D6.1 - CODA - Exploratory Research Plan (ERP)

Deliverable ID:	**D6.1**
Project acronym:	**CODA**
Grant:	**101114765**
Call:	**HORIZON-SESAR-2022-DES-ER-01**
Topic:	**HORIZON-SESAR-2022-DES-ER-01-WA2-1**
Consortium coordinator:	**DBL**
Edition date:	**05 July 2024**
Edition:	**00.02.00**
Status:	**Official**
Classification:	**PU**

Abstract

This document serves as the Exploratory Research Plan for the CODA project (Deliverable 6.1).

It describes the experimental approach, validation objectives, and methods that are planned in the context of the CODA project for performing the foreseen validation activities. Any proposed validation activities, regardless of the simulation technique, are planned to address stakeholders' needs and assess the expected benefits.

This document is the final edition of the CODA TRL 2 Validation Plan.

Authoring & approval

Author(s) of the document

Organisation name	Date
DBL	30/05/2024
BS	25/05/2024
CRIDA	25/05/2024
NLR	25/05/2024
ENAC	25/05/2024

Reviewed by

Organisation name	Date
BS	29/05/2024
IFATCA	29/05/2024
CRIDA	29/05/2024
UPM	29/05/2024
UGR	28/05/2024
ECTL	27/05/2024
NLR	31/05/2024

Approved for submission to the SESAR 3 JU by[1]

Organisation name	Date
BS	29/05/2024
IFATCA	29/05/2024
CRIDA	29/05/2024
UPM	29/05/2024
UGR	29/05/2024
DBL	29/05/2024
ECTL	29/05/2024
ENAC	29/05/2024
NLR	31/05/2024

Rejected by[2]

Organisation name	Date

Document history

Edition	Date	Status	Organisation author	Justification
00.00.01	01/03/2024	Draft	DBL	First Draft

[1] Representatives of all the beneficiaries involved in the project

[2] Representatives of the beneficiaries involved in the project

00.00.02	02/04/2024	Draft	DBL, CRIDA, UPM, BS, DLR	Updated version with input for EXE2 and EXE3
00.00.03	06/05/2024	Draft	DBL, CRIDA, UPM, BS, DLR, UGR, ENAC, ECTL, IFACTA	Updated version after partner's review
00.01.00	30/05/2024	Final	DBL	Final version
00.02.00	05/07/2024	Final	DBL	Sesar comments addressed. 2nd SJU submission

Copyright Statement

Disclaimer

CODA

COntroller adaptive Digital Assistant

CODA

Tables of contents

Abstract 1
1 Executive summary 9
2 Introduction 10
2.1 Purpose of the document 10
2.2 Intended readership 10
2.3 Background 10
2.4 Structure of the document 10
2.5 Glossary of terms 11
2.6 List of acronyms 11
3 Context of the experimental research plan 14
3.1 Experimental Research Plan Context 14
3.2 Scope 15
3.3 Key R&I needs 15
3.4 Estimated performance contributions 16
3.5 Initial and exit maturity levels 17
4 Experimental plan 18
4.1 Experimental plan approach 18
4.2 Stakeholders' expectations and involvement 20
4.3 Validation Objectives 21
4.4 Validation assumptions 25
4.5 Validation exercises list 26
4.6 Validation exercises planning 28
4.7 Deviations with respect to the SESAR 3 JU project handbook 29
5 Validation exercises 30
5.1 Validation exercise #01 plan - Validation of the CODA operational concept, use cases and scenarios 30
5.1.1 Validation exercise description and scope 30
5.1.2 Stakeholder's expectations and benefit mechanisms addressed by the exercise 31
5.1.3 Validation objectives 32
5.1.4 Validation scenarios 33
5.1.4.1 Reference scenario(s) 34
5.1.4.2 Solution scenario(s) 34
5.1.5 Exercise Validation Assumptions 34
5.1.6 Limitations and impact on the level of significance 34

5.1.7 Validation exercise platform/tool and validation technique 35
5.1.7.1 Validation exercise platform/tool characteristics 35
5.1.7.2 Validation exercise technique 36
5.1.8 Data collection and analysis 37
5.1.8.1 Data and data collection methods 37
5.1.8.2 Analysis methods 37
5.1.9 Exercise planning and management 37
5.1.9.1 Activities 37
5.1.9.2 Roles and responsibilities in the exercise 37
5.1.9.3 Time planning 38
5.1.9.4 Identified risks and mitigation actions 38

5.2 Validation exercise #02 plan - Validation of the mental states' prediction models 40
5.2.1 Validation description and scope 40
5.2.2 Stakeholder's expectations and benefit mechanisms addressed by the exercise 40
5.2.3 Validation Objectives 41
5.2.4 Validation scenarios 42
5.2.4.1 Airspace information 42
5.2.4.2 Solution scenario(s) 44
5.2.5 Exercise Validation Assumptions 45
5.2.6 Limitations and impact on the level of significance 45
5.2.7 Validation exercise platform/tool and validation technique 46
5.2.7.1 Validation exercise platform/tool characteristics 46
5.2.7.2 Validation exercise technique 48
5.2.8 Data collection and analysis 48
5.2.8.1 Data and data collection methods 48
5.2.8.2 Analysis methods 50
5.2.9 Exercise planning and management 50
5.2.9.1 Activities 50
5.2.9.2 Roles and responsibilities in the exercise 51
5.2.9.3 Time planning 51
5.2.9.4 Identified risks and mitigation actions 52

5.3 Validation exercise #03 plan - Final Validation of the CODA system 53
5.3.1 Validation exercise description and scope 53
5.3.2 Stakeholder's expectations and benefit mechanisms addressed by the exercise 53
5.3.3 Validation objectives 54
5.3.4 Validation scenarios 55
5.3.4.1 Airspace information 56
5.3.4.2 Reference scenario(s) 56
5.3.4.3 Solution scenario(s) 56
5.3.5 Exercise Validation Assumptions 57
5.3.6 Limitations and impact on the level of significance 57
5.3.7 Validation exercise platform/tool and validation technique 58
5.3.7.1 Validation exercise platform/tool characteristics 58
5.3.7.2 Validation exercise technique 59
5.3.8 Data collection and analysis 59
5.3.8.1 Data and data collection methods 59
5.3.8.2 Analysis methods 60
5.3.8.3 Exercise planning and management 60
5.3.8.4 Activities 61
5.3.8.5 Roles and responsibilities in the exercise 61

5.3.8.6 Time planning....62
5.3.8.7 Identified risks and mitigation actions....62

6 References....64

6.1 Applicable documents....64

6.2 Reference documents....64

List of Figures

Figure 1: validation and pre-validation exercises foreseen in the project....19

Figure 2: STELLAR - validation exercises....28

Figure 3: example of Mentimeter feedback layout in exercise #01....30

Figure 4: simplified overview of the CODA solution....33

Figure 5: detailed overview of the CODA solution....34

Figure 6: example of the slide presented in exercise #01....35

Figure 7: example of the Mentimeter Interface in exercise #01....36

Figure 8: Madrid ACC elementary sectors. Source: Insignia....43

Figure 9: exercise #02 simulated sector....43

Figure 10: exercise #02 main traffic flows....44

Figure 11: ESCAPE simulator....47

Figure 12: ESCAPE - CWP....47

Figure 13: ESCAPE – IPAS and Ground Server....47

Figure 14: ESCAPE- pseudo-pilots....47

Figure 15: ESCAPE- CWP....48

Figure 16 Madrid ACC elementary sectors. Source: Insignia....56

List of tables

Table 1: glossary of terms....11

Table 2: list of acronyms....13

Table 3: CODA R&I 16
Table 4: maturity levels table 17
Table 5: list of validation activities 18
Table 6: list of pre-validation activities 19
Table 7: stakeholders' expectations and involvement 20
Table 8: validation objectives 21
Table 9: complete list of detailed validation objectives and success criteria 23
Table 10: validation exercises variables 24
Table 11: validation assumptions overview 26
Table 12: validation exercise #01 layout 26
Table 13: validation exercise #02 layout 27
Table 14: validation exercise #03 layout 28
Table 15: Advisory Board Meeting CODA partners-participants list (online) in exercise #01 31
Table 16: stakeholders' expectations exercise #01 32
Table 17: validation objectives addressed in validation exercise #01 33
Table 18: exercise #01 Planning activities 37
Table 19: exercise #01 Role and Responsibilities 38
Table 20: detailed exercise #01-time planning 38
Table 21: exercise #01 Risks and mitigation actions 39
Table 22: stakeholders' expectations exercise #02 41
Table 23: validation objectives addressed in exercise #02 42
Table 24: validation exercise #02 assumptions 45
Table 25: exercise #02 data collection 49
Table 26: exercise #02-time planning 50
Table 27: exercise #02 Planning activities 51
Table 28: exercise #02 Role and Responsibilities 51
Table 29: detailed exercise #02-time planning 52

Table 30: exercise #02 risks and mitigation actions (1-low, 2-medium, 3-high)..... 52

Table 31: stakeholders' expectations exercise #03 54

Table 32: Validation objectives exercise #03 55

Table 33: exercise #03 data collection..... 59

Table 34: exercise #03-time planning 61

Table 35: Exercise #03 Planning activities 61

Table 36: Exercise #03 Roles and Responsibilities 62

Table 37: detailed exercise #03-time planning 62

Table 38: exercise #03 risks and mitigation actions (1-low, 2-medium, 3-high)..... 63

1 Executive summary

This Exploratory Research Plan (ERP) provides detailed information on the validation approach and activities planned to advance the CODA project from Technology Readiness Level (TRL) 1 to TRL 2.

The overall objective of the CODA project is **to increase ATM's efficiency, capacity, and safety**, maximizing human-AI teaming by developing a system in which tasks are performed collaboratively by hybrid human-machine teams and dynamically allocated through adaptive automation principles. To do so CODA focuses on developing a solution that predicts relevant mental states of en-route air traffic controllers to anticipate possible problems and trigger specific actions (such as the activation of Digital Assistants).

The experimental research is supported by a **User-Centred Design** iterative approach to the development of the CODA solution.

Three exercises will be organized and performed at different locations to evaluate technical feasibility, integrate prototypes, and assess operational acceptance, feasibility, and performance. Below is the list of the validation exercises planned for the CODA project:

- EXE#01 - **Validation of the CODA operational concept, use cases, and scenarios**, WORKSHOP, (M4) coordinated by DBL
- EXE#02 - **Validation of the mental states prediction models**, Small-Scale Simulation (RTS), (M14) coordinated by CRIDA
- EXE#03 – **Final Validation of the CODA system**, Human-in-the-loop Simulation (RTS), (M19) coordinated by DBL

The validation activities will be conducted throughout the development lifecycle of the solution, starting with the definition of use cases and scenarios. These activities will progress step-by-step, ensuring the **validation of each component of the CODA system individually before integrating them into the complete system.**

Therefore, in addition to the three main validation exercises, other **pre-validation activities** (called “Tests”) have been scheduled to validate specific aspects of the concepts. These include workshops with experts, mock-ups, and technical verifications. In this document, the list of these minor test activities will be listed but not described in detail.

The final **aim of the research plan** is to ensure that the solution developed within the project is in line with stakeholders' needs and expectations, and to collect feedback on specific Key Performance Areas and Indicators, to evaluate the expected impact of the solution.

The CODA solution falls under the ATM operational solution category and it is focused on a specific use case, supporting ATCO's work in an **en-route sector with a congested traffic situation environment** from Spanish airspace.

2 Introduction

2.1 Purpose of the document

The Exploratory Research Plan (ERP) document describes the preparation of the validation activities planned within the CODA project.

The ERP identifies the context of the experimental research plan in terms of scope, Key R&I needs, expected performance contributions, and initial and exit maturity levels.

The ERP presents the experimental research plan approach followed to mature the CODA project to TRL2, and details the stakeholders' expectations and involvement, validation objectives, success criteria, validation exercise list, and validation exercise planning.

This is the Final edition of the CODA ERP.

2.2 Intended readership

This document is intended for various stakeholders in the Air Traffic Management (ATM) community at large, especially those involved in the SESAR Programme. These include:

- CODA consortium members who will need to prepare and execute the validation activities.
- SESAR 3 JU programme management, and related SESAR 3 projects (SynthAIR, ASTRA, TRUSTY, etc.).
- Academic research and Industry researchers who wish to learn about the validation activities behind the CODA solution.

2.3 Background

The CODA project aims to demonstrate the feasibility of developing a system in which tasks are collaboratively performed by hybrid human-machine teams and dynamically allocated through adaptive automation principles.

To achieve this goal, the project **consolidates the work previously undertaken in the SESAR Exploratory Research projects.** The findings from ARTIMATION [1] and MAHALO [2] will be utilized to develop an AI-based adaptable and explainable system, allowing the system to proactively prevent future performance or safety issues. The outcomes of MINIMA [3] and STRESS [4] will be employed for a neurophysiological assessment of mental states. This will enable the system to discern operators' real-time levels of workload, attention, stress, fatigue, and situation awareness. The results from COTTON [5] and comment [6] will contribute to the development of prediction models, anticipating future situations. This enables the system to understand which activities will be undertaken by operators in the future and their potential impact on human factors.

2.4 Structure of the document

- **Section 2** describes the purpose of the document, the intended readership, background, structure of the document and glossary of terms, and the list of acronyms.
- **Section 3** describes the scope, Key R&I needs, expected performance contributions, and initial and exit maturity levels.
- **Section 4** describes the experimental research plan approach, the stakeholders' expectations and involvement, validation objectives, validation assumptions, validation exercise list, and validation exercise planning.
- **Section 5** describes the three validation exercises in detail.
- **Section 6** includes the list of references used in developing the ERP.

