapped or mapped to a smaller subset of parameters to enable the continuation of research. An additional deliverable should be prepared to provide detailed descriptions of methodology, encountered problems, and possible avenues of future research.

Table 37. Exercise TVAL.02.3-AWARE-ASA-TRL2 risks and mitigation actions, [2]

5.5 Validation exercise TVAL.02.4-AWARE-ASA-TRL2 plan

5.5.1 Validation exercise description and scope

The final validation exercise will evaluate the performance of the ASA system with varying level of support, with the goal of supporting ATCOs in managing a 25% increase in peak traffic, reflective of projected future traffic levels. ATCOs serve as the primary operational actors and are divided into four groups, each working with a different level of ASA assistance. The exercise aims to determine whether future high-traffic conditions can be effectively managed with ASA support, and how this compares to ATCO performance without any ASA support.

The scenario incorporates complex conflict situations, interactions with restricted military areas, pilot non-conformances, and multiple requests from pilots and adjacent sectors, ensuring a realistic and challenging environment. The exercise covers use cases U.C.1, U.C.2, U.C.3, and U.C.4. Total of 20 ATCOs divided into four groups will perform this exercise. The reference airspace is the Malmö ACC Sector 23 (ESMM23), which controls en-route traffic. The scenario is based on actual traffic from 4th of August 2024.

Key validation objectives include assessing the ASA system's impact on capacity, operational efficiency, safety, and ATCO workload reduction under significantly higher traffic demands. The exercise will also evaluate ATCO situational awareness, ensuring that controllers remain effectively in the loop while the system performs background tasks. This assessment is essential to determine whether ATCOs can maintain awareness and performance when managing traffic levels 25% above current peak occupancy with ASA support.

The exercise is applicable to TRL2, as it aims to validate the system in a simulated operational environment, moving closer to real-world deployment readiness.

5.5.2 Stakeholder's expectations and benefit mechanisms addressed by the exercise

Stakeholder	Involvement	Why it matters to the stakeholder
ATCOs	ATCOs will need to handle a 25% increase in traffic. Groups with ASA support could relying on ASA system support or task execution.	This exercise is critical for ATCOs to evaluate how the ASA system can reduce their workload, assist with future traffic demands, and maintain safety and efficiency in high-traffic scenarios. It also demonstrates the system's scalability and impact on future traffic operations.
SUPs	SUPs oversee the coordination between ATCOs and the ASA system, ensuring smooth integration.	Supervisors will assess how well the ASA system can support ATCOs in managing future traffic scenarios. The outcome will help them determine whether it enhances oversight and operational safety while reducing ATCO stress and improving decision-making during high traffic peaks.
ATSEPs	ATSEPs will monitor the ASA system's integration with the existing ATM infrastructure during peak loads.	This exercise matters for ATSEPs as it tests the system's robustness and its capability to handle complex traffic situations without introducing technical faults, ensuring system reliability, cybersecurity, and operational safety for future traffic demands.
ANSPs	ANSPs will be involved in assessing overall operational performance and capacity during the simulation.	ANSPs are interested in whether the ASA system can effectively increase airspace capacity and operational efficiency without compromising safety, helping them prepare for future traffic growth while optimising resource use and cost-efficiency.
Airspace Users	Airspace Users will provide feedback on the efficiency and timeliness of flights under increased traffic.	Airspace users expect the ASA system to enable better traffic flow management, which could reduce delays and operational costs, resulting in more fuel-efficient flights and lower environmental impact under higher traffic conditions.

Stakeholder	Involvement	Why it matters to the stakeholder
ATM Research Community	The ATM research community will collect and analyse performance data, focusing on human-machine collaboration under high traffic loads.	This exercise is important to the ATM research community for testing the ASA system's ability to manage future traffic growth and assessing its scalability. The findings will inform future developments in AI-driven air traffic management solutions.

Table 38. Stakeholders' involvement and expectations for exercise TVAL.02.4-AWARE-ASA-TRL2

5.5.3 Validation objectives

SESAR solution validation objective	SESAR solution success criteria	Coverage and comments on the coverage of SESAR solution validation objective in exercise TVAL.02.4-AWARE-ASA-TRL2	Exercise validation objective	Exercise success criteria
Assess the increase in Capacity due to the AWARE solution	AWARE leads to increased Capacity.	Fully covered	Measure ASA system's impact on en-route throughput and traffic handling capability during peak demand hours	ASA system assists ATCOs in maintaining or increasing throughput without performance degradation.
Assess the improvement in Cost Efficiency due to the AWARE solution.	AWARE leads to increased Cost Efficiency.	Fully covered	Assess ASA system's effect on flights handled per ATCO-hour, improving resource allocation.	ASA system results in a higher ratio of flights per ATCO-hour, optimising operational costs.
Assess the improvement in Fuel Efficiency due to AWARE solution.	AWARE leads to reduction in fuel consumption.	Fully covered	Evaluate ASA system's role in reducing fuel consumption and CO_2 emissions by optimising flight management.	ASA system reduces average CO_2 emissions per flight compared to baseline operations.

