Performance based requirements for advanced automation

Deliverable ID:	**D4.2**
Project acronym:	**HUCAN**
Grant:	**101114762**
Call:	**HORIZON-SESAR-2022-DES-ER-0**
Topic:	**HORIZON-SESAR-2022-DES-ER-01-WA1-2**
Consortium coordinator:	**Deep Blue**
Edition date:	**29 November 2024**
Edition:	**01.00**
Status:	**Official**
Classification:	**PU**

Abstract

This document addresses the setting of appropriate performance-based requirements for advanced automation. It aims to focus on desired, measurable, outcomes in terms of Key Performance Indicators (KPIs), rather than on a prescriptive and compliance-based approach to approval and certification. A holistic approach will be targeted, enabling derivation of performance-based requirements to demonstrate safety of automation during normal, impaired operation and recovery phases of service provision.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

Authoring & approval

Author(s) of the document

Organisation name	Date
NLR | Mariken Everdij	26.11.2024
NLR | Sybert Stroeve	26.11.2024
DBL | Elisa Spiller	26.11.2024
EUI | Marco Sanchi	26.11.2024
EUI | Carla Bonacci	26.11.2024

Reviewed by

Organisation name	Date
NLR | Henk Hesselink	22.11.2024
DBL | Paola Lanzi	26.11.2024

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

Organisation name	Date
Deep Blue | Paola Lanzi	28.11.2024
NLR | Mariken Everdij	28.11.2024
EUI | Giuseppe Contissa	28.11.2024
CIRA | Gabriela Gigante	28.11.2024
DLR | Mohsan Jameel	28.11.2024
D-FLIGHT | Edoardo Fornaciari	28.11.2024

Rejected by[2]

Organisation name	Date

Document history

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

[2] Representatives of the beneficiaries involved in the project

Edition	Date	Status	Company Author	Justification
00.01	21.06.2024	Draft	Mariken Everdij | NLR	ToC
00.02	06.08.2024	Draft	Mariken Everdij | NLR	Criteria, Use Cases, initial set of KPI
00.03	08.11.2024	Draft	Elisa Spiller | DBL Marco Sanchi | EUI Sybert Stroeve | NLR	Scope section added, Liability material added, Overall input provided
00.04	13.11.2024	Draft	Mariken Everdij | NLR	Document restructured
00.05	15.11.2024	Draft	Mariken Everdij | NLR Sybert Stroeve | NLR Elisa Spiller | DBL	Significant contributions to chapter 2, 3, 4 and appendices
00.06	22.11.2024	Draft	Mariken Everdij | NLR	Version for internal review
00.07	26.11.2024	Draft	Mariken Everdij | NLR Sybert Stroeve | NLR Elisa Spiller | DBL	Internal review incorporated
00.08	27.11.2024	Draft	NLR, DBL	Final for approval
01.00	29.11.2024	Official	NLR, DBL	Final

HUCAN

HOLISTIC UNIFIED CERTIFICATION APPROACH FOR NOVEL SYSTEMS BASED ON ADVANCED AUTOMATION

HUCAN

This document is part of a project that has received funding from the SESAR 3 Joint Undertaking under grant agreement No 101114762 under European Union's Horizon Europe research and innovation programme.

Table of contents

List of figures

List of tables

1 Introduction

1.1 Objective

For advanced automation to be safely integrated into air traffic operations, performance-based requirements (PBRs) and safety performance indicators (SPIs) as well as other key performance indicators (KPIs) must be defined and monitored to ensure operational effectiveness and maintain high safety standards.

This document addresses the setting of appropriate PBRs for advanced automation. These requirements define how automation should function and perform under various conditions. The focus is on measurable outcomes in terms of SPIs as well as other KPIs, rather than on a prescriptive and compliance-based approach to approval and certification.

A holistic approach will be targeted, enabling derivation of PBRs and KPIs to demonstrate safety of advanced automation with or without AI, during normal phases, impaired operation, and recovery phases of service provision. The approach should be usable for e.g., cockpit automation as well as automation of air traffic management. A specific issue to address is that novel methods like Machine Learning (ML) may learn and adapt their behaviour (in real time) during operation hence the exact behaviour of the automation cannot be predicted in advance. How is safety ensured if not all situations and variations of parameters can be anticipated during the design phase? A possible solution, to be investigated within this task, is to establish specific and additional requirements for safety oversight – of operations and systems based on advanced automation – by the aviation authorities.

1.2 Organisation

This report is organised as follows:

- Chapter 2 aims to set the scene: What do we mean by holistic approach, what would be its scope, how would it relate to EASA's approach for certification of AI and automation. The chapter adopts the Objectives from EASA's Concept Paper giving Guidance for Level 1 & 2 machine-learning applications, and analyses which of these objectives are relevant and applicable to four use cases.
- Chapter 3 explains what we mean by PBR and KPI in this context, why we would need them, and what would be criteria for good requirements, and derives appropriate PBRs and KPIs for advanced automation. The result is a list of KPIs and associated milestones for each of EASA's Objectives.
- Chapter 4 provides conclusions and recommendations.
- Chapters 5 and 6 lists references to sources material used, and provides a list of acronyms.
- The appendices provide input to the main chapters.

2 Scope - Setting the scene

The aim of this chapter is to explain what we mean by holistic approach, what its scope would be, and how it would relate to EASA's approach for certification of AI and automation. To do this, section 2.1 provides context on the need for a novel holistic approach for certification of advanced automation. Section 2.2 introduces the work in EASA's Concept Paper giving Guidance for Level 1 & 2 machine-learning applications (April 2024). This work provides a list of objectives that in this document will form a red thread towards the development of PBR and KPI for advanced automation. Section 2.3 explains how these objectives are used as input to a relevance and applicability analysis on four use cases. Section 2.4 summarises the use cases, Section 2.5 introduces the relevance and applicability analysis, and Section 2.6 gives the results. Section 2.7 gives conclusions.

2.1 Towards a holistic approach for certification of advanced automation

The implementation of advanced automation and AI in operational contexts claims for a paradigm shift in the way technology and process design are approached. Higher levels of automation enabled by these solutions have a profound impact on human-machine interaction, often leading to new forms of collaboration between operators and systems, as well as among operators in scenarios of computer supported cooperative work (CSCW). Recognising that automation places new cognitive and operational demands on human operators, a holistic approach to certification hence aims to encompass the entire operational environment, taking into account all relevant dimensions.

Supported by authoritative guidelines (EASA, 2023; SESAR, 2024b), HUCAN aims to provide an operative framework to support safety, security, ethics and human factors assessments towards certification, starting from the early R&D phases of solutions using advanced automation or AI. The intention is to facilitate the gradual alignment of concepts and technologies with final certification requirements along their development, proactively addressing the relevant issues at the due level of maturity.

This framework is based on the preliminary research carried out by the project (HUCAN, 2024(a, b), on current certification methods available in aviation and on innovative approaches developed for AI and advanced automation, in general. The earlier research, in particular, highlighted the following aspects.

Consolidated certification practices for aviation focus on reliability, relying on traditional assessment methods like fault trees and failure mode and effect analysis, which have their origins in assurance schemes for physical components that may fail/break and for which statistical quality control approaches can be applied. These approaches are known to have limitations for assessing and controlling the safety impact of advanced automated and AI-based systems. The safety impact of a particular component in these systems indeed depends on the dynamic interactions with other systems, humans working with operational procedures, and contextual conditions. This entails that the primary focus in reviewing certification approaches for advanced automation and AI-based solutions is undoubtedly safety, promoting a safety management cycle that can dynamically assess and ensure adequate and effective safety standards during both the design and the operational phases.