2.5 Glossary of terms

Term	Definition	Source of the definition
Air Traffic	All aircraft in flight or operating in the maneuvering area of an aerodrome.	ICAO Annex 11 - ATS
Air Traffic Controller	• Qualified following ICAO Annex 1 – Personnel Licensing and holding a rating appropriate to the assigned functions, • A person authorized to provide air traffic control services.	EUROCONTROL ATM Lexicon
Air Traffic Management	The dynamic, integrated management of air traffic and airspace including air traffic services, airspace management and air traffic flow management – safely, economically, and sufficiently – through the provision of facilities and seamless services in collaboration with all parties and involving airborne and ground-based functions.	ICAO 4444 - ATM
Air Traffic Services	A generic term meaning various, Flight Information Service (FIS), Alerting Service (ALRS), and Air Traffic Control Service (ATC) (area control service, approach control service, or aerodrome control service). In this document, when the term ATS is used, it is usually referring to TWR or AFIS.	ICAO, Annex 11

Table 1: glossary of terms

2.6 List of acronyms

Term	Definition
AB	Advisory Board
AI	Artificial Intelligence
ANSPs	Air Navigation Service Providers
ASM	Assumption
ATM	Air traffic management

ARTIMATION	Transparent Artificial Intelligence and Automation to ATM Systems
ATCO	Air Traffic Controller
BIM	Benefit Impact Mechanism
CODA	Controller Adaptive Digital Assistant
COTTON	Capacity Optimisation for Trajectory-Based Operations
CWP	Controller working position
DES	Digital European Sky
EASA	European Union Aviation Safety Agency
ECAC	European Civil Aviation Conference
eCOMMET	enhanced COMplexity Management Tool
ERP	Exploratory Research Plan
ERR	Exploratory Research Report
EU	European Union
EXE	Exercise
FRD	Functional requirements document
GA	Grant agreement
GDPR	General data protection regulation
HE	Horizon Europe
HMI	Human Machine Interface
ID	Identifier
KPA	Key performance area
KPI	Key performance indicator
MAHALO	Modern ATM via Human/Automation Learning Optimisation
MINIMA	MItigating. Negative Impacts of Monitoring High Levels of Automation
NM	Network Manager
KPI	Key performance indicator
OBJ	Objective
OSED	Operational service and environment description
PI	Performance indicator
RBT	Reference business trajectory
R&D	Research and development
R&I	Research and innovation
RQ	Research question
RTS	Real-time simulation
SA	Situational Awareness
SESAR	Single European Sky ATM research
SESAR 3 JU	SESAR 3 Joint Undertaking
STCA	Short-Term Conflict Alert
STRESS	Human Performance neurometricS Toolbox foR highly automatEd Systems deSign
TA	Transversal area
TRL	Technology Readiness Level
UC	Use case

WP	Work Package
XAI	Explainable Artificial Intelligence

Table 2: list of acronyms

3 Context of the experimental research plan

This chapter describes the context of the experimental research plan for the CODA project. This chapter is considered as an introduction to the following chapter, where the experimental research plan will be detailed.

3.1 Experimental Research Plan Context

The Experimental Research Plan (ERP) aims to present the **different validation experiments foreseen in the CODA project.** The experimental research is supported by a User-Centred Design iterative approach to the development of the CODA solution. Validation exercises are planned at the end of cycles of design going from the initial concept definition up to the development of a working prototype.

The final **aim of the research plan** is to ensure that the solution developed within the project is in line with stakeholders' needs and expectations, and to collect feedback on specific Key Performance Areas and Indicators, to evaluate the expected impact of the solution.

The CODA solution falls under the ATM operational solution category and it is focused on a specific use case, supporting ATCO's work in an **en-route sector with a congested traffic situation environment.** Although the project will focus on the en-route controller use case, the proposed solution is expected to be applicable in the future and to other relevant contexts that have been identified thanks to the support of external experts (e.g. tower ATCOs, Approach ATCOs). For more details, please refer to section 5.2 “Detailed operational Environment” of the D2.2 OSED [7].

It must be highlighted that **the focus of the project (and so the validation activity) is on the adaptability capacity enabled by the solution**. So, the object of the validation will be the adaptation strategy, and all the components enabling that strategy (tasks and mental states prediction models, mental states assessment module, adaptation strategy, and related human-machine interaction components). The specific tools that will be triggered by the adaptation strategy are **out of the scope of the project.**

In the context of CODA, as stated in the D2.2 OSED [7], some assumptions have been made in terms of enablers for the effective use of the proposed system in future operations.

- **Operators (i.e. en route controllers) can be monitored in real-time:** the project proposes a set of tools that can be used for assessing in real-time the mental states of controllers. The actual use of similar tools is impacted by several considerations that are not technical ones. The use of alternative monitoring tools can influence the acceptability of such an assessment, as well as the data management process and privacy management of the acquired data.
- **AI-based Digital assistants and high-level automation tools will be developed and implemented in the ATM domain: CODA makes ATM systems adaptable to specific situations (e.g. high peak workload), a**lso by acting on the behaviour of supporting tools. The more those tools are available, the more CODA will prove its effectiveness in preventing unwanted situations.

CODA focuses on a specific use case (UC1): EXE ATCO monitoring an en-route sector (for more details please refer to D2.2 OSED [7])

3.2 Scope

As stated above, the purpose of this document is to present the Experimental Research Plan (ERP), and it takes input from the WP2, WP3, WP4, and WP5. The work contained in this deliverable is closely linked to the deliverable D2.1 FRD – Functional Requirements Document [8] which explains in detail the CODA system functional description and the necessary logical interfaces with other systems – and to the deliverable D2.2 OSED [7], which not only includes the OSED but also the current state of the art on ATCO roles and contexts and the definition of generic use cases.

The overall objective of the CODA project is **to increase ATM's efficiency, capacity, and safety**, maximizing human-AI teaming by developing a system in which tasks are performed collaboratively by hybrid human-machine teams and dynamically allocated through adaptive automation principles. To do so CODA focuses on developing a solution that predicts relevant mental states of en-route air traffic controllers to anticipate possible problems and trigger specific actions (such as the activation of Digital Assistants).

The end product will be an **adaptable system** that enables the adaptive use of (AI-based) components/tools. For more details related to the concept outline, please refer to the D2.1 FRD [8] and D2.2 OSED [7].

3.3 Key R&I needs

Below is a summary of the outstanding and innovation (R&I) needs that are going to be addressed by the ERP.

R&I ID	R&I needs/questions
CONCEPT DEVELOPMENT	
R&I- 01	Understanding if it is possible to build a Digital Assistant able to allocate tasks dynamically, based on ATCO's mental state and task predictions
HUMAN PERFORMANCE	
R&I- 02	Understanding the impact of a Digital Assistant using an adaptation strategy on ATCO's performance
R&I- 03	Understanding the impact of the CODA solution on human capabilities and limitations (workload, stress, fatigue, attention, situation awareness)
R&I- 04	Investigating the suitability of the CODA system in supporting the tasks of ATCOs
CAPACITY	
R&I- 05	Investigating to what extent the CODA system will impact the amount of traffic/airspace each ATCO can control and therefore airspace capacity
OPERATIONAL EFFICIENCY & SAFETY	
R&I- 06	Understanding the weaknesses and the strengths of the CODA system in terms of safety and operational efficiency and what could be the implications

R&I- 07	Investigating the suitability of the task prediction model in predicting future ATCO tasks based on traffic patterns
R&I- 08	Investigating the suitability of the cognitive status prediction model in assessing and predicting the ATCO's cognitive mental state
R&I- 09	Understand which level of transparency and explainability will give the ATCO the ability to perform the tasks at their best
COST EFFICIENCY	
R&I- 10	Evaluate the impact of the CODA system in terms of cost efficiency (Flights per ATCO Hour on Duty)

Table 3: CODA R&I

3.4 Estimated performance contributions

As already stated, the changes introduced by the CODA system are the adaptability of the ATM En-route system for an executive controller, the adaptability of task allocation between operator and machine (automation, AI), and the adaptability of the interface used by the operator.

At a high level, the outputs of the validation exercises are expected to increase the know-how for future deployment of such a system and are expected to bring performance contributions to the following areas:

- **operational efficiency, cost efficiency, and en-route capacity:** by leveraging AI during peak workloads, such as calculating optimal routes, the CODA system will ensure optimal resource utilization and reduced operational costs (i.e. enhancing flights per hour ATCO on duty). Air Traffic Controllers will be able to focus on critical decision-making while automation will handle other tasks, such as routine tasks, resulting in more efficient flight management operations (i.e. reducing the average difference in actual & Flight Plan or RBT durations). Moreover, AI's ability to process data and make informed decisions will contribute to a more streamlined airspace management system, accommodating higher traffic volumes and then enhancing capacity (i.e. improving en-route throughput in challenging airspace per unit time).

- **safety:** through the adaptable allocation of tasks and the adaptive HMI, the CODA solution is expected to ensure at least the same level of safety in airspace management.

- **and human performance:** by calculating excessive workloads, AI will assist the Air Traffic Controller (ATCO). This will ensure that the level of **fatigue**, **workload**, and **stress will be at least maintained**. Furthermore, thanks to the adaptable Human-Machine-Interface (HMI), the system will maintain the level of ATCO's **attention** and **Situational Awareness** (**SA**), by deciding, based on the workload, which information to display.

As already stated, due to the level of maturity of the CODA project the performance contributions of such a solution are expected to be at a qualitative level. The related KPA/KPIs are described in detail in Section 4.3.

3.5 Initial and exit maturity levels

The purpose of this section is to identify the current level of development of the SESAR Solution and to state the level that the stakeholders expect the related concepts to mature to by the end of the validation process.

The CODA project is classified as **exploratory research**. The activities at this phase cover low TRL research and are divided into two distinct maturity subphases. The first subphase is 'excellent science' and is primarily oriented at creating a coordinated body of ideas using a knowledge transfer network. It consists of early research, which leads to the second subphase, called 'application-oriented research', which takes the most promising ideas and applies them to an area of ATM where there is potential to exploit the concept (s) to deliver future operational benefit. Regarding R&I maturity, the CODA project will start with a TRL (Technology Readiness Level) 1, and it sets the goal of achieving a TRL 2 (as defined by the Maturity Assessment Criteria from SESAR) by the end of the work. In TRL 2 the following are comprised:

- TRL 1: basic principles observed and reported,

- TRL 2: technology concept and/or application formulated.

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
0447	Adaptive System based on Controller Status to enhance Human-AI Teaming	TRL1	TRL2	NA

Table 4: maturity levels table

4 Experimental plan

4.1 Experimental plan approach

The CODA project aims to achieve the TRL2 maturity level through the collection of evidence and results, and through the definition of the validation objectives and success criteria established in section 4.3 and following and in line with the R&I needs.

The experimental plan approach will foresee 3 validation sessions and is as follows:

ID	Exercise	Rationale	Type	Location	Date
VAL.1.1-CODA-0447-TRL2	**Validation of the CODA operational concept, use cases, and scenarios**	Consolidate the CODA concept and define scenarios for testing the solution	Workshops with experts	Online	December 2023
VAL.1.2-CODA-0447-TRL2	**Validation of the mental states' prediction models**	Validate the mental states' prediction models integrated	Small-Scale Simulation (RTS)	CRIDA	October 2024
VAL.1.3-CODA-0447-TRL2	**Final Validation of the CODA system**	Validate the integrated models, final HMI, and final adaptation strategy	Human-in-the-loop Simulation (RTS)	CRIDA	March 2025

Table 5: list of validation activities

This three-step approach will facilitate the refinement of the concept outlined in the D2.2 OSED [7] and gather the necessary background to advance the solution to TRL2. The validation activities are conducted throughout the development lifecycle of the solution, starting with the definition of use cases and scenarios. These activities progress step-by-step, ensuring the validation of each component of the CODA system separately before the integration of them into the complete system.

Along with the three-step approach, the validation plan includes also the execution of **three pre-validation activities, as follows:**

ID	Exercise	Rationale	Type	Location	Date
T1	**Adaptation strategy and HMI**	Gather initial feedback on the adaptation strategy and initial ideas on HMI.	Workshops with experts	Online (NLR + ENAC)	March-May 2024

	requirements gathering				
T2	**Adaptation strategy and HMI pre-validation with ATCOS**	Gather feedback on the adaptation strategy and present the CODA's interaction modes to help participants understand how the system will function.	Presentations and simulations with mock-ups	ENAC	September 2024
T3	**Tasks prediction model verification**	Validate the task prediction model by comparing the task predicted by the AI with the traffic data provided by CRIDA.	Data verification	UPM	July 2024

Table 6: list of pre-validation activities

More information on how those tests are conducted and their results **will be available in D3.3 – Models Validation; D4.3 – Integration with the Prediction Models (HMPE); D5.5 – Prediction Models and Adaptation Strategy Integration.**

The timeline of the validation and pre-validation activities is designed as follows:

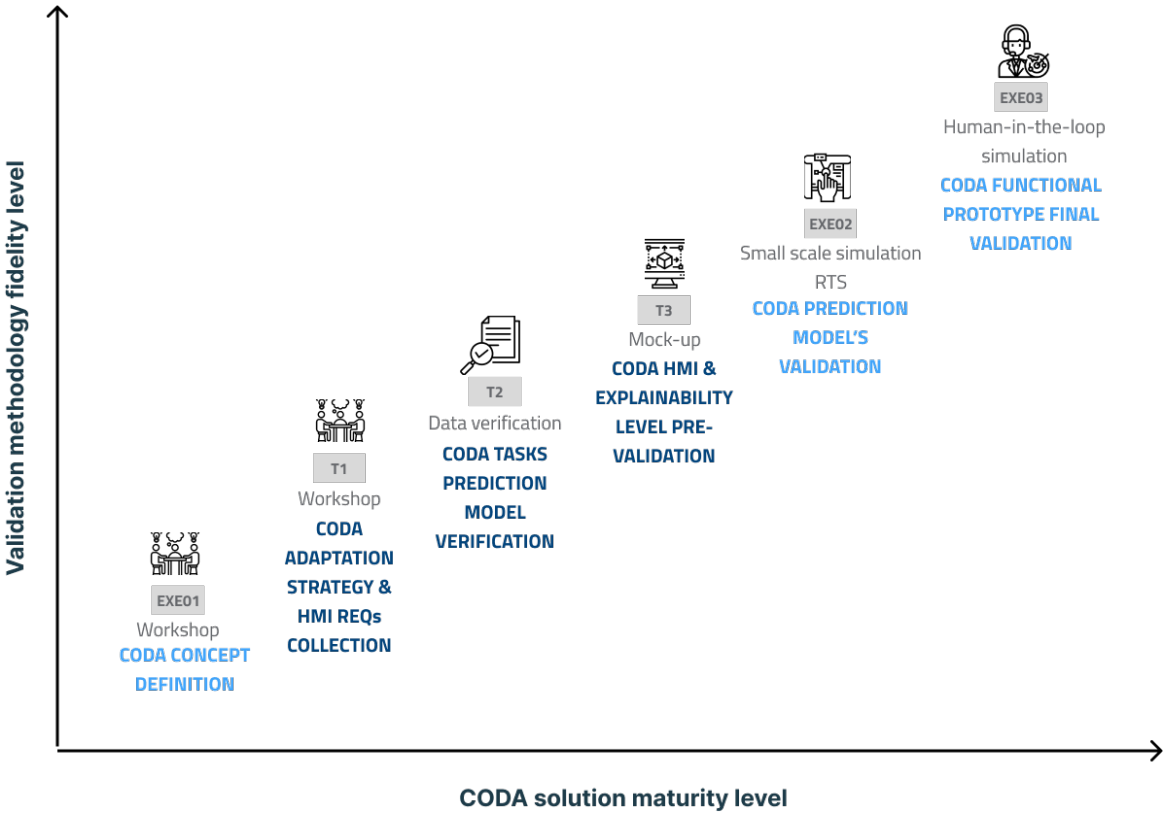

Figure 1: validation and pre-validation exercises foreseen in the project

4.2 Stakeholders' expectations and involvement

Stakeholder	Involvement	Why it matters to the stakeholder
Scientific community (Universities, research institutions, EU projects, educational institutions)	To disseminate the project objectives and results of the project, potentially finding synergies with other projects/approaches focused on the implementation of higher levels of automation and AI within ATM through scientific dissemination, workshops, etc.	To raise mutual awareness of approach results potentially resulting in further collaborations. To make students and academia aware of advanced AI solutions in ATM, helping to entice them to this knowledge area.
Institutional bodies (EU and EC, European Joint Undertakings, EASA, Policy makers, Regulatory and safety agencies, Standard making bodies, National bodies, Certification bodies).	To provide an actionable roadmap for the progressive and controlled deployment of AI solutions through direct links with the Advisory Board and dissemination events	To help facilitate the progressive and safe introduction of AI solutions in ATMs in a controlled manner, by implementing actual examples of potential uses to build confidence among users and regulators and to help define future safety, security, and privacy constraints for their implementation.
Industry (ATM automation systems providers).	To make industry aware of the potential and results of the CODA project for final implementation using direct links with industries and dissemination events	To provide a clear path for the higher TRLs implementation of the solution, and potentially open the way to more AI-based solutions implementation.
General public, media.	To raise awareness of CODA objectives and results in the general public	To clarify the value of the project and its benefits for citizens, especially in terms of safety and economic impacts.
Users (ATCOs, ANSPs, NM).	To promote solution benefits and potential applications to the ANSPs and Network Management stakeholders at large using NM fora and communication channels, and direct relations with some ANSPs and ATCO organizations (e.g., ENAIRE, IFATCA).	To deploy, in the future, a CODA-based solution, with the help of the industry To inform decision-makers and incentivize operational stakeholders to adopt the CODA solution

Table 7: stakeholders' expectations and involvement

4.3 Validation Objectives

This section lists the validation objectives required for the exploratory research plan.

The detailed validation objectives are decomposed into Key Performance Areas (KPAs), Success Criteria, Key Performance Indicators (KPIs), and Performance Indicators (PIs).

The following table describes the relevant KPAs and related Detailed Validation Objectives.

Identifier	Validation Objectives	R&I Need
OBJ-CODA-TRL2-ERP-001	Assess the impact of the CODA system on ATCO's **human performance** in nominal conditions	2,3,4
OBJ-CODA-TRL2-ERP-002	Assess the impact of the CODA system on **capacity** in nominal conditions	5
OBJ-CODA-TRL2-ERP-003	Assess the impact of the CODA system on **operational efficiency** in nominal conditions	6,7, 8, 9
OBJ-CODA-TRL2-ERP-004	Assess the impact of the CODA system on **safety** in nominal conditions	6
OBJ-CODA-TRL2-ERP-005	Assess the impact of the CODA system on **cost efficiency** in nominal conditions	10

Table 8: validation objectives

[OBJ]

Identifier	OBJ-CODA-TRL2-ERP-001
Objective	Assess the impact of the CODA system on ATCO's human performance in nominal conditions
Title	Human Performance Assessment
Category	<Human Performance>
Key environment conditions	En-route
TRL level	TRL 2

[OBJ Suc]

Identifier	Success Criterion
CRT-CODA-TRL2-ERP-001-001	The majority of ATCOs' feedback (60%) states that they understand their **responsibilities and tasks** with the new human-AI teaming allocation in nominal conditions.
CRT-CODA-TRL2-ERP-001-002	The majority of ATCOs' feedback (60%) states that they understand the CODA system and the new **procedures** that are introduced in nominal conditions
CRT-CODA-TRL2-ERP-001-003	The majority of ATCOs' feedback (60%) states that they are supported by the new **operating methods** that specify procedures/requirements for human-AI collaboration/communication in nominal conditions
CRT-CODA-TRL2-ERP-001-004	The majority of ATCOs' feedback (60%) states that they can achieve their **tasks** while being supported by the CODA system in nominal conditions

CRT-CODA-TRL2-ERP-001-005	The majority of ATCOs' feedback (60%) states that they can perceive the CODA system **HMI** as capable of supporting them in carrying out their tasks in nominal conditions
CRT-CODA-TRL2-ERP-001-006	The majority of ATCOs' feedback (60%) states that their mental **workload** is positively supported by the CODA system and relevant procedures in nominal conditions
CRT-CODA-TRL2-ERP-001-007	The majority of ATCOs' feedback (60%) states that their **situational awareness** is positively supported by the CODA system in nominal conditions
CRT-CODA-TRL2-ERP-001-008	The majority of ATCOs' feedback (60%) states that their **attention** is positively supported by the CODA system in nominal conditions
CRT-CODA-TRL2-ERP-001-009	The majority of ATCOs' feedback (60%) states that their **stress** is positively supported by the CODA system in nominal conditions
CRT-CODA-TRL2-ERP-001-010	The majority of ATCOs' feedback (60%) states that their **fatigue** is positively supported by the CODA system in nominal conditions

[OBJ]

Identifier	OBJ-CODA-TRL2-ERP-002
Objective	Assess the impact of the CODA system on capacity in nominal conditions
Title	Capacity assessment
Category	<Capacity>
Key environment conditions	En-route
TRL level	TRL 2

[OBJ Suc]

Identifier	Success Criterion
CRT-CODA-TRL2-ERP-002-001	En-route throughput per unit time (CAP1) is slightly improved in nominal conditions

[OBJ]

Identifier	OBJ-CODA-TRL2-ERP-003
Objective	Assess the impact of the CODA system on operational efficiency in nominal conditions
Title	Operational Efficiency Assessment
Category	<Operational feasibility>
Key environment conditions	En-route
TRL level	TRL 2

[OBJ Suc]

Identifier	Success Criterion
CRT-CODA-TRL2-ERP-003-001	Average Difference (deviation) in actual & Flight Plan or RBT durations (PRD1) is slightly decreased in nominal conditions
CRT-CODA-TRL2-ERP-003-002	Average en-route fuel burnt per flight (FEFF1) is slightly decreased in nominal conditions
CRT-CODA-TRL2-ERP-003-003	En-route time (TEFF6) is slightly decreased in nominal conditions

[OBJ]

Identifier	OBJ-CODA-TRL2-ERP-004
Objective	Assess the impact of the CODA system on safety in nominal conditions
Title	Safety Assessment
Category	<Safety>
Key environment conditions	En-route
TRL level	TRL 2

[OBJ Suc]

Identifier	Success Criterion
CRT-CODA-TRL2-ERP-004-001	The new operational procedures/CODA system maintains the level of safety in nominal conditions.

[OBJ]

Identifier	OBJ-CODA-TRL2-ERP-005
Objective	Assess the impact of the CODA system on cost efficiency in nominal conditions
Title	Cost Efficiency Assessment
Category	<Performance>, <Operational feasibility>
Key environment conditions	En-route
TRL level	TRL 2

[OBJ Suc]

Identifier	Success Criterion
CRT-CODA-TRL2-ERP-005-001	Flight per ATCO hour on duty (CEF2) is slightly increased in nominal conditions

Table 9: complete list of detailed validation objectives and success criteria

Attention was paid to providing a clear view of dependent and independent variables which provides a very structured approach to validate the solutions in a harmonized way:

KPA	Detailed validation objectives	Success criteria	Variables		
			Dependent	Independent	Control
Human Performance	OBJ-CODA-TRL2-ERP-001	CRT-CODA-TRL2-ERP-001-001-10	ATCO Workload, SA, Stress, Fatigue, Attention	Use of CODA (without CODA; with CODA)	Age Years of experience Type of Scenario Airspace structure ATCO instructions per flight
Capacity	OBJ-CODA-TRL2-ERP-002	CRT-CODA-TRL2-ERP-002-001	En-route throughput	Use of CODA (without CODA; with CODA)	Airspace structure Separation management
Operational Efficiency	OBJ-CODA-TRL2-ERP-003	CRT-CODA-TRL2-ERP-003-001-3	Average difference in actual and flight plan or RBT duration	Use of CODA (without CODA; with CODA)	Airspace structure Separation management
Safety	OBJ-CODA-TRL2-ERP-004	CRT-CODA-TRL2-ERP-004-001	Number of Mid-Air Collisions	Use of CODA (without CODA; with CODA)	Airspace structure Separation management
Cost Efficiency	OBJ-CODA-TRL2-ERP-005	CRT-CODA-TRL2-ERP-005-001	Flight per ATCO hours on duty	Use of CODA (without CODA; with CODA)	Airspace structure Separation management

Table 10: validation exercises variables

4.4 Validation assumptions

Assumption ID	Assumption title	Assumption description	Justification	Impact Assessment
ASM-CODA-TRL2-ERP-001	Training and competencies	ATCOs have appropriate training and competencies.	To have relevant results, ATCOS must be familiar with the concept, and procedures relevant to the operational environment chosen for the validation and with the platform tools.	High
ASM-CODA-TRL2-ERP-002	Considered Airspace	The operational environment used for testing the solution is representative of the environment in current operations.	The validated operational environment (En-route) shall be as close to the environment as in the current operations to quantify the benefits of the solution accurately, considering the expected airspace characteristics at horizon 2035.	High
ASM-CODA-TRL2-ERP-003	Neurophysiological Assessment	Operators (i.e. en-route controllers) can be monitored in real-time	The project proposes a set of tools that can be used for assessing in real-time the mental states of controllers. The actual use of similar tools is impacted by several considerations that are not technical ones. The use of alternative monitoring tools can influence the acceptability of such an assessment, as well as the data management process and privacy management of the acquired data.	High
ASM-CODA-TRL2-ERP-004	AI feasibility	AI-based Digital assistants and high-level automation	CODA makes ATM systems adaptable to specific situations (e.g.	High

		tools will be developed and implemented in the ATM domain	high peak workload), also by acting on the behaviour of supporting tools. The more those tools are available, the more CODA will prove its effectiveness in preventing unwanted situations.	

Table 11: validation assumptions overview

4.5 Validation exercises list

Identifier	VAL.1.1-CODA-0447-TRL2
Title	**Validation of the CODA operational concept, use cases and scenarios**
Description	The validation consisted of a workshop conducted with the members of the CODA Advisory Board. In this workshop, the content of the D2.2 Operational Service and Environment Definition (OSED) [7] and D2.1 Functional Requirements Document (FRD) [8] was discussed.
KPA/TA addressed	Human Performance
Addressed expected performance contribution(s)	Technical exercise to prepare the VAL1.3, so no EPC
Maturity level	TRL2
Use cases	UC1 (+ additional possible future use cases for CODA application)
Validation technique	Expert Group Workshop
Validation platform	Zoom meeting
Validation location	Online
Start date	15/12/2023
End date	15/12/2023
Validation coordinator	Validation WP6 DBL, WP2 ENAC
Status	<in progress>
Dependencies	VAL1.2 and VAL1.3

[EXE Trace]

Linked Element Type	N/A
SESAR Solution	SESAR Solution - CODA
Project	CODA
Sub-Operating Environment	N/A
Validation Objective	OBJ-CODA-TRL2-ERP-001, OBJ-CODA-TRL2-ERP-006

Table 12: validation exercise #01 layout

Identifier	VAL.1.2-CODA-0447-TRL2
Title	**Validation of the mental states' prediction models**
Description	The mental state prediction models will be validated through a small-scale simulation with Air Traffic Controllers (ATCOs) provided by IFATCA.
KPA/TA addressed	<Human Performance>
Addressed expected performance contribution(s)	Technical exercise to prepare the VAL1.3, so no EPC
Maturity level	TRL2
Use cases	UC1
Validation technique	Small-scale simulation with ATCOs
Validation platform	ESCAPE
Validation location	CRIDA, BS
Start date	21st October 2024
End date	25th October 2024
Validation coordinator	WP6 DBL, WP3 CRIDA
Status	<in progress>
Dependencies	VAL1.1 and VAL1.3

[EXE Trace]

Linked Element Type	N/A
SESAR Solution	SESAR Solution - CODA
Project	CODA
Sub-Operating Environment	N/A
Validation Objective	OBJ-CODA-TRL2-ERP-001

Table 13: validation exercise #02 layout

Identifier	VAL.1.3-CODA-0447-TRL2
Title	**Final Validation of the CODA system**
Description	To validate the integrated models, final HMI and final adaptation strategy, a human-in-the-loop simulation will be conducted, involving professional ATCOs.
KPA/TA addressed	<Capacity>, <Safety>, <Human Performance>, <Operational Efficiency>, <Cost Efficiency>
Addressed expected performance contribution(s)	Improvement in en-route capacity (CAP2) Safety neutral (SAF1.1) Human Performance not degraded (HP1, HP2)

	Improvement in Operational Efficiency (PRD1) Improvement in Cost Efficiency (CEF2)
Maturity level	TRL2
Use cases	UC1
Validation technique	Human-in-the-loop simulation
Validation platform	ESCAPE
Validation location	CRIDA
Start date	17th March 2025
End date	21st March 2025
Validation coordinator	Validation WP6 DBL
Status	<in progress>
Dependencies	VAL1.1 and VAL1.2

[EXE Trace]

Linked Element Type	N/A
SESAR Solution	SESAR Solution - CODA
Project	CODA
Sub-Operating Environment	N/A
Validation Objective	OBJ-CODA-TRL2-ERP-001; OBJ-CODA-TRL2-ERP-002; OBJ-CODA-TRL2-ERP-003; OBJ-CODA-TRL2-ERP-004; OBJ-CODA-TRL2-ERP-005, OBJ-CODA-TRL2-ERP-006

Table 14: validation exercise #03 layout

4.6 Validation exercises planning

Below is the screenshot of all the validation exercises' schedules captured in STELLAR which are under the scope of this ERP.
These validation exercises will be detailed in Chapter 5.