Co-funded by the European Union

SESAR solution validation objective	SESAR solution success criteria	Coverage and comments on the coverage of SESAR solution validation objective in exercise TVAL.02.4-AWARE-ASA-TRL2	Exercise validation objective	Exercise success criteria
Assess the improvement of Safety due to the AWARE solution.	AWARE leads to fewer conflicts, resulting in increased safety.	Fully covered	Analyse the reduction in conflict events and overall safety risks with ASA system support during peak traffic.	ASA system support lowers the number of conflict-induced safety incidents.
Assess the impact on the required Human Performance levels due to AWARE solution.	AWARE leads to lower workload levels.	Fully covered	Assess the reduction in ATCO workload and stress levels with ASA system assistance during high traffic.	ATCOs report lower workload and improved task performance with ASA system support.
Assess the improvement on the required Operational Efficiency due to the AWARE solution.	AWARE leads to increased Operational Efficiency.	Fully covered	Evaluate ASA system's effect on operational fuel efficiency during high-traffic scenarios.	ASA system improves fuel efficiency, reducing the average fuel burn per flight.

Table 39. Validation objectives addressed in validation exercise TVAL.02.4-AWARE-ASA-TRL2

5.5.4 Validation scenarios

The validation scenario for Exercise TVAL.02.4-AWARE-ASA-TRL2 involves simulating a 25% increase in peak traffic occupancy, representing the highest traffic demand in all experiments. This scenario evaluates the ASA system's capabilities under projected future traffic conditions. ATCOs in the validation group with level of ASA support 2B could rely on the system to autonomously manage routine tasks such as frequency transfers, direct routing or sector coordination, ATCOs can focus on complex decision-making. This setup will assess the ASA system's effectiveness in enhancing capacity, operational efficiency, safety, and workload management in response to future airspace demands.

5.5.4.1 Reference scenario(s)

The reference scenario for this validation exercise will be drawn from the Group 1 of the TVAL.02.4-AWARE-ASA-TRL2 where ATCOs managed peak traffic levels without any ASA system support. This approach will provides a valuable point of comparison to assess how the ASA system impacts key performance areas, such as capacity, operational efficiency, and ATCO workload, under typical traffic conditions.

EUROPEAN PARTNERSHIP
Co-funded by the European Union

Although Group 1 will be required to perform a scenario with traffic count that exceeds operational limitations, this will be valuable insight into whether, and to what extent, the ASA system can improve situational awareness, reduce workload and improve other KPA's when applying 2A or 2B level of support. This experiment will help determine whether Human–AI teaming can raise existing operational limits by maintaining or improving performance relative to current validation objectives.

5.5.4.2 Solution scenario(s)

The solution scenario for Validation Exercise TVAL.02.4-AWARE-ASA-TRL2 mirrors the setup of previous experiment in terms of airspace configuration, operational parameters, and ASA system support level. The key difference is the introduction of a 25% increase in peak traffic, combined with different high complexity elements such as conflict resolution, interactions with restricted military area, pilot errors, pilot requests, and requests from adjacent sectors.

5.5.5 Exercise validation assumptions

Assumption ID	Assumption title	Assumption description	Justification	Impact Assessment
VA2.1	ASA System Intermediate Support	The ASA system will provide intermediate support to ATCOs but will not perform any autonomous actions.	This assumption ensures that the focus remains on how ASA assists ATCOs without taking over any tasks autonomously, aligning with the experiment's objectives.	A direct impact on workload and situational awareness results is predicted
VA2.2	ASA System Autonomous Support	The ASA system will provide fully autonomous support to ATCOs.	The focus will be on how ASA performs autonomous tasks, and how it will choose which tasks to take over from the ATCO.	A direct impact on workload and situational awareness results is predicted