While the technical aspects of AI safety and the criteria for assessing them are important, they are not the only priorities. It is also important to address potential risks associated with over-reliance and reduced human autonomy, such as automation bias, particularly in the context of human factors

analysis. Advanced automation systems, including AI-driven architectures that include human operators, must be designed to avoid these pitfalls. Incorporating the human element is therefore critical to maintaining safety and trust in aviation automation. Integrating these considerations promotes trust, ethics and the overall effectiveness of AI systems in aviation. It is therefore imperative that future certification efforts prioritise these human-centred aspects to create a robust regulatory framework.

This human-centred, cross-cutting approach to designing and implementing solutions with high levels of automation in critical sectors is strongly supported by the EU AI Strategy and the EU AI Act (Reg. (EU) 2024/1689), as well as by EASA's AI Roadmap 2.0 (EASA, 2023). Indeed, all three initiatives - albeit with different levels of granularity - make multidisciplinary collaboration a key compliance milestone that should be integrated throughout the value chain and lifecycle of systems. This directive implicitly underscores the importance of a holistic approach: multidisciplinarity should be seen not as compartmentalised but as an integrated method, addressing various aspects from diverse perspectives and ensuring cohesive alignment among these viewpoints throughout the process, up to the final validation of systems.

In this context, HUCAN is pursuing an approach to facilitate the integration of this holistic perspective throughout the development phases of advanced automation. While effective compliance with legal and regulatory requirements is not mandated until solutions achieve a high level of maturity and are prepared for real-world testing, regulations implicitly encourage gradual alignment from the initial design stages. This proactive approach aims to prevent significant setbacks later on, which can arise from early poor conceptual choices.

This initiative is not taking place in isolation, but is aligned with the existing frameworks established by EASA, with the intention of infusing the spirit of certification within funded research projects.

In this regard, EASA is promoting a proactive strategy to facilitate the approval or certification of products, parts, and appliances that incorporate AI/ML technologies. To assist applicants in introducing AI/ML into systems used for safety- or environment-related applications across all domains covered by the EASA Basic Regulation (Regulation (EU) 2018/1139), the Agency is providing a set of practical objectives. Accordingly, the objectives identified by EASA cover not only AI assurance, safety, and risk mitigation, but also call for rethinking and redesigning human factors paradigms, especially before redefining authority in human-AI teaming. A preliminary AI trustworthiness analysis is introduced as a key step in assessing certification requirements, enabling early and comprehensive evaluation of technical and operational risks across different levels of automation.

Building on this background, HUCAN aims at contributing to this ongoing rule-making process by promoting the holistic attitude of this approach on a diachronic dimension, testing the building blocks objectives in concepts having different levels of automation and maturity. Indeed, the process of aligning values and requirements throughout the gradual journey toward certification or authorization remains unclear. While certain technical and organisational aspects must be addressed from the early design stages, others can be tackled later in the process.

In light of these objectives, the following paragraphs present the approach adopted in this document to evaluate the objectives outlined by EASA, aiming to detect redundancies, identify gaps, and highlight areas warranting further exploration. Accordingly, the assessment preparation begins with an overview of the objectives prescribed for the various levels of AI defined by the EASA taxonomy.

2.2 How to use EASA objectives for research and development purposes

Notably, the regulatory approach embraced by EASA is founded on different levels of AI, which contribute to characterising the level of automation of given solutions. Level 1 AI corresponds to "cognitive human assistance" and includes human augmentation functions (level 1A) and human assistance (level 1B). Level 2 AI focuses on human-AI teaming, distinguishing between cooperation (level 2A) and collaboration (level 2B). Eventually, Level 3 AI considers advanced automation, encompassing safeguarded (level 3A) and non-supervised applications (level 3B).

In view of the characterisation of the solutions, the EASA AI Roadmap currently covers different objectives and expected means of compliance, which gradually increase according to the level of automation to be achieved. So far, the EASA Concept Papers, Issues 1 and 2, comprehensively define 142 objectives (including corollary objectives), ranging from Level 1A to Level 2B. The definition of specific objectives for systems incorporating advanced automation in the forms described at levels 3A and 3B is underway and a dedicated concept paper is expected in 2025.

The objectives and anticipated means of compliance support the goals of the individual building blocks within the EASA AI Trustworthiness Framework and, more broadly, advance the pursuit of a human-centred approach to AI in aviation. The table below provides an overview of the objectives covered by the respective building blocks (BBs) and shows how many objectives are included at the different levels of automation, from Level 1A to Level 2B.

EASA BBs	Objectives Classification and Codes	1A	1B	2A	2B
Trustworthiness Analysis	Characterization *(CO/CL)*	7	7	7	7
	Safety assessment *(SA)*	3	3	3	3
	Information and security *(IS)*	3	3	3	3
	Ethics-based assessment *(ET)*	N/A	N/A	8	8
AI Assurance	Learning assurance *(DA, DM, LM, IMP, CM, QA, RU, SU)*	56	56	56	56
	Development and post-ops AI explainability *(EXP)*	9	9	9	9
Human Factors for AI	AI operational explainability *(EXP)*	2	10	10	10
	Human-AI teaming (HF)	N/A	N/A	5	11
	Modality of interaction and style of interface (HF)	N/A	N/A	6	16
	Error management *(HF)*	N/A	N/A	5	5
	Failure management *(HF)*	N/A	N/A	N/A	4
AI safety risk mitigation	AI safety risk mitigation *(SRM)*	2	2	2	2
	Organisation *(ORG)*	8	8	8	8
	Tot.	**90**	**98**	**124**	**142**

Table 1: EASA AI Roadmap 2.0, Concept Paper - Issue 2, Objectives

As specified by EASA, in principle, the trustworthiness analysis is always required and all its elements are important prerequisites for the development of any system developed with or embedding AI. The objectives belonging to the other three building blocks indicate how the depth of guidance could be adapted depending on the classification of the application (EASA, 2024b).

From a practical perspective, the Agency clarified that the purpose of this framework, along with the guidance provided under its AI Roadmap, is to offer stakeholders involved in the research and development of AI-based solutions - whether grounded in ML or other advanced automation enablers - a foundational set of references to guide strategic development choices. The four building blocks and their objectives should therefore be understood and considered in light of their inherent interdependence (EASA, 2023).

Looking at the holistic approach of this framework, it is interesting to note that while the technical requirements apply to all applications, specific human factors objectives only apply to solutions with a higher level of automation (Level 2A/2B). More details on the distribution of objectives and the expected means of meeting them are given in the extended version of this table, as reported in Appendix A. That assessment confirms that EASA's attention to date has been focused on the technical requirements for reconciling AI and advanced automation-based solutions with operational and societal expectations, while aspects related to human factors and operations could be further explored.

In light of these objectives, the following paragraphs describe the approach adopted in this report to test the objectives outlined by EASA so far, with the aim of detecting redundancies and gaps and identifying the areas that could deserve further exploration.

2.3 Methodological approach used for this chapter

One of the starting points for developing a holistic approach in line with the evolving regulatory framework is the material produced by EASA under the AI Roadmap 2.0 (EASA, 2023), together with the deliverables published to date for the application of ML (EASA, 2024a) and AI Levels 1 and 2 (EASA, 2024b).

As noted by the Agency, the objectives and anticipated means of compliance outlined in these documents aim to progressively align the development of solutions based not only on ML, but also on advanced automation enabled by other technologies with certification objectives and requirements (EASA, 2024b,c). Furthermore, in view of the ongoing renewal of the current aviation regulatory ecosystem, a broad process of participation and discussion on the adequacy, relevance and comprehensiveness of the work done so far in relation to specific case needs is encouraged. (EASA, 2024b).