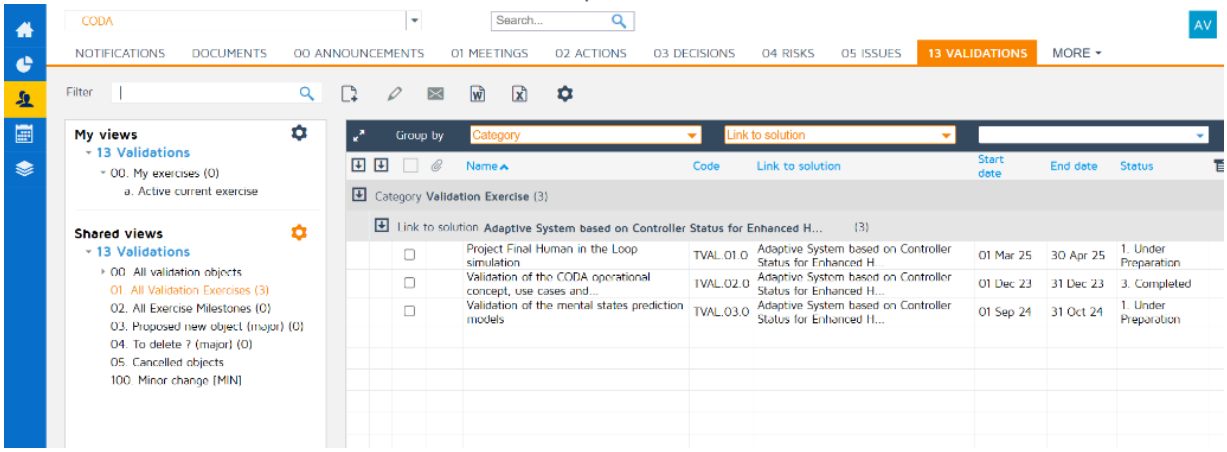

Figure 2: STELLAR - validation exercises

4.7 Deviations with respect to the SESAR 3 JU project handbook

There are no deviations with respect to the SESAR 3 JU project handbook.

5 Validation exercises

5.1 Validation exercise #01 plan - Validation of the CODA operational concept, use cases and scenarios

This section describes the validation plan for exercise #01, according to the validation exercise plan template provided hereunder.

5.1.1 Validation exercise description and scope

The validation exercise #01 consisted of a workshop conducted with the members of the CODA Advisory Board. The scope of the workshop was to consolidate the CODA concept and define scenarios for testing the solution. To achieve this, participants were asked to provide their feedback, with Mentimeter, on various aspects of the project. An example of feedback can be found in the Figure below.

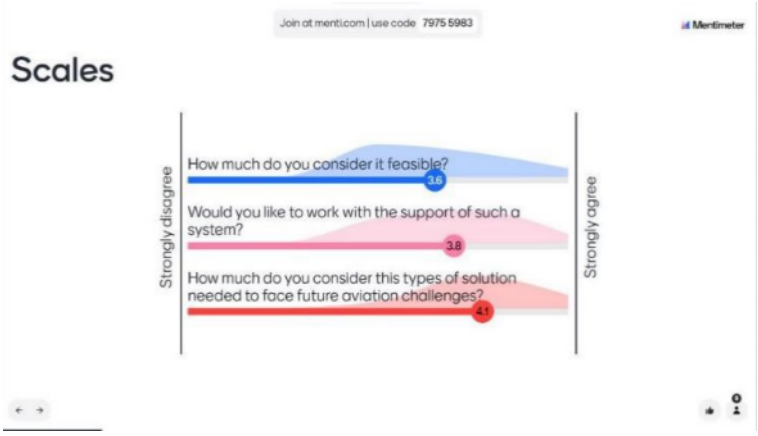

Figure 3: example of Mentimeter feedback layout in exercise #01

Specifically, the following feedback was asked:

- Insights on the CODA Concept, including its anticipated impact and any potential issues
- Insights on the actual predictability of the ATCO's tasks and the approach proposed for the HMI and its explainability; Insights on the relevance of the selected cognitive functions (fatigue, workload, vigilance, stress) for the mental state prediction model
- Insights on the feasibility of ATCOs' real-time mental state assessment
- Insights on en-route scenarios to be defined for validating the CODA system
- Insights on how to set the strategy the CODA system will follow to adapt to the foreseen situations (e.g. better delineate instances where the adaptation strategy might undermine confidence in the system)

The stakeholders considered in the AB include the research community and SESARJU, ANSPs, aircraft manufacturers and system industries, institutional and regulatory bodies, and ATCO associations. A more detailed list is provided in the table below.

Company/Entity
DBL
BS
CRIDA
UPM
NLR
UGR
DBL/ HAIKU Project
ANACNA/ENAV
EASA
Mälardalen University / TRUSTY Project
IFATCA
FOXATM

Table 15: Advisory Board Meeting CODA partners-participants list (online) in exercise #01

The outcome of the workshop was reported in D2.2 OSED [7].

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

Stakeholder	Involvement	Why it matters to the stakeholder
Users (ATCOs, ANSPs, NM).	Direct	The exercise matters to ATCOs because it helps to define the concept of the solution which ultimately helps drive the CODA project through all its phases. End users expect that the CODA system will enable ANSPs to improve their possibility of achieving their business targets of providing enough capacity for the demand most efficiently and safely and improving effectively ATM management.
Scientific community (Universities, research institutions, EU projects, educational institutions)	Direct	The scientific community, as part of the consortium, expects that the solution concept validated through this exercise is accurate and complete. This will ensure that the CODA system is developed to meet the expectations of end users. Moreover, is expected that the CODA system, as designed and developed based on the validated requirements, will be effective in supporting ATM operations.

Institutional and regulatory bodies (EU and EC, European Joint Undertakings, EASA, Policy makers, Regulatory and safety agencies, Standard bodies, National bodies, Certification bodies).	Direct	Regulatory bodies facilitated the definition of the CODA concept by granting a safe introduction of AI solutions in the ATM domain. They provided feedback on how to implement the CODA system in a controlled way by helping anticipate future safety, security, and privacy constraints for its implementation.

Table 16: stakeholders' expectations exercise #01

5.1.3 Validation objectives

SESAR solution validation objective	SESAR solution success criteria	Coverage and comments on the coverage of the SESAR solution validation objective in exercise #01	Exercise validation objective	Exercise success criteria
OBJ-CODA-TRL2-ERP-001: Assess the impact of the CODA system on ATCO's **human performance** in nominal conditions	Success criteria (same as in section ***¡Error! No se encuentra el origen de la referencia.***)	Partially covered	OBJ-CODA-TRL2-ERP-EXE01-001: Assess the opinion of experts (AB) on the potential impact of the CODA system on ATCO's human performance	CRT-CODA-TRL2-ERP-EXE01-001: The majority of AB experts's feedback (60%) states that they perceive the CODA system HMI as capable of supporting ATCOs in carrying out their tasks CRT-CODA-TRL2-ERP-EXE01-002: The majority of AB experts's feedback (60%) states that they perceive ATCOs can achieve their tasks while being supported by the CODA system. CRT-CODA-TRL2-ERP-EXE01-003: The majority of AB experts's feedback (60%) states that they understand the CODA system and the new procedures that are introduced. CRT-CODA-TRL2-ERP-EXE01-004: The majority of AB experts's feedback (60%)

				states that they perceive that SA, vigilance, stress, workload, and fatigue are at least maintained or are not degraded

Table 17: validation objectives addressed in validation exercise #01

5.1.4 Validation scenarios

At the beginning of the workshop, the CODA solution was introduced to the participants. To illustrate how the CODA system could work, participants were asked to imagine a scenario where an Air Traffic Controller (ATCO) is managing a complex traffic situation with a medium level of workload.

The system, capable of assessing the operator's current neurophysiological status, including workload, stress, vigilance, and fatigue, predicts that upcoming tasks will increase the workload beyond what an operator can handle.

In anticipation of this issue, the system determines an adaptation strategy: it may choose to increase the level of automation by enabling additional AI-based tools, propose sector splitting, or contact network management to or even contact network management to ask for some regulations to be issued.

The following images were shown to the participants to have a simplified overview of the CODA solution.

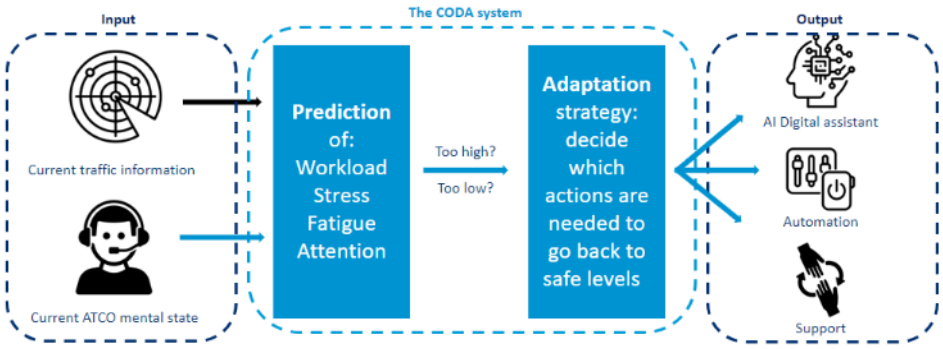

Figure 4: simplified overview of the CODA solution

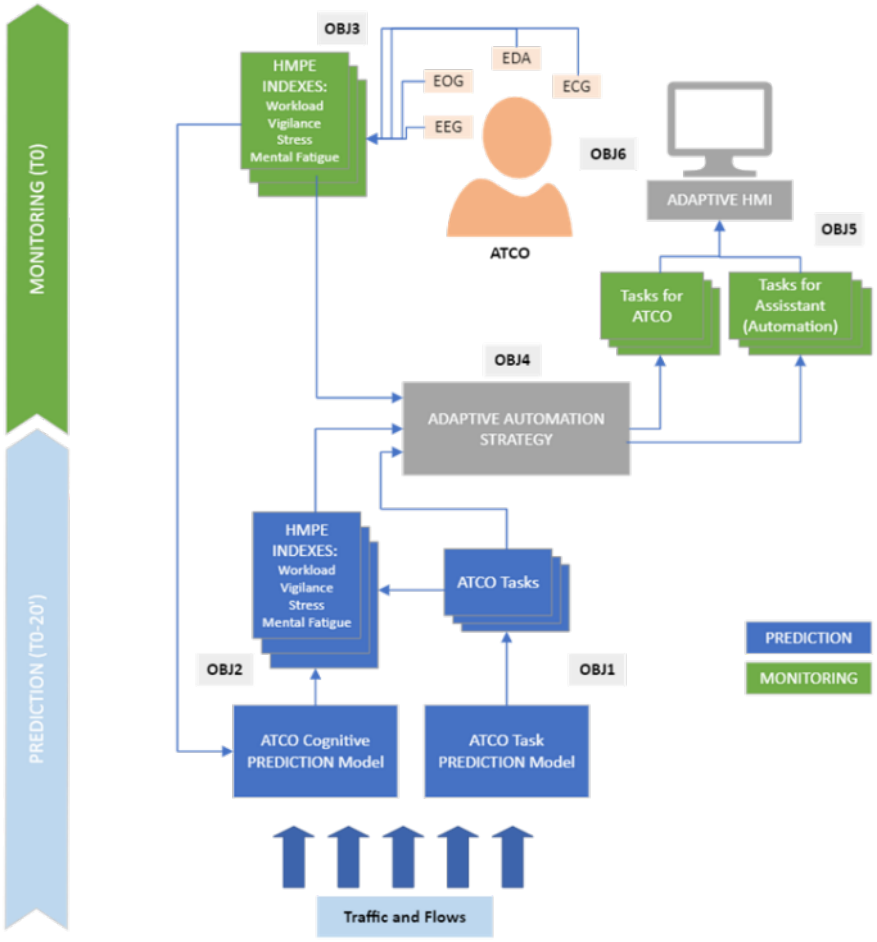

Figure 5: detailed overview of the CODA solution

5.1.4.1 Reference scenario(s)

The baseline scenario represents the current state of operations, where ATCOs operate using the existing systems, displays, procedures, etc. This scenario will serve as a reference point for comparison with the scenario in which the CODA system is introduced.

5.1.4.2 Solution scenario(s)

The solution scenario matches with the validation scenario. See section 5.1.4

5.1.5 Exercise Validation Assumptions

The assumptions that guided the definition and execution of the exercise are those identified for the entire solution, as outlined in section 4.4.

5.1.6 Limitations and impact on the level of significance

The workshop with experts (AB) is the tool selected for the validation of the CODA concept and represents the first step within the solution design cycle. It has been proposed as a tool to gather information regarding the concept of the CODA solution.

The main limitation of this exercise relies on its dependence on high-level feedback collection to validate the OSED and gather initial information for the development of the OSED itself. Indeed, the collected data are qualitative and provide a broad overview of the entire solution, without delving into the detailed aspects of each element from which the CODA system emerges.

5.1.7 Validation exercise platform/tool and validation technique

In this section, the validation platform and validation technique used in Exercise 1 are discussed.