Assumption ID	Assumption title	Assumption description	Justification	Impact Assessment
VA2.3	Stable Simulation Environment	The simulation environment will remain stable without introducing unexpected disruptions (e.g., technical issues).	A stable environment is necessary to ensure accurate data collection and isolate the impact of the ASA system's full support.	High: Any disruptions may compromise the reliability of results
VA2.4	ATCO Familiarity with ASA System	ATCOs participating in the exercise have received adequate training on the ASA system's intermediate functions.	Ensures ATCOs are comfortable using the ASA system, minimizing variability in results due to unfamiliarity.	Moderate: Insufficient training could affect performance outcomes.
VA5.1	25% Increase as Representative	The 25% traffic increase accurately represents expected future air traffic demand.	This assumption supports the use of the scenario to simulate realistic future conditions, enhancing the relevance of findings.	Moderate: Impacts the validity of future scalability analysis.

Table 40. Validation exercise TVAL.02.4-AWARE-ASA-TRL2 assumptions

5.5.6 Limitations and impact on the level of significance

Validation Exercise TVAL.02.4-AWARE-ASA-TRL2, simulating a 25% increase in peak traffic, has several limitations impacting the significance and generalizability of the results:

- The simulation environment cannot fully replicate the variability and unpredictability of real-world air traffic operations.
- HITL simulations may exclude certain dynamic factors that shape ATCO decision-making during peak traffic, such as sudden weather disruptions or complex traffic flow deviations.
- A small, specialized group of trained ATCOs may not fully represent broader ATCO demographics, impacting generalizability.
- Although the ASA system operates at EASA level 2B, it remains at an early development stage (TRL2), which limits the extent to which conclusions about its operational effectiveness can be drawn.

- The introduction of a new operational environment (Polaris and the ASA system), combined with the absence of a planner position, may influence ATCO situational awareness, operational efficiency, and workload.
- Although scenarios are randomized to reduce bias, some residual learning or fatigue effects may influence ATCO performance.
- Limited simulation time restricts assessment of sustained performance under prolonged high-traffic conditions.
- While the 25% traffic increase aligns with projections, variations in actual future conditions may affect scalability at the ECAC level.

These limitations suggest that while insights into ASA system performance are valuable, further validation in operational environments is necessary for broader application.

5.5.7 Validation exercise platform / tool and validation technique

5.5.7.1 Validation exercise platform / tool characteristics

Validation Exercise TVAL.02.4-AWARE-ASA-TRL2 will use the POLARIS ATM system integrated with the ASA system to simulate en-route airspace with a 25% increase in peak traffic. Each Group will manage traffic with specifically defined level of ASA support. ATCOs in Groups 3 and 4 will be offered the system in which routine task (such as frequency transfers, quality of service tasks or sector coordination) are handled autonomously. This allows ATCOs to focus on higher-level decision-making under high-demand conditions.

5.5.7.2 Validation exercise technique

The chosen validation technique for this validation exercise is a HITL simulation. This technique enables ATCOs to interact directly with the ASA system in a controlled, high-fidelity simulation environment, allowing for real-time assessment of the system's impact on ATCO workload, operational efficiency, and capacity management under a 25% increase in peak traffic.

HITL simulation is particularly suited to this exercise, as it replicates realistic air traffic scenarios while providing detailed insights into human-system interactions. This approach enables the accurate measurement of the ASA system's effectiveness in supporting ATCOs during peak traffic loads, meeting the exercise's validation objectives by capturing data on both system performance and human factors.

5.5.8 Data collection and analysis

5.5.8.1 Data and data collection methods

This validation exercise will collect both quantitative and qualitative data to assess the ASA system's impact on ATCO workload, operational efficiency, safety, and capacity in a high-traffic environment. Quantitative data will capture metrics such as conflict resolution rates, frequency transfers, and workload scores, while qualitative data will include ATCO feedback on system usability and workload perceptions.

Data collection methods include system logs, real-time monitoring, and post-exercise questionnaires. Quantitative data will be analysed to identify performance trends, while qualitative data will undergo

Co-funded by the European Union

thematic analysis to capture key insights from ATCO feedback. Data will be gathered in multiple iterations to strengthen the reliability of findings, with all data processed and cleaned for consistency. Recorded outputs will include file logs and video captures, providing a detailed view of ASA system performance.

5.5.8.2 Analysis methods

The analysis for Validation Exercise TVAL.02.4-AWARE-ASA-TRL2 will involve both quantitative and qualitative approaches to assess the ASA system's effectiveness in meeting success criteria under high traffic conditions. Quantitative data, such as conflict resolution rates and workload scores, will be analysed to identify trends and measure performance improvements and situational awareness against baseline metrics. Qualitative feedback from ATCOs will undergo thematic analysis to reveal insights on system usability and perceived workload reduction.