HUCAN uses a case-based approach to assess the relevance and applicability of predefined objectives, considering the varying levels of automation and maturity in the project's Use Cases (UCs). The goal of the remainder of this chapter is to evaluate whether and how these objectives can be practically implemented during development. The evaluation aims to identify redundancies and pinpoint areas needing further integration to ensure comprehensive coverage of complementary building blocks for a broad scope of advanced automation techniques, including ML as well as other methods.

The methodology proposed in this chapter consists of 4 steps and covers all building blocks in the EASA Roadmap. After a brief description of the objectives and specificities of the UCs (section 2.4), we define the scope of the analysis taking into account the preliminary characterisation of each concept and the corresponding TRL (section 2.5). Against this background, we analyse the actual relevance of the EASA objectives in each scenario, taking into account the enabling technologies, the nature of the human-machine interaction and the prospective impact of the solutions in operational procedures. In parallel, we assess the applicability of the objectives in the development phase at the current TRL, with the aim of evaluating if some issues can and should be addressed during the development process to progressively align the solution with certification requirements.

2.4 A summary of the UCs

For the sake of clarity, it is essential to bear in mind that HUCAN includes 4 UCs that focus on capacity on demand and each of them includes technical solutions based on AI and advanced automation. In D4.1 (HUCAN, 2024c), the project defined the respective scenarios of these UCs and assessed the levels of automation of functions and concepts according to the EASA/SESAR integrated taxonomy for AI.

The main characteristics of the UCs can be summarised as follows:

- **UC1 – Dynamic configuration of airspace**

UC1 focuses on dynamic airspace sectoring with the aim of improving the use of the medium airspace by dynamically optimising the airspace sector. More specifically, this case study aims to support the design of the sector collapsing/decollapsing configuration for a given planned traffic in a performance-based environment for air traffic controller (ATCO) workload optimisation, capacity optimisation and flow management optimisation.

In this UC, the role of advanced automation is to provide automated support for the design of a new ATM concept to achieve the required performance levels. Accordingly, this case considers the development of two potentially complementary solutions: a simulation-based and a scenario-based decision support system.

The first of these is a simulation-based decision support system based on computational intelligence techniques which allow to carry out offline simulations for performing what-if analyses of ATM changes and for supporting the design of new solutions aimed at ATM system optimisation. According to the analysis carried out in D4.1, this solution is classified at level 1A of the EASA AI taxonomy, as it is intended for human assistance, more specifically human augmentation functions.

The second is a scenario-based decision support system, which enables a clearer understanding of the scenarios, relying on a description of the reference operating environment, including: a set of actors; a set of available actions; a set of processes; the relationships between the previous elements and their formalisation as a flow of information, representing the dynamics to allow the system to perform a mission or a service. The scenario integrates the change to be simulated and evaluated for the ATM system of interest. As it provides a more insightful contribution to decision making, it has been classified as level 1B of the EASA AI taxonomy, as it essentially provides cognitive support to the human operator.

- **UC2 – AI-Powered Digital Assistant in TMA**

UC2 focuses on optimising the application of advanced continuous descent operations in the TMA through a Digital Assistant (DA) for spacing, scheduling and conflict detection and resolution (CDR). The expected safety benefits include better application of ICAO longitudinal/lateral separations, maximisation of runway capacity and optimisation of pilot and ATCO workload. In addition, the application of this concept could also contribute to minimising fuel consumption and environmental impact.

The main objective is to provide an AI-based DA to assist ATCOs in effectively managing inbound traffic and ensuring continuous descent operations. Given the nature of human-machine interaction, the analysis performed in D4.1 suggested classifying this solution as a level 2A of the EASA AI taxonomy.

- **UC3 – Dynamic Airspace Reconfiguration Service for U-Space**

The dynamic airspace reconfiguration (DAR) service involves modifying U-space volumes and exchanging information between ATM and U-space to temporarily create airspace boundaries. In controlled airspace, ANSPs remain responsible for providing air navigation services to manned aircraft operators. ANSPs also conduct dynamic reconfiguration of U-space airspace to ensure the safe segregation of manned and unmanned aircraft. In this context, air traffic control (ATC) units will temporarily limit areas within designated U-space airspace where unmanned aircraft system (UAS) operations can occur to accommodate short-term changes in manned traffic demand by adjusting the lateral and vertical limits of U-space airspace. They will also ensure timely and effective notification of relevant U-space service providers and single common information service providers (sCISPs) regarding the activation, deactivation, and temporary limitations of designated U-space airspace.

Supporting tools and AI applications will assist ATCOs in determining the best solutions and configurations for managing operations. These tools will process data from various sources (ATM, USSP) to provide optimal settings in terms of capacity, predictability, safety, efficiency, and environmental sustainability.

An AI could play the role of a DAR Manager, or at least, as a support, by leveraging its capabilities in data analysis, pattern recognition, predictive modelling, and decision-making.

Accordingly, the solutions here are classified at level 1B of the EASA taxonomy, as the output generated by the AI will ultimately be used as an input for decision making by the ATCO (at least for the moment).

- **UC4 – Dynamic Allocation of Traffic between ATCO and System**

ARGOS (Dynamic Allocation of Traffic between ATCO and System) is a solution based on deterministic algorithms for the improvement of upper airspace utilisation by means of dynamic allocation of traffic between the ATCO and ARGOS. Objectives are to dynamically support the ATCOs in managing the traffic in the sector, by means of issuing operational clearances to safely handle basic traffic situations and aid controllers in handling complex traffic situations. ARGOS has 3 modes of use, enabling corresponding different levels of automation. It can serve as a decision-making support tool, just providing the ATCO the best plan for the considered flights (L3). Alternatively, it can be delegated to manage a specific set of flights under the monitoring of the ATCO (L5). Eventually, it can be set to autonomously manage all the flights alerting the ATCO when monitoring is required (L8).

Considering the three modes of use of ARGOS, the characterisation of this system in terms of the EASA AI level was considered to be threefold. When used as a decision support tool, the solution corresponds to level 1A, while when used at L5 it can be classified at EASA level 2B. Finally, when flight management is fully delegated to the tool, it can reach level 3A.

The figure below provides an overview of the distribution of the UCs according to the taxonomy provided by EASA for the characterisation of concepts and solutions.

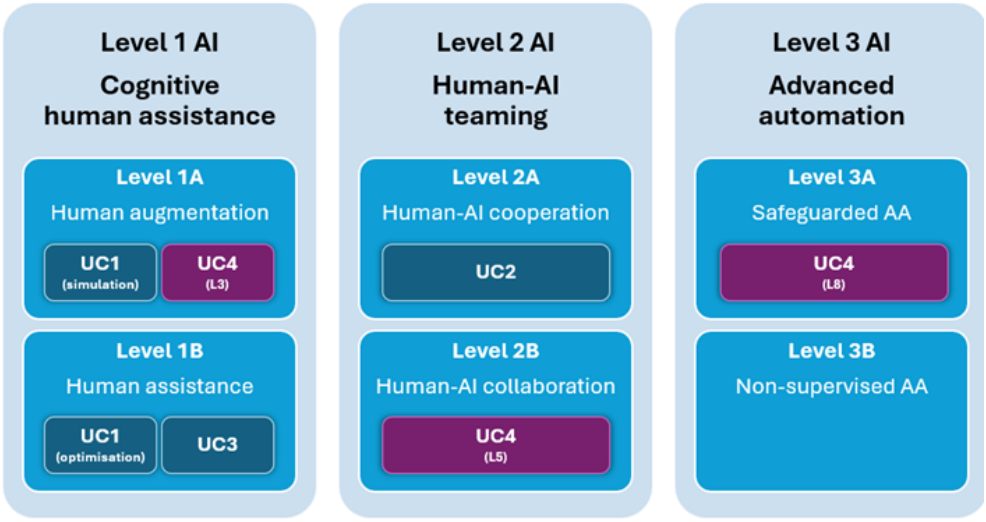

Figure 1: Overview of the distribution of the UCs according to the taxonomy provided by EASA for the characterisation of concepts and solutions

It is noteworthy that the four HUCAN use cases cover almost all levels of automation outlined in the EASA AI Roadmap, with the sole exception of level 3B ("non-supervised advanced automation"). However, it should be noted that UC1, UC2 and UC3 are based on AI solutions using machine learning (ML) (highlighted in blue). UC4, on the other hand, is based on deterministic algorithms (highlighted in purple).