5.1.7.1 Validation exercise platform/tool characteristics

Exercise 1 was held on the online Zoom video meeting platform. Slides were used to present the CODA concept and to give the participants further information related to the project. Below is an example:

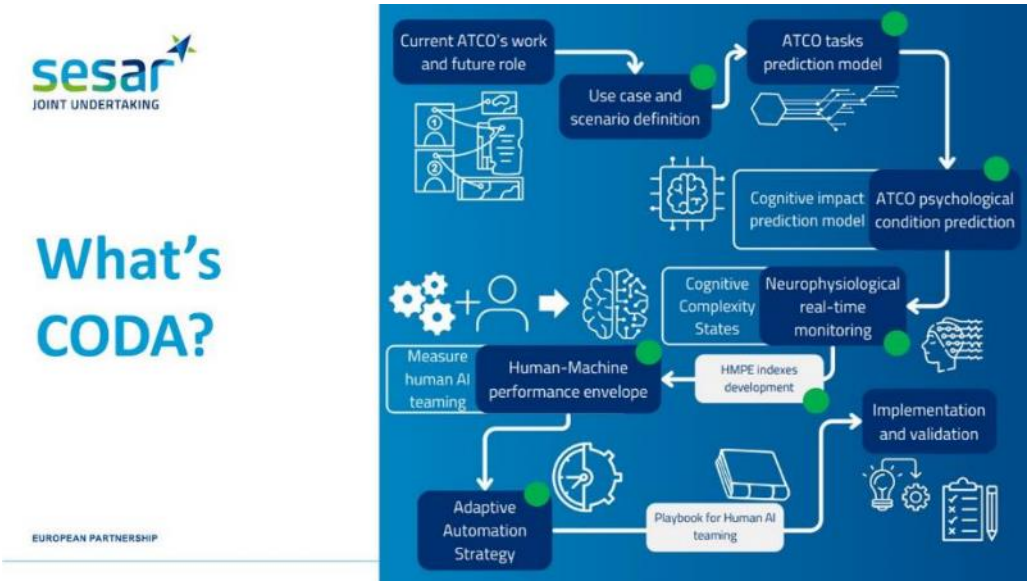

Figure 6: example of the slide presented in exercise #01

The Mentimeter tool was used to articulate the discussion and to obtain feedback from the AB. Below are some examples of the used platform.

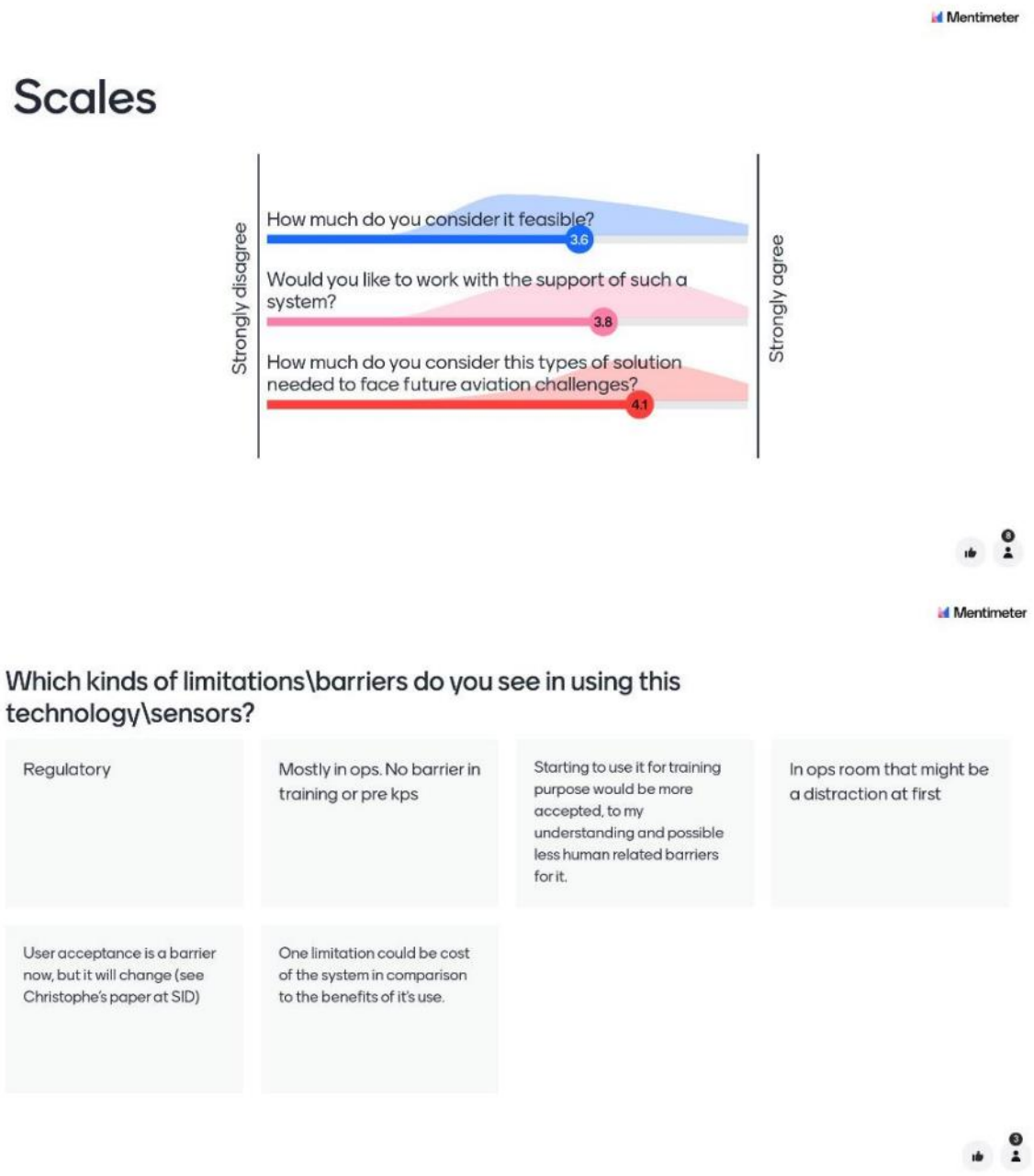

Figure 7: example of the Mentimeter Interface in exercise #01

5.1.7.2 Validation exercise technique

This validation exercise utilized AB workshop validation, which involves having domain experts (CODA AB) review and assess the concept of the solution against their knowledge and experience. This technique is effective in ensuring that the requirements are accurate, complete, and aligned with the project's scope and objectives.

Before conducting the CODA Concept Validation Workshop a live discussion (Focus Group) with air traffic controllers (ATCOs) was conducted by ENAC to consolidate the CODA concept and define scenarios for testing the solution. This approach leveraged collaborative input from multiple individuals gathered simultaneously and allowed to identify five scenarios that subsequently were presented during the workshop with the Advisory Board.

5.1.8 Data collection and analysis

5.1.8.1 Data and data collection methods

CODA AB participants were asked to express their feedback on the CODA concept. The qualitative feedback was collected through the Mentimeter platform, (AND NOTE) and it will be reported in the Exploratory Research Report (ERR).

5.1.8.2 Analysis methods

The qualitative data collected from the CODA AB feedback was analysed by the partners during the following meetings.

5.1.9 Exercise planning and management

5.1.9.1 Activities

Phase	Activities	Responsibilities
Preparation	Workshop Preparation	Exercise Leader
Exercise	Briefing: Present the project, the consortium, and the validation exercise to the Advisory Board	Exercise Leader / ATCOs
	Execute according to plan	Exercise Leader / ATCOs
	Debriefing: Get required feedback from participants via debriefing sessions and focus group	Exercise Leader / ATCOs
Post-Exercise	Analysis of results from focus groups, discussions, and debriefing	Exercise Leader / Project's Consortium

Table 18: exercise #01 Planning activities

5.1.9.2 Roles and responsibilities in the exercise

This subsection describes the roles and responsibilities of the CODA consortium in preparing, conducting, and analysing validation exercise #01.

Actor	Role/responsibility
Exercise Leader	Managed the involvement of the AB members Managed the organization of the workshop, collection, and analysis of data, and report of the validation exercise #1 in the D6.2 Experimental Research Report (ERR)

	Generated the final analysis of the AB workshop results
Advisory Board	Feedback during discussions and debriefings Provided opinions through a survey
Validation Team	Collected feedback on the expected impact and possible issues of the CODA solution (NLR) Collected feedback on the mental states selected for the solution and gathered suggestions on other interesting mental states to be considered in future projects (CRIDA) Collected feedback on the actual predictability and explainability of the ATCOs' tasks (UPM) Collected feedback on the feasibility of ATCOs' mental state assessment (BS) Collected feedback on en-route scenarios and on the adaptation strategy the system should follow to adapt to the predicted situations (ENAC)

Table 19: exercise #01 Role and Responsibilities

5.1.9.3 Time planning

The planning is set as follows:

- The preparatory phase of the validation exercise #01 started with the CODA CONCEPT ADVISORY BOARD INITIAL MEETING preparation in November 2023 (M3)
- Execution phase of exercise #1 was in December 2023 (M4).
- Post-execution phase, namely the analysis of the AB feedback on the CODA concept, was in December 2023 (M4).

The following table indicates the start and end dates of the preparatory, execution, and post-execution phases.

Activity	Week				
	1	2	3	4	5
Preparation					
Execution					
Analysis					

Table 20: detailed exercise #01-time planning

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**

Risk 12: The proposed CODA concepts get negative reviews from the AB members.	Medium	Medium	Medium	In case of negative feedback collected from the Advisory Board (AB), the concept would have been reworked, and the revised version would have been presented again for evaluation by the advisory board.

Table 21: exercise #01 Risks and mitigation actions

5.2 Validation exercise #02 plan - Validation of the mental states' prediction models

5.2.1 Validation description and scope

The mental state prediction models will be validated through a **small-scale simulation with Air Traffic Controllers (ATCOs)** provided by IFATCA.

The validation activities of mental models will be conducted in presence in Madrid at CRIDA facilities in October 2024 (M14).

One of the objectives of the CODA project is to be able to **predict the mental states of the controllers through the creation of a mental model**. The mental model created within WP3 is going to provide information regarding the controllers' workload, stress, fatigue, and vigilance decrements. The mental model has been modelled using psychology theories adapted to the En-Route executive controller role. More information can be found in the D3.3 Models Validation available at M14. The model has been initially calibrated using information from real operations within Spanish airspace but needs to be verified by using a set of scenarios that provoke the target mental reactions on the controllers. The mental model prediction and the actual measurements will then be compared to gather information regarding the performance of the models outside the theoretical field in laboratory conditions.

Exercise #2 will be used to perform the validation of the **mental model prediction** through a series of scenarios within a **small-scale Real-Time Simulation**, RTS. The scenario will be an En-Route sector of Spanish airspace. The exercise is within the TRL2 target technology readiness level. The platform that will be used is ESCAPE, a scalable EUROCONTROL ATM real-time simulation platform that supports small- and large-scale simulations in airspace design for both En-Route and TMA. CODA plans to use the ESCAPE Light platform version, designed to be used in small real-time simulations such as the case of CODA's validation exercise. The validation platform will consist of a single En-route controller working position allowing a controller to monitor the simulated air traffic in a dedicated airspace sector, a pseudo-pilot position, and a monitoring computer.

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

Stakeholder	Involvement	Why it matters to the stakeholder
Users (ATCOs from ANSPs, NM).	ATCOs will be firsthand involved in the exercise. They will use a small-scale simulation to assess the tasks and the mental state prediction models integrated.	The exercise matters to ATCOs because it helps to refine the functionality of the models which ultimately can help adjust ATCOs' workload and optimize their mental state.

Scientific community (Universities, research institutions, EU projects, educational institutions)	Running tests on Mental models with end users	To see if the mental models are working according to the expectations or not (if the predictions are accurate, adjust to the ATM environment...)

Table 22: stakeholders' expectations exercise #02

5.2.3 Validation Objectives

SESAR solution validation objective	SESAR solution success criteria	Coverage and comments on the coverage of the SESAR solution validation objective in exercise #02	Exercise validation objective	Exercise success criteria
OBJ-CODA-TRL2-ERP-001: Assess the impact of the CODA system on ATCO's human performance in nominal conditions.	Success criteria (same as in section ***¡Error! No se encuentra el origen de la referencia.***)	**Partially covered,** because full functionalities of the CODA Solution will be available only at validation exercise #03	OBJ-CODA-TRL2-ERP-EXE02-001: Assess if the neurophysiological-based mental states indexes correlate with the ones provided by the prediction models (for all the selected mental states: workload, stress, vigilance, fatigue) (Prediction reliability)	CRT-CODA-TRL2-ERP-EXE02-001: The tendency of ATCOs' current mental workload indexes/ questionnaires is coherent with the ones predicted by the mental state models integrated. CRT-CODA-TRL2-ERP-EXE02-002: The tendency of ATCOs' current stress indexes/ questionnaires is coherent with the ones predicted by the mental state models integrated. CRT-CODA-TRL2-ERP-EXE02-003: The tendency of ATCOs' current

				vigilance indexes/ questionnaires is coherent with the ones predicted by the mental state models integrated. CRT-CODA-TRL2-ERP-EXE02-004: The tendency of ATCOs' current mental fatigue indexes/ questionnaires is coherent with the ones predicted by the mental state models integrated.

Table 23: validation objectives addressed in exercise #02

5.2.4 Validation scenarios

The selected validation scenario is an En-route sector from Spanish airspace. The Spanish airspace has been selected due to CRIDA's accessibility of the traffic data needed to perform the verification of the different models.

5.2.4.1 Airspace information

Madrid ACC, LECM, Figure 7, is a class C airspace categorized as a medium density/complexity environment. The main communication means between controllers and flight crew is radio communication.

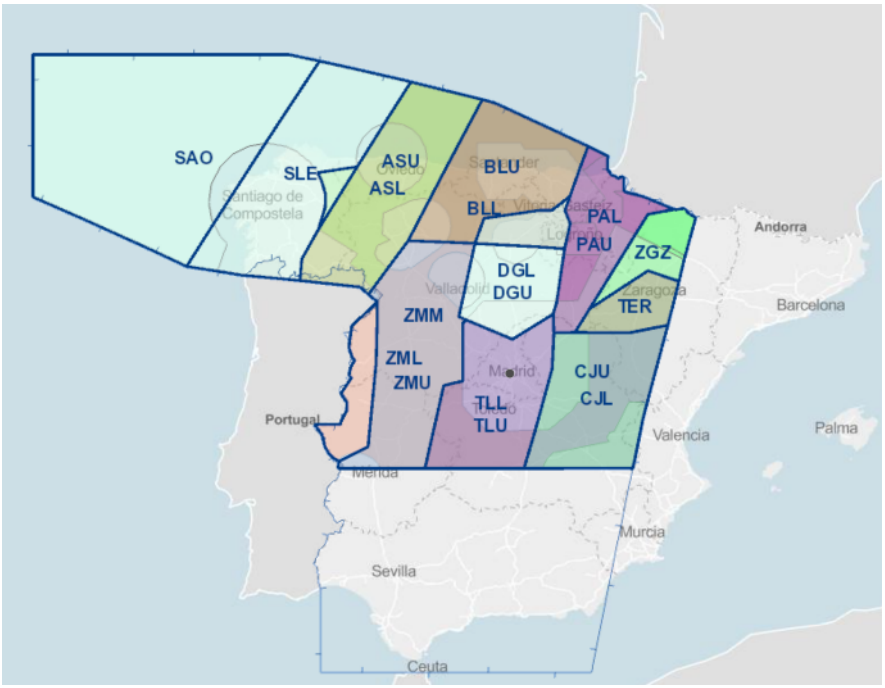

Figure 8: Madrid ACC elementary sectors. Source: Insignia

The selected sector is Domingo Alto, DGU, which provides control service to all aircraft from FL345 to FL 660, **¡Error! No se encuentra el origen de la referencia.**8.