Post-processing will include cleaning and normalizing the data to ensure accuracy and consistency. The combined analysis will address key research questions regarding the ASA system's impact on capacity, operational efficiency, and ATCO workload, providing a comprehensive assessment of its performance under increased traffic demands.

5.5.9 Exercise planning and management

5.5.9.1 Activities

This validation exercise is split into three phases of activities:

Preparatory phase (PREP):

- Setting up the validation infrastructure
- Integrating the ASA system with the POLARIS ATM platform.
- Conducting training sessions for ATCOs on using the ASA system

Execution phase (EXEC):

- Running the HITL simulations with a 25% increase of peak traffic
- Monitoring ATCO performance and system functionality in real time

Post-execution phase (POST-EXEC):

- Data processing: All the data collected needs to be cleaned and analysed.
- Preparing the exercise report

5.5.9.2 Roles and responsibilities in the exercise

In validation exercise TVAL.02.4-AWARE-ASA-TRL2, several consortium members are involved and hold distinct responsibilities essential for the preparation, execution, and analysis of the exercise.

ZHAW will oversee the development, implementation, and analysis of the interviews and questionnaires.

LFV will provide both the ATCO participants, and the physical space required for conducting the exercise.

UkSATSE and IFATCA will contribute additional ATCOs, enhancing the exercise's diversity by ensuring that participants bring a range of operational expertise.

TERN and FTTS will supply technical support, focusing on the integration of the ASA system into the Polaris platform.

5.5.9.3 Time planning

Activity	Week	
	1	2
Setting up validation infrastructure (PREP)		
Integrating the ASA system with the POLARIS ATM platform (PREP)		
Conducting training sessions for ATCOs (PREP)		
Running HITL simulations with a 25% increase of peak traffic and data collection (EXEC)		
Evaluating simulation outcomes (POST-EXEC)		
Preparing the exercise report (POST-EXEC)		

Table 41. Detailed exercise TVAL.02.4-AWARE-ASA-TRL2 time planning

5.5.9.4 Identified risks and mitigation actions

Risks	Impact (1-low, 2-medium, 3-high)	Likelihood (1-low, 2-medium, 3-high)	Criticality (calculated based on likelihood and impact)	Mitigation actions
Ukrainian ATCOs cannot leave Ukraine or UkSATSE becomes unavailable.	2	2	2	LFV and IFATCA provide more controllers and higher input than previously planned. In extreme case, simulation campaign is halved.

Risks	Impact (1-low, 2-medium, 3-high)	Likelihood (1-low, 2-medium, 3-high)	Criticality (calculated based on likelihood and impact)	Mitigation actions
Inability to assess ATCO intent from traffic situation semantics, attention tracking, and HMI tracking	3	1	1	ATCO intent will be either partially mapped or mapped to a smaller subset of parameters to enable the continuation of research. An additional deliverable should be prepared to provide detailed descriptions of methodology, encountered problems, and possible avenues of future research.

Table 42. Exercise TVAL.02.4-AWARE-ASA-TRL2 risks and mitigation actions, [2]

6 References

6.1 Applicable documents

This ERP complies with the requirements set out in the following documents:

Validation

[1] SESAR Joint Undertaking, DES HE requirements and validation/demonstration guidelines, SESAR Joint Undertaking library, Edition 03.00, 2023

Project and programme management

[2] [101167442] AWARE Grant Agreement, [25/06/2024]

[3] SESAR 3 JU Project Handbook – Programme Execution Framework, [11/04/2022], [1.00]

6.2 Reference documents

[4] SESAR Joint Undertaking, European ATM Master Plan, Luxembourg, Publications Office of the European Union, 2019.

[5] CRIDA/ENAIRE, AIRBUS, BOEING, DSNA, DEEPBLUE, DLR, ECTL, ENAV, INDRA-ALG, ONB4, TELEPAZIO, DES Performance Framework, SESAR Joint Undertaking library, Editon 00.01.04, 2023.

[6] ON (B4), EUROCONTROL, AIRBUS, ENAIRE, TELESPAZIO, DES Common Assumptions, SESAR Joint Undertaking library. Edition 00.02.01, 2023

[7] AWARE consortium, AWARE Economic Evaluation (ECO-EVAL)

[8] AWARE consortium, AWARE Operational Services and Environment Description (OSED)