2.5 Introduction to relevance and applicability assessment

As proposed by EASA, AI applications should comply with applicable requirements throughout their lifecycle. This implies that some requirements need to be considered from the early stages of the development process and be progressively met in the subsequent stages according to the maturity of the solution in question.

Against this background, HUCAN here proposes two complementary assessments to evaluate the relevance and applicability of the available objectives in the light of the characterisation and maturity of the UCs covered by the project. More specifically, the two analyses (relevance and applicability) run in parallel and aim to test the objectives defined by EASA for the different levels of AI. On the one hand, we assess the relevance of the EASA objectives in each scenario, taking into account the enabling technologies, the nature of the human-machine interactions and the expected impact on operational procedures. On the other hand, we examine the applicability of these objectives during the

development phase at the current TRL to identify issues that can and should be addressed to progressively align the solution with certification requirements.

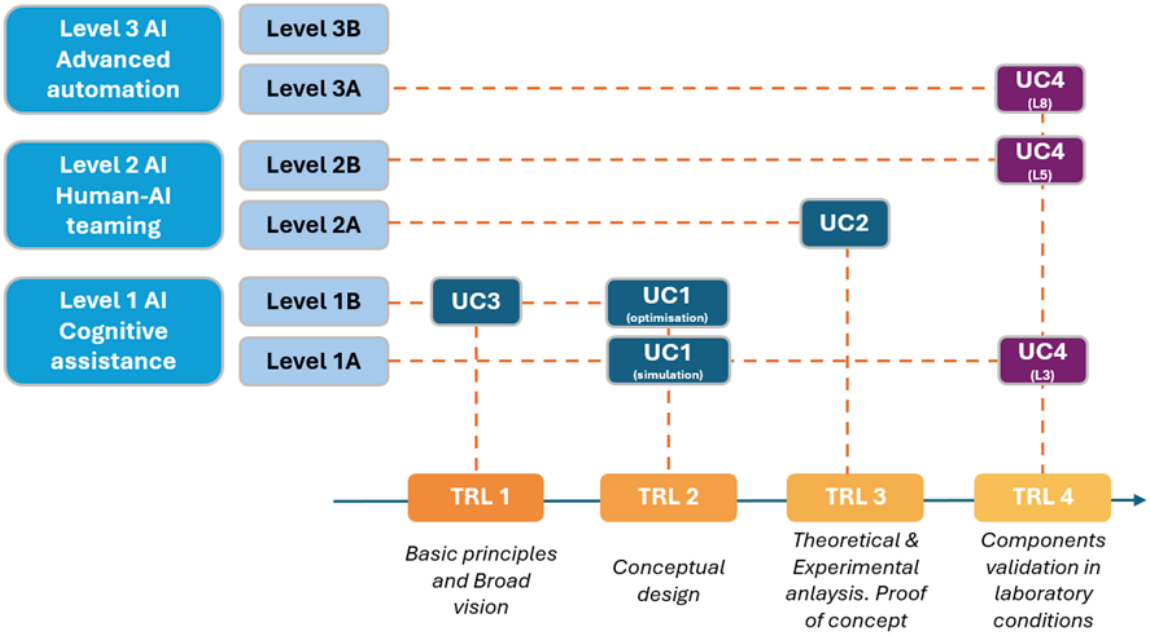

Figure 2: Overview of the distribution of the UCs according to the LoA and TRL

In light of these assumptions, HUCAN has performed the two assessments for each UC. The extended version of the results is available in Appendix B.

Considering the different technological enablers, HUCAN aims to assess the relevance of the objectives outlined by EASA not only for solutions explicitly based on ML, but also for approaches that aim to achieve high levels of automation through different enablers. Accordingly, the relevance assessment will be based on the centrality of the objectives for the purposes of the UCs, considering the level of automation achieved more than the enabler used to deliver it.

To assess the relevance of the objectives, we asked the UC owner to evaluate whether the objectives proposed by EASA were relevant to the realisation of the final solution, taking into account the concept defined so far and the expected level of automation. This evaluation covered the operational objectives, the available technical alternatives, the reasonably expected human factors (HF) implications - both in terms of human-machine interaction and wider organisational aspects. In addition, we suggested that ethical implications should be assessed in all cases, as these are an essential part of the trustworthiness analysis. Respondents were free to include objectives not explicitly related to the level of automation of their solution(s), if they felt they could be relevant for the purposes of trustworthiness.

Once this initial screening had been completed, we asked for an assessment of which objectives applicable to the solution could be practically considered at its current level of maturity, taking into account the stage of software development and any additional HF-related evaluations that had been carried out.

In the following paragraphs, the results of these assessments are presented in a table that organises the objectives using the same structure and order proposed by EASA. For each solution evaluated, the table highlights the number of applicable requirements based on the EASA guidelines for the level of automation considered. This is followed by a quantitative presentation of the relevance and applicability ratings. Qualitative feedback on specific findings from the assessments for each case is provided after the table.

Finally, the assessments carried out for each UC are followed by general conclusions on the emerging gaps identified through these case-based assessments in relation to the objectives and scope of a holistic certification approach as outlined at the beginning.

2.6 Results of the relevance and applicability assessment and discussion

2.6.1 UC1 – Dynamic configuration of airspace

As mentioned above, the relevance (R) and applicability (A) assessments for the UCs have been performed for each of the two solutions covered by the concept according to the respective levels of automation.

The table below illustrates the results obtained for the decision-making support tool based on simulation (TRL2, Level 1A). Columns Ref. and Subject refer to the section number and title in the EASA concept paper. Column 1A provides the number of objectives proposed by the EASA concept paper for level 1A. Columns R and A indicate how many of those objectives are considered relevant and applicable for UC1. The results are highlighted in red if a negative deviation from EASA's recommendations is recorded (fewer objectives than suggested) and in green if an addition is proposed (to include more objectives). For details see Appendix B.

Ref.	Subject	1A	A	R
C2.1	Characterization (CO/CL)	7	7	3
C2.2	Safety assessment (SA)	3	3	0
C2.3	Information and security (IS)	3	3	0
C2.4	Ethics-based assessment (ET)	N/A	**2**	2
C3.1	Learning assurance (DA, DM, LM, IMP, CM, QA, RU, SU)	56	**44**	9
C3.2	Development and post-ops AI explainability (EXP)	9	**5**	4
C4.1	AI operational explainability (EXP)	2	2	0
C4.2	Human-AI teaming (HF)	N/A	0	0
C4.3	Modality of interaction and style of interface (HF)	N/A	0	0
C4.4	Error management (HF)	N/A	0	0
C4.5	Failure management (HF)	N/A	**1**	1
C5	AI safety risk mitigation (SRM)	2	2	0
C6	Organisation (ORG)	8	8	4

Tot.	**90**	**74 (+3)**	**20 (+3)**

Table 2: UC1 - Decision-making support tool based on simulation. Relevance and applicability assessments

It is interesting to note how the relevance assessment highlighted that while some technical objectives are not of immediate utility for the solution in question, others, despite being designed for higher levels of automation, can contribute to the development of a holistic solution, particularly regarding the ethical impact of the solutions and failure management.