Figure 9: exercise #02 simulated sector

The sector will be simulated by one executive controller in a single operation position (without a planning controller).

The scenario has a high percentage of overflight traffic in the north/south axis with traffic from/to central Europe to /from the Canary Islands and South America.

The sector also has traffic evolving in the arriving/departing routes from/to Adolfo Suarez Madrid-Barajas Airport (LEMD), the main airport in Spain with over 60 million passengers in 2023 according to AENA, the airport operator [10].

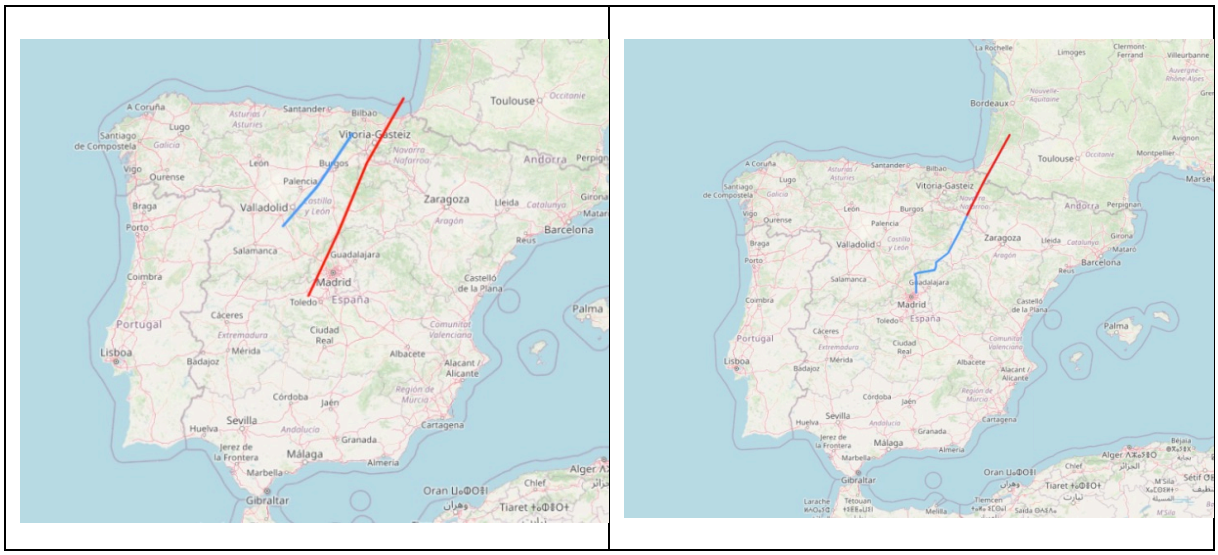

Figure 10: exercise #02 main traffic flows

The CODA system is not implemented in this exercise. Controllers execute their tasks as current operations.

No reference scenario will be used in this exercise. The mental model prediction and the actual mental measurements will be compared to gather information regarding the performance of the models.

The outcome of this exercise will be used as a reference for exercise 3 if needed.

5.2.4.2 Solution scenario(s)

As the objective of this exercise is to validate the mental model developed within CODA, the scenario will be common for all the RTS exercises but several traffic samples will be used. These traffic samples will be selected from real traffic from 2023 to meet a set of characteristics that will allow the CODA team to produce the target mental states in the controller. The resulting traffic samples will vary in terms of traffic demand levels and evolution in time. This will allow the team to test different scenarios ranging from low to high complexity:

- ES1 traffic sample presents low complexity and low traffic demand.
- ES2 traffic sample presents medium complexity, medium traffic.

- ES3 traffic sample presents high complexity and high traffic demand.
- ES4 traffic sample presents a steady increase in complexity and traffic demand.
- ES5 traffic sample presents a steady decrease in complexity and traffic demand.
- ES6 traffic sample presents a steep increase in complexity and traffic demand.
- ES7 traffic sample presents a steep decrease in complexity and traffic demand.
- ES8 traffic sample presents one steep increase followed by one steep decrease in complexity and traffic demand.
- ES9 traffic sample presents a steady increase followed by a steady decrease in complexity and traffic demand.

Additionally, another traffic sample will be prepared to be used during the initial runs for training purposes. This sample will be similar to the ES2 sample in terms of conflicts and traffic demand.

The other scenarios will be 50 minutes long. Two of the scenarios will be 90 minutes long to foster the appearance of fatigue in the controller if needed. Another option is to place the fatigue-inducing scenarios by the end of the day.

5.2.5 Exercise Validation Assumptions

Assumption ID	Assumption title	Assumption description	Justification	Impact Assessment
ASM-CODA-TRL2-ERP-003	Neurophysiological Assessment	Operators (i.e. en-route controllers) can be monitored in real-time	The project proposes a set of tools that can be used for assessing in real time the mental states of controllers. The actual use of similar tools is impacted by several considerations that are not technical ones. The use of alternative monitoring tools can influence the acceptability of such an assessment, as well as the data management process and privacy management of the acquired data.	High

Table 24: validation exercise #02 assumptions

5.2.6 Limitations and impact on the level of significance

The following limitations have been identified within this exercise:

- The sample of individuals is low, 2 controllers, which limits the representativeness of the outcome. The controllers that will participate will be selected by IFATCA with the demographics, expertise, and characteristics needed for the validation. The level of significance in this regard is considered as high.

- Due to time and budget constraints the exercise is limited to one week. The team has decided to prioritize the validation of the mental models by creating several different mental states over the number of individuals. Therefore, the level of significance regarding the reliability of the mental model prediction is considered as high, but the level of significance regarding the sensitivity of the exercise is low.

5.2.7 Validation exercise platform/tool and validation technique

5.2.7.1 Validation exercise platform/tool characteristics

As mentioned above, the platform that will be used is ESCAPE, a scalable EUROCONTROL ATM real-time simulation platform that supports small- and large-scale simulations in airspace design for both en-route and TMA. It is used in:

- the evaluation of new operational concepts and controller tools
- the pre-operational validation (live trial)
- the controller training,
- Research & Development.

CODA plans to use the ESCAPE Light platform version, designed to be used in small real-time simulations such as the case of CODA's validation exercise. One of the strengths of ESCAPE is its modular design. ESCAPE is a full PC/Linux platform based on Corba middleware for dynamic communications between ESCAPE components. ESCAPE provides a broad set of capabilities such as:

- Ground capabilities: SYSCO, AMAN, MCTD, MONA, TBS, RTA, STCA, APW, MTAPW, TCT, etc,
- Air capabilities: high fidelity aircraft performance, RTA, CDA, ASAS, Datalink, etc,
- HMI tool kit (ECHOES) which allows the development of specific and elaborated controller HMI(s),
- Interface with the industrial tool: AMAN (Thales Maestro, Harris OSYRIS), Thales TBS, etc,
- Interoperability with ATM systems using SVS or Pitch.

The simulator is located at CRIDA's premises at the Aeronautical and Space Engineering School at Politecnica University in Madrid. The implementation in the school features four role positions: the IPAS (Integrated Preparation and Analysis System) position, the Ground Server position, the controller's working position, CWP, (in the current version, up to three controllers can work simultaneously), and the pseudo-pilot position (currently, there are two pilot positions available). See Figures 10 to 13 below. In exercise 2 only on CWP and one pseudo-pilot position will be used.

Figure 11: ESCAPE simulator

Figure 12: ESCAPE - CWP

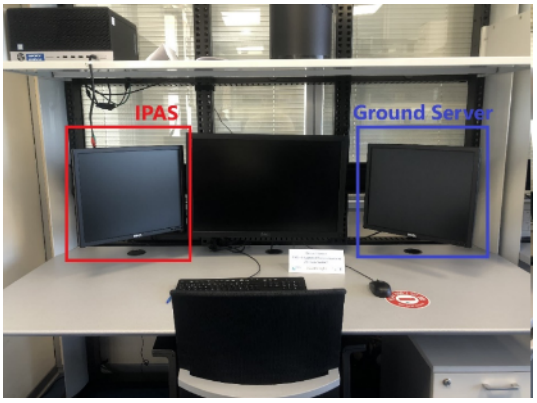

Figure 13: ESCAPE – IPAS and Ground Server

Figure 14: ESCAPE- pseudo-pilots

The IPAS position is used to prepare the scenarios. The ground server position launches and controls the simulation. Regarding the controller working position, in Figure 10, two positions can be used as an executive or executive+planning controller, while the ground server can be used as a feeder for the sectors if needed.

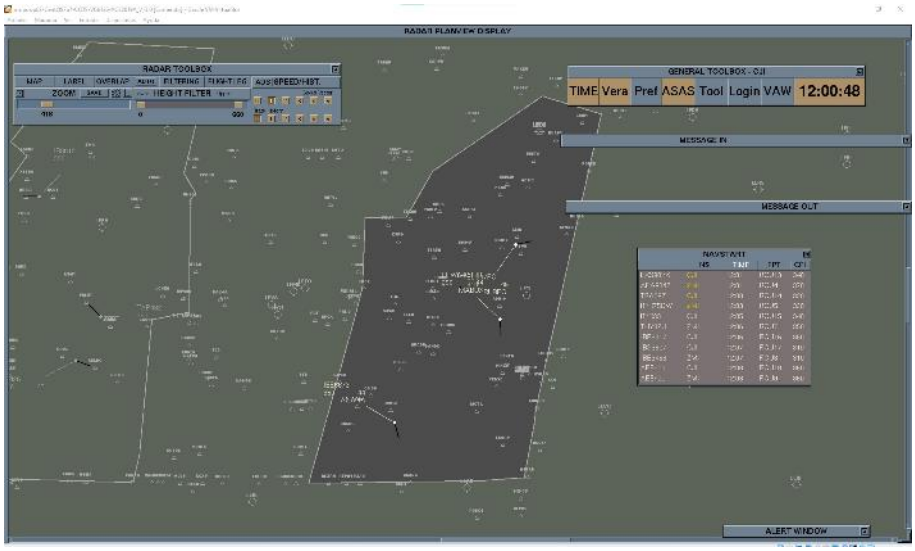

Figure 15: ESCAPE- CWP

The platform supports En-Route low, medium, and high complexity areas as well as TMA and approach. During the exercise next functional blocks of the validation platform will be used:

- Flight planning; Lyfe cycle Management and data distribution
- Controller Human-Machine Interaction management in ENR/APP
- Correlation management
- Flight coordination and transfer
- Short-term conflict detection
- Data recording

Apart from the validation simulator, the platform for this exercise is also composed of the next elements:

- A computer where the mental prediction models will run in parallel. These models will be fed with traffic information.
- A set of neurophysiological sensors that will record in real-time the reaction of the controllers.
- A tablet that will be used to gather feedback from controllers.

5.2.7.2 Validation exercise technique

As mentioned above, the tasks and the mental state prediction models integrated will be validated through a small-scale real-time simulation with Air Traffic Controllers (ATCOs) provided by IFATCA.

5.2.8 Data collection and analysis

5.2.8.1 Data and data collection methods

A mixture of quantitative and qualitative data shall be used to assess the objectives. For this level of maturity, it is necessary to assess the objectives with an emphasis on quantitative data, supported by qualitative data.

Data Collection Methods	Qualitative / Quantitative	Objective / Subjective
Neurophysiological sensors	Quantitative	Objective
Questionnaires	Qualitative & Quantitative	Subjective
Debriefings	Qualitative	Subjective
Over the Shoulder observations	Qualitative	Subjective
System Data Collection	Quantitative	Objective

Table 25: exercise #02 data collection

Quantitative data from **neurophysiological sensors** will record controller mental state indexes to be compared with the prediction from the models. These data will be the main input for the validation of the mental models.

Qualitative data will be used to complement, nuance, and provide context to the mental measurements:

- **Individual questionnaires:** standard questionnaires recognized by the scientific community will be used to gather feedback from the controllers regarding the mental states. The concrete questionnaires will be selected during the preparation phase, but among others, the following questionnaires will be considered: NASA TLX [11], Bedford Scale [12].
- **Debriefing sessions:** after each run the exercise will be discussed among all the participants (operational and simulation staff).
- **Over-the-shoulder observations**: **Direct and non-intrusive** over-the-shoulder observation will be carried out by human factors experts, during the runs. This non-intrusive observation will have the purpose **of providing detailed, complete, and reliable information** on the way the activity is carried out, especially if further commented on and discussed with the observed users.

Further quantitative data will be obtained from **system data** recorded during each session. These data contain information on radar tracks and flight performance.

NOTE1: These data from exercise #02 will be available for exercise #03 if needed for **comparison against a reference scenario** (more information on this are available below in EXE03 description).

NOTE2: To further validate the ATCO task prediction model and start developing the final integrated CODA system there will be a tablet-based interface to be used by an observer to **collect in real-time the tasks performed** by the ATCO, following a model like the one provided in ATON (CRIDA). In this validation exercise, it will be just stored, and the results will help partially validate this experimentation tool for the last validation exercises, where it will be integrated with the CODA system in real-time to facilitate real-time task allocation to the ATCO or the automation system. It should be emphasized, in any case, that the task prediction model is validated offline based on available traffic recordings, the results obtained here will not serve to validate the CODA system itself or any of its parts, but to progress in the integration and support tools for the final validation.