Regarding applicability, for this use case (which is at TRL2) it is observed that, from a practical perspective, only 20 of the 90 prescribed requirements can effectively be considered. These are primarily distributed among the preliminary trustworthiness assessment and the technical requirements for learning assurance. Objectives related to explainability, safety risk mitigation, and organisational aspects can also begin to be considered, albeit marginally. It is worth noting that in addition to the 90 objectives proposed by EASA, the UC owner suggested that 3 additional objectives be included, incorporating ethics and failure management considerations.

The table below illustrates the results obtained for the decision-making support tool based on optimisation (TRL2, Level 1B).

Ref.	Subject	1B	A	R
C2.1	Characterization (CO/CL)	7	7	7
C2.2	Safety assessment (SA)	3	3	3
C2.3	Information and security (IS)	3	3	0
C2.4	Ethics-based assessment (ET)	N/A	**3**	3
C3.1	Learning assurance (DA, DM, LM, IMP, CM, QA, RU, SU)	56	**30**	9
C3.2	Development and post-ops AI explainability (EXP)	9	9	4
C4.1	AI operational explainability (EXP)	10	**8**	1
C4.2	Human-AI teaming (HF)	N/A	**1**	1
C4.3	Modality of interaction and style of interface (HF)	N/A	0	0
C4.4	Error management (HF)	N/A	**4**	0
C4.5	Failure management (HF)	N/A	**2**	2
C5	AI safety risk mitigation (SRM)	2	2	0
C6	Organisation (ORG)	8	8	6
	Tot.	**90**	**98**	**70 (+10)**

Table 3: UC1 - Decision-making support tool based on optimisation. Relevance and applicability assessments

The relevance assessment shows that, although operating at a higher level of automation (Level 1B) compared to the first solution related to UC1 (Level 1A), only 30 of the 56 Learning assurance-related

objectives identified by EASA are considered relevant for the use case. Specifically, while broader general objectives may still apply, for the given scenario and technology, several objectives related to learning assurance, data management and learning process implementation are not considered relevant.

Although the gap is significantly smaller than for Learning assurance, it is worth noting that not all AI operational explainability goals appear to be relevant, particularly those related to the customisation of recommended abstraction levels for system use and the timing of explainability.

It is interesting to note that, from the UC owner's point of view, a few objectives should be included that should not be applicable to its concepts under the EASA directives. More specifically, three ethics-related objectives emerge as relevant: respect for privacy, environmental impact and well-being, and assessment of medium to long-term re-skilling and up-skilling needs (see Appendix B for details). The same applies to HF profiles related to human-AI teaming, in particular the system's ability to propose alternative solutions to those already proposed, and objectives related to error management.

The outcome is that, out of a suggested compliance framework comprising 98 objectives, 80 are deemed relevant by the UC owner, and at the current maturity level of the solution, 36 of these can already be addressed.

2.6.2 UC2 – AI-Powered Digital Assistant in TMA

The table below illustrates the results obtained for this solution, currently TRL3, Level 2A.

Ref.	Subject	2A	A	R
C2.1	Characterization (CO/CL)	7	**6**	4
C2.2	Safety assessment (SA)	3	3	0
C2.3	Information and security (IS)	3	3	0
C2.4	Ethics-based assessment (ET)	8	8	0
C3.1	Learning assurance (DA, DM, LM, IMP, CM, QA, RU, SU)	56	**40**	12
C3.2	Development and post-ops AI explainability (EXP)	9	9	4
C4.1	AI operational explainability (EXP)	10	10	0
C4.2	Human-AI teaming (HF)	5	**2**	2
C4.3	Modality of interaction and style of interface (HF)	6	**0**	0
C4.4	Error management (HF)	5	**0**	0
C4.5	Failure management (HF)	N/A	0	0
C5	AI safety risk mitigation (SRM)	2	**0**	0
C6	Organisation (ORG)	8	**0**	0
	Tot.	**124**	**81**	**22**

Table 4: UC2. Relevance and applicability assessments

Extending the analysis to an automation level that incorporates human-AI teaming, it is interesting to note that, given the current definition and consideration of the UC2 concept, only 81 out of the 124 objectives outlined by EASA are considered relevant. Furthermore, several objectives, particularly those related to modes of interaction, error and failure management, and safety risk mitigation, do not introduce substantial innovations compared to the status quo.

For the analysis carried out on HUCAN UCs, it is also noteworthy, in terms of applicability, that a higher maturity level (such as TRL3) does not immediately enable the consideration of more objectives, whether from a technical, programming and interaction, or organisational perspective.

2.6.3 UC3 – Dynamic Airspace Reconfiguration Service for U-Space

The table below illustrates the results obtained for this solution, currently TRL1, Level 1B.

Ref.	Subject	1B	A	R
C2.1	Characterization (CO/CL)	7	7	7
C2.2	Safety assessment (SA)	3	3	1
C2.3	Information and security (IS)	3	3	0
C2.4	Ethics-based assessment (ET)	N/A	0	0
C3.1	Learning assurance (DA, DM, LM, IMP, CM, QA, RU, SU)	56	**27**	0
C3.2	Development and post-ops AI explainability (EXP)	9	9	0
C4.1	AI operational explainability (EXP)	10	**6**	0
C4.2	Human-AI teaming (HF)	N/A	**5**	0
C4.3	Modality of interaction and style of interface (HF)	N/A	0	0
C4.4	Error management (HF)	N/A	0	0
C4.5	Failure management (HF)	N/A	0	
C5	AI safety risk mitigation (SRM)	2	2	0
C6	Organisation (ORG)	8	8	0
	Tot.	**98**	**80 (+5)**	**8**

Table 5: UC3. Relevance and applicability assessments

Looking at this concept, classified at level 1B, the relevance assessment provides two main results. On the one hand, looking at the objectives set by EASA for solutions aiming to achieve this level of automation, most of them are essential for development and certification purposes. For example, when looking at learning assurance, 27 of them are essential for the development of the solution . On the other hand, it is interesting to note that objectives intended for solutions that already involve some form of human-AI team could also be important for solutions with a lower level of automation.

For a concept at TRL1, the analysis of the applicability of the EASA objectives to the proposed solution provided limited results. As the data show, the level of definition of the intended use of the decision

support tool so far allows a stable classification according to the EASA taxonomy and a preliminary analysis of the safety risks that could be correlated to the introduction of this innovation. The assessment of more specific aspects, ranging from learning assurance to HF and organisational implications, is premature.

An important point to note is that this preliminary concept analysis, guided by EASA standards, has allowed for a more refined evaluation and definition of certain design choices. Grounding the envisaged technological solution in a concept has allowed a more informed exploration of the risks and benefits associated with alternatives proposing different levels of automation. In the light of the findings of this initial analysis, it has been possible to redefine the objectives, scope and intended operational deployment of the solution in a more confident and coherent manner.

2.6.4 UC4 – Dynamic Allocation of Traffic between ATCO and System

As previously mentioned, relevance (R) and applicability (A) assessments for the UC4 were conducted for each of the three solutions outlined in the concept, based on their respective levels of automation. Since UC4 relies on deterministic algorithms that the UC owner does not classify as AI models, the evaluation of the EASA objectives differed slightly from previous assessments. In light of these characteristics, the focus has been primarily on the BBs of trustworthiness analysis, HF for AI, and minimization of safety risks. Learning assurance has not been considered.

The table below illustrates the results obtained for the decision-making support tool (TRL4, Level 1A).

Ref.	Subject	1A	A	R
C2.1	Characterization (CO/CL)	7	7	7
C2.2	Safety assessment (SA)	3	3	1
C2.3	Information and security (IS)	3	3	1
C2.4	Ethics-based assessment (ET)	N/A	**6**	0
C3.1	Learning assurance (DA, DM, LM, IMP, CM, QA, RU, SU)	56	N/A	0
C3.2	Development and post-ops AI explainability (EXP)	9	9	0
C4.1	AI operational explainability (EXP)	2	**10**	0
C4.2	Human-AI teaming (HF)	N/A	**5**	0
C4.3	Modality of interaction and style of interface (HF)	N/A	**2**	0
C4.4	Error management (HF)	N/A	**2**	0
C4.5	Failure management (HF)	N/A	**4**	0
C5	AI safety risk mitigation (SRM)	2	2	0
C6	Organisation (ORG)	8	8	0
	Tot.	**90**	**34 (+27)**	**9**

Table 6: UC4 - L3 Decision-making support tool. Relevance and applicability assessments

As mentioned above, it is important to note that for the relevance assessment, objectives related to AI assurance were not considered. Therefore, with regard to the L3 function, which corresponds to an AI level 1A, only 34 of the 90 objectives specified by EASA were evaluated.

In this context, it is immediately apparent that the analysis of the UC includes several objectives beyond those prescribed, which could contribute to strengthening the overall trustworthiness of the concept. As the results show, this applies not only to the ethical evaluation, but also to many HF dimensions, such as AI operational explainability (EXP); Human-AI teaming (HF); Modality of interaction and interface style (HF); Error management (HF); Failure management (HF).

In terms of applicability, the objectives of immediate utility are those related to the trustworthiness analysis, while others cannot yet be adequately considered, even at TRL4.

The table below illustrates the results obtained when ARGOS is delegated to manage specific flights under the monitoring of the ATCO (TRL4, Level 2B).

Ref.	Subject	2B	A	R
C2.1	Characterization (CO/CL)	7	7	7
C2.2	Safety assessment (SA)	3	3	1
C2.3	Information and security (IS)	3	3	1
C2.4	Ethics-based assessment (ET)	8	**6**	0
C3.1	Learning assurance (DA, DM, LM, IMP, CM, QA, RU, SU)	56	N/A	0
C3.2	Development and post-ops AI explainability (EXP)	9	9	0
C4.1	AI operational explainability (EXP)	10	10	0
C4.2	Human-AI teaming (HF)	11	**10**	0
C4.3	Modality of interaction and style of interface (HF)	16	**2**	0
C4.4	Error management (HF)	5	**4**	0
C4.5	Failure management (HF)	4	4	0
C5	AI safety risk mitigation (SRM)	2	2	0
C6	Organisation (ORG)	8	8	0
	Tot.	**142**	**68**	**9**

Table 7: UC4 - L5 Delegation of management of specific flights under the monitoring of the ATCO. Relevance and applicability assessments

In line with what was mentioned earlier, for the L5 function, which is classified as level 2B in the EASA taxonomy, the number of objectives considered for the relevance assessment is 86, as the 142 initially outlined by EASA exclude those related to AI assurance. In the light of these considerations, it is noteworthy that 68 objectives are found to be relevant. The negative deviations concern, in particular, certain aspects of ethics (in particular with regard to the protection of personal data) and the modality of interaction and style of interface (HF). Again, in terms of applicability, the objectives related to the trustworthiness analysis are the ones that can be adequately addressed.

The table below illustrates the results obtained when ARGOS is delegated to manage flights requesting ATCO monitoring only in case of necessity (TRL4, Level 3A). Considering that the EASA AI Roadmap and the Concept Papers still did not specifically address this level of AI, the reference objectives are those provided for level 2B.

Ref.	Subject	2B	A	R
C2.1	Characterization (CO/CL)	7	7	7
C2.2	Safety assessment (SA)	3	3	1
C2.3	Information and security (IS)	3	3	1
C2.4	Ethics-based assessment (ET)	8	**6**	0
C3.1	Learning assurance (DA, DM, LM, IMP, CM, QA, RU, SU)	56	N/A	0
C3.2	Development and post-ops AI explainability (EXP)	9	9	0
C4.1	AI operational explainability (EXP)	10	10	0
C4.2	Human-AI teaming (HF)	11	11	0
C4.3	Modality of interaction and style of interface (HF)	16	**2**	0
C4.4	Error management (HF)	5	**4**	0
C4.5	Failure management (HF)	4	4	1
C5	AI safety risk mitigation (SRM)	2	2	0
C6	Organisation (ORG)	8	8	0
	Tot.	**142**	**69**	**10**

Table 8: UC4 - L8 Delegation of management flights requesting ATCO monitoring only in case of necessity. Relevance and applicability assessments

There are clear analogies between the results of the L8 and L5 evaluations, although the two functions correspond to different levels of automation (2B and 3A in the EASA taxonomy). In terms of relevance, 69 of the 86 objectives considered have been taken into account. In terms of applicability, 10 objectives are considered addressable. In addition to the reliability analysis, the concept design has also considered and addressed failure management aspects.

2.7 Conclusions

The overall results of these evaluations show that not all learning assurance objectives are critical in terms of relevance. Specifically, the detailed data show that not all objectives related to data management, learning process management and development, and AI post-operational explainability are relevant to the scenarios considered. Objectives related to AI/ML model reuse and surrogate modelling are also marginal in the context of the project.

As a general observation, it appears that the integration of certain aspects related to ethical evaluation and human factors is important, in particular with regard to error and failure management and some human-AI teaming profiles. On the other hand, objectives related to interaction modality and interface

style do not seem to be crucial for the applications considered. This suggests that there is room and interest for further exploration in these areas.

In terms of applicability, it is not surprising that for concepts at a relatively low TRL, the objectives that can realistically be considered during development vary. What is more curious is that solutions having similar TRLs sometimes are ready to take on board different objectives with a considerable variance in terms of distribution (technical and non-technical objectives). These data also suggest an interest in exploring whether and how the alignment process to objectives can be standardized across different maturity levels of a concept, to facilitate the overall standardization of the design and development process towards certification.

3 PBRs and KPIs for advanced automation

The aim of this chapter is to develop performance indicators for the objectives outlined in the EASA Concept Paper. This is done in several steps. The chapter first (section 3.1) explains what we mean by performance-based requirements (PBRs) and key performance indicators (KPIs), and why we would need them. Next, it provides (section 3.2) criteria for good performance indicators. Subsequently (section 3.3), it identifies aspects of liability and human factors to be taken into account for the development of PBRs and KPIs, since these are particularly relevant in the context of advanced automation. Next, section 3.4 shows that PBRs and KPIs for advanced automation need to be developed following a holistic approach. Finally, section 3.5 uses these results to develop KPI for the EASA objectives.

3.1 What are PBRs and KPIs and why do we need them?

HUCAN's holistic framework not only addresses the key issues and areas critical to the certification of AI and advanced automation in aviation, but also considers the most appropriate approaches and requirements to address these areas throughout the development process of these solutions. The adaptive nature of advanced automated systems, including AI-based solutions, poses significant challenges to traditional certification regulations, which rely on the premise that a system's correct behaviour must be fully specified and verified prior to operation, and that its response remains invariant in the long run. In response, regulatory initiatives are emerging to extend these certification frameworks by incorporating more flexible training and testing methods, which may foster increased trust in the reliable performance of these systems over time (HUCAN, 2024b).

Many of these initiatives substantially combine traditional prescriptive requirements by a performance-based approach to certification and operational safety. This scheme shifts the focus from strict compliance with predefined rules to a more flexible method for assessing the achievement of specific, measurable outcomes. Unlike prescriptive standards, which dictate exactly what must be done and how, PBRs cover a set of objectives, targets, and indicators to assess whether a system or service meets the desired performance outcomes (section 1.2 in ICAO, 2018).