5.2.8.2 Analysis methods

Direct comparison between the predicted and recorded mental states will be used as an analysis method. Discrepancies between predicted and measured mental states will be moduled considering the feedback from controllers in questionnaires and debriefings.

If necessary, the thresholds in the mental model between the different states will be adjusted considering the outcome of the exercise. To perform this adjustment, the flight plans and radar tracks recorded during the exercise will be one of the information sources used as input.

5.2.9 Exercise planning and management

Exercise #02 validations are scheduled from the 21st to the 25th of October 2024. Two controllers will perform the nine scenarios. The order of the scenarios will be decided considering the objective (e.g. scenarios looking for fatigue will likely be performed by the end of the day). Each simulation exercise will take around 50 minutes and will be followed by a ten-minute questionnaire session.

Participants will be experienced controllers. Information regarding the platform and airspace will be provided in advance. After the introduction of the scope of the experiments, they will be trained on the simulation platform.

Table 26 presents the expected schedule of the week.

	Monday	Tuesday	Wednesday	Thursday	Friday
9:00	Intro	Briefing	Briefing	Briefing	Briefing
10:00	Test Run	Run#4	Run#8	Run#2	Run#6
11:00	Run#1	Run#5	Run#9	Run#3	Run#7
12:00	BREAK	BREAK	BREAK	BREAK	BREAK
12:30	Run#2	Run#6	Test Run	Run#4	Run#8
13:30	LUNCH	LUNCH	LUNCH	LUNCH	LUNCH
14:30	Run#3	Run#7	Run#1	Run#5	Run#9
15:30	Debriefing	Debriefing	Debriefing	Debriefing	Debriefing
16:00	End	End	End	End	End

Table 26: exercise #02-time planning

5.2.9.1 Activities

Phase	Activities	Responsibilities
Preparation	Development of the validation plan; preparation of the scenarios in the ESCAPE platform (selection of traffic samples; creation of scenarios); creation of simulation sessions,	Exercise Leader

	questionnaires, and debriefing guidelines; selection of controllers; creation of Escape and scenarios guidelines for controllers and pseudo pilot.	
	Development of the mental and task models mock-up; calibration of neurophysiological sensors and creation of questionnaires.	Validation Team
Exercise	Execution of the simulation exercises; execution of the prediction mental model and data gathering	Exercise Leader / Validation Team
Post-Exercise	Analysis of data and questionnaires and models update.	Exercise Leader / Project's Consortium

Table 27: exercise #02 Planning activities

5.2.9.2 Roles and responsibilities in the exercise

Actor	Role/responsibility
Exercise Leader (CRIDA)	Prepare the platform and scenarios to be carried out; prepare the manuals for controllers and pseudo-pilots; participate in elaboration on the questionnaires; ensure the execution of the exercise according to the schedule; present the objectives of the exercise to the controllers; train the controllers on the platform; launch the simulation exercises and the mental model predictions; record the data from the platform log; record controller's task, participate in the debriefing sessions; participate in the analysis for results.
ATCOs (IFACTA)	Getting familiar with the aerodromes and position Feedback during debriefings Complete questionnaires throughout the validation
Validation Team	Participate in elaboration on the questionnaires; present the objective of the project to the controllers; participate in the debriefing sessions and the analysis of results (DBL). Participate in elaboration on the questionnaires; explain the mental model and states to the controllers; participate in the debriefing sessions and the analysis of results (UGR). Explain the neuro-psychological index recording system to the controllers; take care of headset mounting and calibration; participate in the data gathering, debriefing session, and analysis of results (BS). Develop a tablet interface to collect tasks in real-time during the execution of the exercise (UPM).

Table 28: exercise #02 Role and Responsibilities

5.2.9.3 Time planning

The Table below presents the timeline where the preparation, execution, and analysis tasks start. The exercise will take place from the 21st to the 25th of October 2024.

Activity	Week (2024)

	1-04	8-04	…	7-10	14-10	21-10	28-10	4-11	…	16-12	23-12
Preparation											
Execution											
Analysis											

Table 29: detailed exercise #02-time planning

5.2.9.4 Identified risks and mitigation actions

Risks	Impact	Likelihood	Criticality (calculated based on likelihood and impact)	Mitigation actions
Risk 8: Controllers are not available for the validation dates.	3	1	4	Early identification of target participants Early invitation of participants.
Risk 13: High variability of mental measurements between the two participants	2	2	4	Debriefings and questionnaire feedback

Table 30: exercise #02 risks and mitigation actions (1-low, 2-medium, 3-high)

5.3 Validation exercise #03 plan - Final Validation of the CODA system

This section describes the validation plan for exercise #03, according to the validation exercise plan template provided hereunder.

5.3.1 Validation exercise description and scope

The third validation exercise aims at validating the CODA system, by validating the integrated models, the final HMI, and the final adaptation strategy.

To do so a human-in-the-loop simulation will be conducted, involving professional ATCOs as experimental subjects. These simulations will take place on the ESCAPE platform, a scalable EUROCONTROL ATM real-time simulation platform available at CRIDA facilities. During the simulations, controllers will engage in simulated operational scenarios to evaluate the effectiveness of the CODA system using different explainable Artificial Intelligence (XAI) solutions. The validation platform will consist of a single En-route controller working position allowing a controller to monitor the simulated air traffic in a dedicated airspace sector, a pseudo-pilot position, a monitor where HMI related to the XAI will be presented, and a monitoring computer. There is a simulated radio-communication system for communication between the controller and the pseudo-pilot during the simulation. The controller and the pseudo-pilot will behave according to real traffic situations to obtain representative conclusions from the validation exercise. It must be highlighted that the final platform will integrate, in a way in which subjects can realistically interact with them, the prediction models, the real-time measurement of current cognitive status, and the adaptation strategy. So, the platform will be able to simulate the adaptation strategy proposed by the AI (e.g. some tasks will start to be carried out by the system, warnings informing that network actions, such as splitting sectors, will be implemented, and so on). The AI tools enabled by the adaptation strategy are out of the scope of the project, they may be simulated using AI already available in the consortium or faked using the Wizard of Oz technique.

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

Stakeholder	Involvement	Why it matters to the stakeholder
Users (ATCOs, ANSPs, NM).	ATCOs will be directly involved in the exercise. Within the exercise, they will participate in human-in-the-loop simulations, engaging in simulated operational scenarios to assess the effectiveness of the CODA system. The validation	The exercise is significant for Air Traffic Controllers (ATCOs) because it provides them with the opportunity to offer early feedback regarding the CODA solution, which could potentially be adopted by them in the future. They can firsthand test and express their opinions on the overall raw functioning of the system, in terms of the Human-Machine Interface (HMI), as well as the system's

	platform will comprise a single En-route controller working position, enabling a controller to monitor simulated air traffic within a dedicated airspace sector, a pseudo-pilot position, and a monitoring computer.	capability to predict the controllers' mental states and tasks and adjust its strategy accordingly. This will enable them to assess the capability of the CODA solution in improving en-route capacity and enhancing or maintaining safety, human performance, and operational efficiency.
Scientific community (Universities, research institutions, EU projects, educational institutions)	Direct participation in the validation with end users	The exercise is significant for the scientific community to raise mutual awareness of approach results potentially resulting in further collaborations and to make students and academia aware of advanced AI solutions in ATM.
Institutional bodies (EU and EC, European Joint Undertakings, EASA, Policy makers, Regulatory and safety agencies, Standard making bodies, National bodies, Certification bodies).	Direct participation in the validation with end users	The exercise is significant for the Institutional bodies (EASA, EUROCONTROL) to help facilitate the progressive and safe introduction of AI solutions in ATM in a controlled way, by implementing actual examples of potential uses to help confidence be built on the users and regulators and helping define future safety/security/privacy constraints for its implementation.

Table 31: stakeholders' expectations exercise #03

5.3.3 Validation objectives

SESAR solution validation objective	SESAR solution success criteria	Coverage and comments on the coverage of the SESAR solution validation objective in exercise #03	Exercise validation objective	Exercise success criteria
OBJ-CODA-TRL2-ERP-001: Assess the impact of the	CRT-CODA-TRL2-ERP-001-10 Success criteria (same as in section	Fully covered	Same solution validation objectives	Success criteria (same as in section ***¡Error! No se encuentra el***

CODA system on ATCO's human performance in nominal conditions	***¡Error! No se encuentra el origen de la referencia.***)			***origen de la referencia.***)
OBJ-CODA-TRL2-ERP-002: Assess the impact of the CODA system on capacity in nominal conditions	CRT-CODA-TRL2-ERP-002-001 Success criteria (same as in section ***¡Error! No se encuentra el origen de la referencia.***)	Fully covered	Same solution validation objectives	Success criteria (same as in section ***¡Error! No se encuentra el origen de la referencia.***)
OBJ-CODA-TRL2-ERP-003: Assess the impact of the CODA system on operational efficiency in nominal conditions	CRT-CODA-TRL2-ERP-003-001-003 Success criteria (same as in section ***¡Error! No se encuentra el origen de la referencia.***)	Fully covered	Same solution validation objectives	Success criteria (same as in section ***¡Error! No se encuentra el origen de la referencia.***)
OBJ-CODA-TRL2-ERP-004: Assess the impact of the CODA system on safety in nominal conditions	CRT-CODA-TRL2-ERP-004-001 Success criteria (same as in section ***¡Error! No se encuentra el origen de la referencia.***)	Fully covered	Same solution validation objectives	Success criteria (same as in section ***¡Error! No se encuentra el origen de la referencia.***)
OBJ-CODA-TRL2-ERP-005: Assess the impact of the CODA system on cost efficiency in nominal conditions	CRT-CODA-TRL2-ERP-005-001 Success criteria (same as in section ***¡Error! No se encuentra el origen de la referencia.***)	Fully covered	Same solution validation objectives	Success criteria (same as in section ***¡Error! No se encuentra el origen de la referencia.***)

Table 32: Validation objectives exercise #03

5.3.4 Validation scenarios

The scenarios will present different types of traffic and complexity. They will be selected from the ones created for exercise 02, considering the adaptative strategy proposed within CODA, and the limitations of the simulation platform to present the strategy and integrate the different systems. The initially selected scenarios are:

- ES4 traffic sample presents a steady increase in complexity and traffic demand.

- ES6 traffic sample presents a steep increase in complexity and traffic demand.
- ES8 traffic sample presents one steep increase followed by one steep decrease in complexity and traffic demand.

The scenarios will be lightly adapted if necessary for the validation of the CODA system.

5.3.4.1 Airspace information

The selected Validation scenario is an En-Route sector from Spanish airspace. The sector is the same as in Exercise 2 (see section 5.2.4 Validation scenarios). The sector is Domingo Alto, DGU, which provides control service to all aircraft from FL345 to FL 660, in light blue in the figure below.

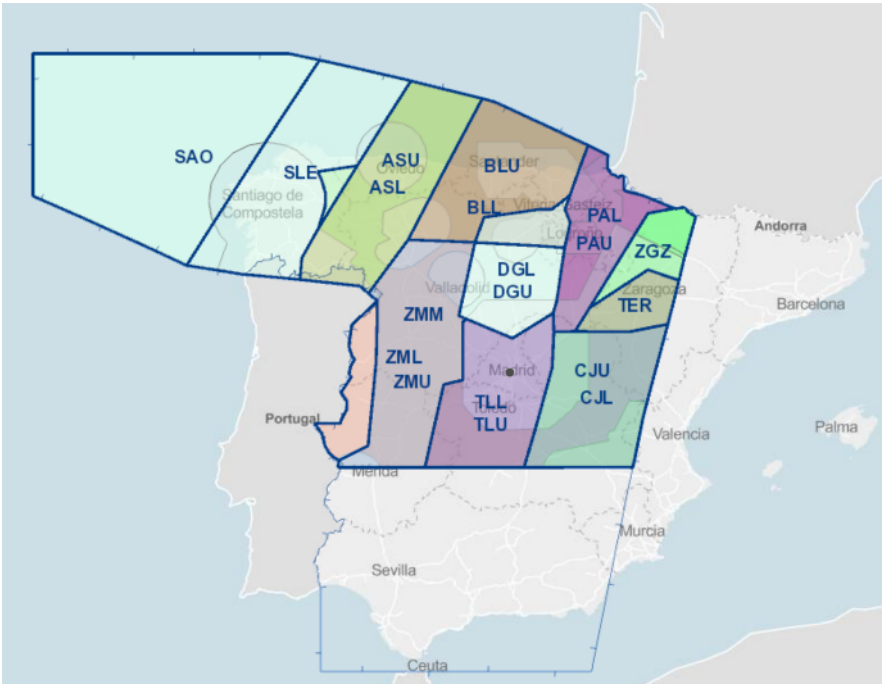

Figure 16 Madrid ACC elementary sectors. Source: Insignia

5.3.4.2 Reference scenario(s)

Current operation reference scenarios are the ones executed during exercise 02 where the mental model was validated.

5.3.4.3 Solution scenario(s)

The solution scenario consists of the joint team controller and Artificial intelligence assisted by the CODA system. The CODA system is composed of the following elements:

- Controller prediction task model
- Controller mental state prediction model
- Envelop index
- Adaptative strategy

- Adaptative HMI
- Controller mental state measurement

In the solution scenario, the envelop index uses the prediction models to monitor the performance of the controller plus artificial intelligence teaming and proposes strategies to the controller to optimize the joint performance. The strategies proposed are presented to the controller using an HMI that fosters the explainability of the solution. The reaction of the controller to the proposals is monitored to ensure the desired effect is achieved.

CODA proposes a solution where all the elements will be integrated and adaptable to the team being supported: controller role and Artificial Intelligence supporting the human. The solution being implemented in this exercise is limited by the level of maturity and integration capabilities of the platform and models.

It should be noted that the Artificial intelligence-based tools supporting the ATCOs are out of the scope of the solution. This means that AI assistants activated or deactivated by the CODA systems are not developed within the project: AI solutions available within the consortium will be used or AI assistants will be simulated in a Wizard of Oz approach. For more information, refer to D5.1 Adaptation and Human-AI interaction strategy.

5.3.5 Exercise Validation Assumptions

The assumptions that will guide the definition and execution of the exercise are those identified for the entire solution, as outlined in section 4.4.

5.3.6 Limitations and impact on the level of significance

The following characteristics have been identified in this exercise:

- The sample of individuals is 4 controllers, which is considered adequate representativeness of the viability of the concept considering the level of maturity of the project, TRL2. Nevertheless, the sample is low for statistical significance.