This performance-based approach encourages innovation, as it allows for practical, iterative evaluation of whether a solution meets safety expectations in real-world scenarios. The key to implementing this approach effectively lies in correctly identifying appropriate thresholds to assess safety performance. This includes defining relevant KPIs that align with overarching safety objectives and other objectives and establishing data-driven parameters for ongoing monitoring and evaluation (sections 8.3.5.2 - 8.3.5.13 in ICAO, 2018).

Accordingly, for advanced automation to be safely integrated into air traffic operations, PBRs and KPIs must be defined and monitored to ensure operational effectiveness and maintain high safety standards. PBRs outline how automated systems should function and perform under varying conditions, while KPIs are used to measure the safety and effectiveness of these systems. This approach emphasises the need for desired, measurable outcomes rather than strict compliance with prescriptive rules in the certification process. KPIs provide ongoing assessment of automated systems' safety and effectiveness, across normal, impaired, and recovery operational phases.

In this regard, EASA in its Roadmap already opens consultations on the new requirements that should inform the certification of AI-based solutions. Parallel to the activities carried out by the EUROCAE WG-114 on Artificial Intelligence, the Agency in the Concept Papers #1 and #2 progressively defined new anticipated Means of Compliance (Anticipated MOCs) to orient ongoing AI-based research projects with emerging certification expectations (EASA, 2024b). As emphasised in safety management guidelines, fostering a positive safety culture remains critical, with safety being the foremost priority in certification processes. Given the challenges posed by the certification of advanced automation and AI, it is essential to undertake a comprehensive and wide-ranging assessment of the scope that performance targets should encompass, the attributes that requirements must have to effectively evaluate performance, and the indicators to be used as benchmarks. In fact, while pressing issues underscore the need for a holistic approach to certification, key considerations—particularly regarding Human Factors and liability—require careful examination of the methodologies and parameters to be employed for certification purposes.

KPIs for advanced automation aim at evaluating the effectiveness, efficiency, safety, and overall impact of automated systems on air traffic operations. The literature distinguishes between leading indicators (or process indicators) and lagging indicators (also content or outcome indicators). Leading indicators are oriented towards a process or activity and usually refer to positive things that an organisation wants to strengthen, such as activities that improve safety. Lagging indicators are oriented towards an outcome of (for example) a scenario and usually refer to negative things an organisation wants to avoid, such as accidents. A set of indicators should include both leading and lagging indicators. Lagging indicators can often more directly describe the effectiveness of a risk management measure, but they are less sensitive to changes in the very safe aviation system. Leading indicators can be used to proactively evaluate the impact of risk management measures before serious incidents or accidents occur.

3.2 Criteria for good KPIs

Criteria for good KPIs are:

Sensitive. The indicator should be sufficiently sensitive to variations in what is to be measured. If an indicator is not sensitive to changing conditions then it will not be able to provide information, e.g., on the trend of a risk or on the effectiveness of risk management measures. The sensitivity of an indicator is thus an important prerequisite for being able to control the value it provides. This includes having the right level of detail, having a range that is sufficient to measure all variations, and having the opportunity to measure frequently enough in order to capture those variations.

Rational. An indicator is rational if it measures what it is intended to measure. Rational means that the relationship between the indicator and that which is to be measured is based on (empirical) evidence or logical reasoning. There is a clear rationale of how the indicator connects to the objective.

Unambiguous. An indicator is unambiguous if it is clear from the indicator's description what is being measured. Unambiguous means that only one interpretation is possible, so that there can be no confusion about it and no other things are measured than are needed to achieve the objective. The interpretation of the indicator should not depend on culture, knowledge or experience.

Measurable. An indicator is measurable if it can be expressed in a measure and unit. The measurability of an indicator also depends on the availability of data. An indicator is measurable if supporting data is available or can be made available. This also enables statistical analysis.

Reproducible. An indicator is reproducible if there is minimal variability if measured under the same conditions. An indicator may in principle be measurable, but that does not mean that the underlying data is accurately reliable and reproducible. An indicator is accurate and reliable if the measurement results relate to the actual value. This includes the quality of the data and the reliability of the measuring instruments. It also requires capabilities to determine accuracy: the margin of uncertainty is known and understood.

Acceptable. The application of a particular indicator can only be successful if it is accepted by those directly involved. An indicator is acceptable if the persons or organisations that have to work with it (or that have to deal with consequences of applying the indicator, in particular the aviation sector and citizens) accept the description and application. The identification and formulation of the indicator should take into account the capabilities and limitations of the organisations involved.

Manipulation-proof. Manipulation is the deliberate act of making the indicator value look different without changing the underlying factors. An indicator that is susceptible to manipulation may give a value that is not an accurate reflection of reality. As a result, the (manipulated) value of the indicator loses its validity. It is important that the possibility of manipulation of an indicator's values should be avoided as much as possible. Because many indicators can be manipulated to some extent, this criterion is also about how easy that is and how strong the control mechanisms are. Manipulation resistance is also dependent on the process by which the indicator value is determined.

Time-valid. The meaning and validity of an indicator should not change over time. For a variety of reasons the validity of an indicator may expire. It is therefore important that the validity of an indicator is periodically checked.

Cost-efficient. An indicator is cost-efficient if the costs of applying the indicator are not disproportionate to the benefits of using the indicator. Costs include the collection of the data, and the time and resources required to apply the indicator. Benefits include the quality of the data obtained, often compared to the quality of data from other indicators. Cost-efficiency is also an important consideration for the acceptance of an indicator.

Simple. The system within which an indicator is designed can be complex. Indicators are meant to provide insight despite this complexity. Indicators should therefore be simple. Simple means that the indicator should be understandable without much specialist knowledge. The documentation of the process by which the indicator was created should also be clear.

Manageable. This criterion is not about individual indicators, but about the complete set. The set of indicators must remain manageable. They should be meaningful (i.e. not just easy to measure) and simple, and there should not be too many of them.

3.3 Human factors and liability

This section identifies aspects of liability and human factors to be taken into account for the development of PBRs and KPIs. In particular, the analysis first identifies and defines relevant

stakeholders as the bearers of responsibility and accountable entities, including their interrelationships. For instance, **operators** (such as air traffic controllers, pilots, and ground handling staff), **system developers** (such as software engineers and AI system designers), and **maintenance organisations** (including technical support teams and hardware maintenance providers) each play a critical role in supporting system functionality and safety.

These stakeholders' responsibilities are interconnected and influenced by various levels of automation, each presenting specific challenges and limitations that affect liability and human factors considerations. Continuing, the influence of different levels of automation is considered, stressing the challenges and limitations emerging at different levels, and how they influence liability and human factors. Finally, mitigation strategies are additionally included, to support the research of solutions in an effort to tackle the complications highlighted with this analysis, and directly supporting the drafting of PBRs and KPIs.

To begin with, a liability analysis within ATM systems involves a complex interplay between operators, systems developers, and maintenance organisations. For the purposes of definitions of the stakeholders involved (as **organisations**, **entities** and\or **individuals**) please refer to the roles and responsibility mentioned in Chapter 5 of HUCAN D4.1 (HUCAN, 2024c).

In this regard, when identifying relevant stakeholders, regulatory authorities have been considered out of scope. This evaluation depends on the consideration that among EU Member States and at international level the liabilities of public entities and administrations – such as regulatory authorities involved in the sector - is based on assumptions and requires proof of very different conditions (which may vary in each jurisdiction) than those to be considered in identifying liabilities of private entities and individuals.