The controllers that will participate will be selected by IFACTA and ENAIRE with the demographics, expertise, and characteristics needed for the validation. The level of significance in this regard is considered as high.

The platform used, ESCAPE light, is a light version of the complete platform adequate to simulate concepts with low levels of maturity. Nevertheless, the platform presents some limitations, especially regarding the integration of external systems. The limited integration of the CODA system elements in the exercise should be understood as a limitation of the platform and TRL level rather than a limitation of the CODA system proposed.

Due to time and budget constraints, the exercise is limited to one week. The team has decided to prioritize the number of samples over the thorough analysis of the possible mental states. Therefore, the level of significance regarding the reliability of the Coda system is considered as medium as well as the level of significance regarding the sensitivity of the exercise

5.3.7 Validation exercise platform/tool and validation technique

5.3.7.1 Validation exercise platform/tool characteristics

The validation platform will be composed of several pieces:

- The real-time simulator ESCAPE light. The simulator is adequate for exploring new ATM controller-related concepts. This simulator is the same one that will be used in exercise 2. See 5.2.7 Validation exercise platform/tool and validation technique. It is composed of two positions to prepare and supervise the exercise, and up to three CWP and 2 pseudo-pilots. For exercise #03, one of the CWPs will be used for the controller under evaluation and another CWP will be used to implement some of the Artificial Intelligence actions when needed.

- A computer where the mental prediction models will run. The computer will use as input information from the task prediction model, the strategy from the envelop-index, and provide, as output, the prediction on workload, stress, fatigue, and vigilance decrement.

- High-resolution monitors and touchscreens will be employed to provide clear, interactive displays, allowing operators to efficiently interact with the system. These tools will enable precise adjustments to the visual layout and functionality of the HMI. Usability testing stations equipped with advanced input devices, such as customized keyboards and control panels, will replicate the operational environment, allowing for realistic interaction scenarios. These devices will collectively ensure that the HMI is not only user-friendly but also capable of presenting transparent and accurate data visualizations critical for the operators' tasks.

- A tablet with a tablet-based interface to be used by an observer to collect in real-time the tasks performed by the ATCO, following a model like the one provided in ATON (CRIDA). This validation exercise will be integrated with the CODA system in real time to facilitate real-time task allocation to the ATCO or the automation system. The collected tasks will be provided to the task prediction system and the automation strategy through an asynchronous information distribution mechanism.

Apart from the validation simulator, the platform for this exercise is also composed of the next elements:

- A set of neurophysiological sensors that will estimate the ATCOs' mental states in real time while dealing with the ATC simulation. To achieve this capability, the ATCOs' brain (Electroencephalogram – EEG), heart (Photoplethysmography – PPG) activity, and skin (Electrodermal activity – EDA) conductance will be acquired during the task execution by using wearable, wireless, non-invasive and reliable sensors. In particular, the device for the EEG data collection is the Mindtooth Touch system (https://mindtooth-eeg.com/) with 8 water–-based electrodes (5 frontal and 3 parietal channels) and Bluetooth low-energy (BLE) connectivity. This system has been used in several real contexts as described in these works [14][15][16][17][18][19].
- For the PPG and EDA data collection, the Shimmer3 GSR+ (https://shimmersensing.com/product/shimmer3-gsr-unit/) will be used or the Research Ring (https://www.biopac.com/product/research-ring) according to the results coming from T4.2. In any case, the selected technology will guarantee reliable signal quality and comfort to be compliant with both ATC settings and validation requirements.

5.3.7.2 Validation exercise technique

The CODA system will be validated through a small-scale real-time simulation with Air Traffic Control Officers (ATCOs) provided by IFATCA and ENAIRE.

5.3.8 Data collection and analysis

5.3.8.1 Data and data collection methods

A mixture of quantitative and qualitative data shall be used to assess the objectives. For the CODA level of maturity, mainly qualitative feedback backed up with quantitative information will be provided.

Data Collection Methods	Qualitative / Quantitative	Objective / Subjective
Neurophysiological sensors	Quantitative	Objective
Questionnaires	Qualitative & Quantitative	Subjective
Debriefings	Qualitative	Subjective
Over the Shoulder observations	Qualitative	Subjective
System Data Collection	Quantitative	Objective

Table 33: exercise #03 data collection

Qualitative data will be used to gather feedback from end users, and controllers, from the CODA concept and its different components.

- **Individual questionnaires:** standard questionnaires recognized by the scientific community will be used to gather feedback from the controllers regarding the mental states. The concrete questionnaires will be selected during the preparation phase, but among others, the following questionnaires will be considered: NASA TLX [11], Bedford Scale [12], SHAPE Teamwork (STQ-s) [13]. Based on the result of the preliminary Scoping & Change Assessment, specific questionnaires will be developed (e.g. understanding of the CODA system and new procedures introduced) to evaluate human-machine teaming and understand the impact of the CODA system on human performance and safety.
- **Debriefing sessions:** after each run the exercise will be discussed among all the participants (operational and simulation staff), to incise in benefits and improvement points identified.
- **Over-the-shoulder observations**: **Direct and non-intrusive** over-the-shoulder observation will be carried out by human factors experts, during the runs. This non-intrusive observation will have the purpose **of providing detailed, complete, and reliable information** on the way the activity is carried out, especially if further commented on and discussed with the observed users.

Data coming from the **neurophysiological sensors** will consist of EEG, PPG, and EDA. This data will endow additional and objective information about ATCO's mental states while dealing with the validation scenarios. The temporal resolution of the mental states' indicators (mental workload, stress, vigilance, and mental fatigue) will be adjusted according to other measurements so that correlation analyses will be performed to validate the CODA system.

Further quantitative data will be obtained from **system data** recorded during each session. These data contain information on radar tracks and flight performance. These data will be used to analyse the performance of the system (AI + controller). Data from exercise #02 will be used for comparison when necessary.

5.3.8.2 Analysis methods

Data from the neurophysiological sensors will be used to estimate ATCO's mental state and analyse if the CODA model implementation has supported the controller/AI teaming as expected regarding workload, stress, fatigue, and vigilance decrement.

The questionnaires from the different runs will be analysed statistically and complemented with the feedback from debriefings, and observations

Radar tracks and flight performance are processed to gather information regarding capacity, operational efficiency, and cost efficiency. Comparison with current operations (exercise 2) will be performed to estimate the increase or decrease in performance.

5.3.8.3 Exercise planning and management

The simulation execution is planned for the week of 17th to 21st March 2025. Each simulation exercise will take around 50 minutes and will be followed by a ten-minute questionnaire session.

Participants will be experienced controllers. Information regarding the platform and airspace will be provided in advance. After the introduction of the scope of the experiments, they will be trained on the simulation platform.

The table below presents the planned activities:

	Monday		Tuesday		Wednesday		Thursday		Friday	
9:00	Intro (all ATCOs)		Briefing		Briefing		Briefing		Briefing	
10:00			Run# 1	Dig. Assist Support	Run#4	Dig. Assist Support	Test Run	Test Run	Run# 3	Dig. Assist Support
11:00			Debriefing		Debriefing		Debriefing test run		Debriefing	
12:00	BREAK									
12:30	Test Run	Test Run	Dig. Assist support	Run #2	Dig. Assist support	Run #5	Dig. Assis Support	Run #1	Dig. Assist Support	Run #4

13:30	LUNCH									
14:30	Debriefing Test Run		Debriefing		Debriefing		Debriefing		Debriefing	
15:30	Test Run	Test Run	Run#3	Dig. Assist Support	Run#6	Dig. Assist Support	Run#2	Dig. Assist Support	Run#5	Dig. Assist Support
16:30	Debriefing Test Run		Debriefing		Debriefing		Debriefing		Debriefing	
17:00	End									

Table 34: exercise #03-time planning

5.3.8.4 Activities

Phase	Activities	Responsibilities
Preparation	Lead validation plan description; lead questionnaires preparation, debriefing guidelines; support selection of controllers.	Exercise Leader
	Creation of CODA mock-up system; support Integration of models. Participate in validation description, questionnaires, and debriefing preparation. Preparation of the scenarios in the ESCAPE platform (selection of traffic samples; creation of scenarios and simulation sessions).	Validation Team
Exercise	Execution of the simulation exercises.	Exercise Leader / Validation Team
Post-Exercise	Analyse data and questionnaires and report on conclusions.	Exercise Leader / Project's Consortium

Table 35: Exercise #03 Planning activities

5.3.8.5 Roles and responsibilities in the exercise

Actor	Role/responsibility
Exercise Leader (DBL)	Ensure the execution of the exercise according to the schedule; participate in elaboration on the questionnaires; present the objective of the project to the controllers; lead the debriefing session and the analysis of results.

ATCOs (IFACTA & ENAIRE)	Getting familiar with the aerodromes and position Feedback during debriefings Complete questionnaires throughout the validation
Validation Team	Participate in elaboration on the questionnaires; participate in the debriefing sessions and the analysis of results (ENAC, BS, NLR). Participate in elaboration on the questionnaires; explain the mental model and states to the controllers; participate in the debriefing sessions and the analysis of results (UGR). Participate in elaboration on the questionnaires; launch the task prediction model, and the task collection application; participate in debriefing sessions and in the analysis of results (UPM). Prepare the platform and scenarios to be carried out; prepare the manuals for controllers and pseudo-pilots; participate in elaboration on the questionnaires; train the controllers on the platform; launch the simulation exercises and the mental model predictions; record controller's task in the application; record the data from the platform log; participate in the debriefing sessions and the analysis of results (CRIDA). Explain the neuro-psychological index recording system to the controllers; take care of headset mounting and calibration; participate in the data gathering (BS).

Table 36: Exercise #03 Roles and Responsibilities

5.3.8.6 Time planning

The Table below presents the timeline where the preparation, execution, and analysis tasks start. The exercise will take place from the 17th to the 21st of March 2025.

Activity	Week (2025)										
	6-01	13-01	…	3-03	10-03	17-03	24-03				
Preparation											
Execution											
Analysis											

Table 37: detailed exercise #03-time planning

5.3.8.7 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
Risk 8: Controllers are not available for the validation dates.	3	1	4	Early identification of target participants Early invitation of participants.

Risk 7: Platform limitations impact feedback on the CODA concept.	1	2	3	Early coordination of inputs/outputs between different models. -Identification of scenarios that support the concept and can be simulated

Table 38: exercise #03 risks and mitigation actions (1-low, 2-medium, 3-high)

6 References

6.1 Applicable documents

SESAR solution pack

[1] ARTIMATION (2022). Final Project Report. SESAR JU. https://cordis.europa.eu/project/id/894238

[2] MAHALO (2022). Final Project Report. SESAR JU. http://mahaloproject.eu/ and https://cordis.europa.eu/project/id/892970/results

[3] MINIMA (2018). Final Project Report. SESAR JU. https://cordis.europa.eu/project/id/699282

[4] STRESS (2018). Final Project Report. SESAR JU. https://cordis.europa.eu/project/id/699381

[5] COTTON (2018). Final Project Report. SESAR JU. https://cordis.europa.eu/project/id/783222. doi: 10.3030/783222

[6] eCOMMET (2019). Final Project Report. SESAR JU. https://cordis.europa.eu/project/id/731730

Content development

[7] CODA D2.2 Operational Services and Environment Description (OSED), ed. 00.02.00

[8] CODA D2.1 Functional Requirements Document (FRD), ed. 00.02.00

6.2 Reference documents

[9] EUROCAE ED-78A Guidelines for Approval of the Provision and Use of Air Traffic Services supported by Data Communications, December 2020.

[10] Estadisticas de trafico aereo. Año 2023. Aeropuerto Adolfo Suarez Madrid-Barajas. AENA. https://www.aena.es/es/estadisticas/inicio.html. Webpage accessed on May 2024.

[11]Nasa Task Load Indez (TLX). Human Performance Group. NASA Ames Research Center. Moffet Field.

[12]Roscoe, Ellis. A subjective rating scale for assessing pilot workload in flight, Royal Aircraft Establishment, 1990

[13]FT Durso; C.A. Hackworth; T.R. TruittDurso, F. T., Hackworth, C. A., Truitt, T. R., Crutchfield, J., Nikolic, D., & Manning, C. A. (1999). Situation Awareness as a predictor of performance in en-route air traffic controllers (Final Report No. DOT/FAA/AM-99/3). Washington: Federal Aviation Administration.

[14]Ronca, V., Brambati, F., Napoletano, L., Marx, C., Trösterer, S., Vozzi, A., ... & Di Flumeri, G. (2024). A Novel EEG-Based Assessment of Distraction in Simulated Driving under Different Road and Traffic Conditions. Brain Sciences, 14(3), 193.

[15]Borghini, G., Giorgi, A., Ronca, V., Mezzadri, L., Capotorto, R., Aricò, P., ... & Babiloni, F. (2023, November). Cooperation and mental states neurophysiological assessment for pilots' training and expertise evaluation. In 2023 IEEE International Workshop on Technologies for Defense and Security (TechDefense) (pp. 77-82). IEEE.

[16]Di Flumeri, G., Giorgi, A., Germano, D., Ronca, V., Vozzi, A., Borghini, G., ... & Aricò, P. (2023). A neuroergonomic approach fostered by wearable EEG for the multimodal assessment of drivers trainees. Sensors, 23(20), 8389.

[17]Ronca, V., Uflaz, E., Turan, O., Bantan, H., MacKinnon, S. N., Lommi, A., ... & Borghini, G. (2023). Neurophysiological Assessment of An Innovative Maritime Safety System in Terms of Ship Operators' Mental Workload, Stress, and Attention in the Full Mission Bridge Simulator. Brain Sciences, 13(9), 1319.

[18]Pugh, J., Goman, M., Abramov, N., Borghini, G., De Vissche, I., Granger, G., ... & Rooseleer, F. (2023). Airfox Upset Prevention & Recovery Training (UPRT) Flight Simulation with Wake Vortex Encounter Events.

[19]Borghini, G., Ronca, V., Aricò, P., Di Flumeri, G., Giorgi, A., Bonelli, S., ... & Babiloni, F. (2023). Teamwork objective assessment through neurophysiological data analysis: a preliminary multimodal data validation. In Neuroergonomics and Cognitive Engineering (Vol. 102).