Each stakeholder's actions and decisions contribute to the overall safety and reliability of the system, and understanding these responsibilities is essential for mitigating risks and addressing potential liabilities. In the context of advanced automation, the issues of responsibility, accountability, and liability take on a complex dimension, requiring an in-depth analysis of the level of automation, human factors and their interactions with automated systems. This section examines how such dynamics manifest at different levels of automation, how they impact the liabilities of the subject mentioned above (some of them or all) and how they can inform the development of Performance-Based Requirements (PBRs) and Key Performance Indicators (KPIs).

Interdependence between with PBRs and KPIs

Implications for Performance-Based Requirements

Considerations of responsibility and liability must inform the development of PBRs. These requirements should consider not only the technical capabilities of the system but also how operators interact with it. For example, a PBR could require including regular training sessions for operators to ensure they maintain the necessary skills to intervene effectively when required.

Key Performance Indicators

KPIs must reflect not only the effectiveness of the automated system but also the interaction between the system and human operators. It is crucial to include indicators that assess operator fatigue,

recovery capacity, and working conditions to ensure that automated systems not only function correctly but that operators can make safe and informed decisions.

Considerations of Human Factors

The analysis of human factors must be at the centre of all considerations of responsibility and liability. Elements such as fatigue, stress, and operator training can significantly influence performance and decision-making. For example, in high-automation scenarios, dependence on automation can lead to decreased operator attention, increasing the risk of errors. Therefore, it is essential that PBRs and KPIs include measures to continuously monitor and improve operator well-being, ensuring they can maintain control and responsibility even in complex contexts.

Complex interrelationships exist between responsibility, liability, and human factors in scenarios of advanced automation. The considerations raised must guide the formulation of PBRs and KPIs, ensuring not only that automated systems are safe and reliable, but also that human operators can act responsibly and competently.

Liability and Human Factor Analysis

Appendix C gives the details of the Liability and Human Factor Analysis for eight cases:

i. Loss of system control
ii. Human-computer interaction
iii. Lack of information or data misinterpretation
iv. Regulatory violation or non-compliance with certification standards
v. Over-reliance on automation
vi. Human oversight
vii. Inadequate training and skill gaps
viii. Difficulties in the allocation of responsibility or unclear responsibility during automation failures

The described cases highlight various potential risks and liability scenarios that can arise in the use of automated ATM systems, based on the level of automation of the systems. Each category emphasises the importance of clear communication, proper training, and ongoing vigilance among all stakeholders involved in the operation, development, and maintenance of these systems.

However, to fully connect and understand the dynamics described in the above cases, in order to develop coherent Performance-Based Requirements (PBR) and Key Performance Indicators (KPI) it should be considered that:

The increasing level of automation may affect the roles and responsibilities of the stakeholders involved. For instance, in lower levels of automation, operators might be more directly accountable, but as automation increases, the focus might shift toward the accountability of system developers and maintenance organisations (details in Appendix C). PBRs can be developed that account for the gradual reduction in human involvement and the increased reliance on automated systems.

Human factors, such as reduced oversight or misinterpretation of automated data, can directly inform the creation of KPIs. For example, safety indicators could monitor operator alertness and engagement

in semi-automated systems or evaluate the frequency and causes of human error in interpreting automated data.

The identified liabilities and cases of risks should be tied to operational scenarios directly and on PBR development. For example, scenarios like human oversight failures or data misinterpretation, should influence PBRs by requiring systems to have built-in fail-safes, clearer data interfaces, or better operator training programs to prevent those specific types of errors.

Similarly, KPI should measure not only system reliability but also the human-system interaction performance. For example, KPIs might track how frequently operators successfully intervene during system failures, how well operators understand data from the system, or how maintenance organisations respond to system issues.

The exercise conducted in this section will link each case and hypothesis to specific PBRs and KPIs in section 3.4, to properly consider the automation-related liabilities and risks and integrate them into the PBRs and KPIs developed.

3.4 Holistic KPIs for advanced automation

This section uses the results of the previous sections to introduce KPIs for advanced automation. These KPIs address the performance of the automated system in a holistic scope for operations by the overall sociotechnical system. They aim to focus on maintaining operational excellence while integrating ethical standards, accountability, human oversight, uncertainty, safety, public oversight, sustainability, and data governance. This is the broad-scope, holistic view that was recommended following the review in D3.2 (HUCAN, 2024b). In this section, implications for KPIs are discussed for the recommended topics of the holistic certification approach. Next, Section 3.5 provides an overview of specific KPIs for the EASA objectives discussed in Chapter 2.

Uncertainty. A robust certification approach should account for the inherent uncertainties in various key aspects, including the technology itself, the data used, operational scenarios, environmental factors, and unforeseen behaviour in the context of autonomy and automation. This evaluation goes beyond assessing if the approach considers basic uncertainties and component failures and is particularly critical when considering the highest levels of automation and the relationship of all of the above with accountability. It also assesses how the certification approach facilitates the development of contingency plans for unforeseen events, major failures, or security breaches. Examples of KPIs for this topic include:

- *Adopted assumptions.* Measuring uncertainty starts with identifying and maintaining the list of assumptions adopted, related to the design and evaluation of the system.
- *Identified varying conditions.* Disturbances and types of performance variability that can influence operations of the sociotechnical system have been identified and assessed.
- *Entropy and information gain.* Measures the amount of uncertainty or information in a probability distribution or a data set.

Safety. Evaluate the effectiveness of the certification approach in supporting comprehensive risk control strategies. Posing the focus on safety management should facilitate robust feedback mechanisms to learn from operational occurrences involving advanced automation, as well as tackle technological safety tools in the strict sense. This includes identifying suitable indicators that

effectively capture potential risks and dangerous autonomous or automated behaviour. The certification approach should support integrated risk management practices, encompassing not only safety but also security-related interfaces for key performance areas like environmental, service-oriented and organisational security. This evaluation should consider the level of detail provided by the safety risk assessment, including the types of qualitative or quantitative results generated and the means of compliance included.

Associated categories are:

- Redundancy: Any incident attributed to the failure of an AI-based system is a safety concern, requiring investigation and possible corrective action. There should be a clear mechanism for determining responsibility, whether it is human or machine-based. To guarantee that the operation continues safely, even in the event of a system failure, the system must include redundant subsystems to ensure fail-safe operations and to prevent single points of failure. Human controllers should be able to override the system in case of malfunctions. Liabilities related to the described case of loss of control (section 3.3) can be mitigated by redundant subsystems ensuring seamless transitions to backup systems in case of failure.
- Reliability and availability: To ensure continuous operation during all phases, to ensure minimal disruptions in the management of air traffic, and to minimise the risk of system failures that could lead to safety incidents, the automated systems should be reliable and should be available without interruptions. Downtime or system failures can severely disrupt operations, leading to delays or compromised safety.
- Robustness and resilience: To ensure safety and efficiency across diverse flight and traffic situations and in diverse environmental and non-standard conditions (e.g., extreme weather, turbulence, dense traffic), the highly automated sociotechnical system should function effectively in diverse operating conditions and handle disturbances efficiently.
- Security and cyber resilience: With increased automation comes the risk of cyberattacks and system vulnerabilities. Certification processes must ensure that automation systems are robust against cyber threats by assessing measures to prevent unauthorised access or control of the systems, ensuring that data processed by the system is accurate and secure, and evaluating how well the system can resist attempts to disable or manipulate it (hacking, malicious interference).

Examples of KPIs for this topic include:

- *System uptime.* Measures the percentage of time that automated systems are available and functioning without interruption.
- *Automation failure rate*. Tracks the number of failures by automated systems and the severity of their consequences.
- *Human-automation interaction failures*. Measures failures in coordination between human operators and automated systems that may lead to hazardous situations.
- *Adaptability to changing airspace conditions.* Measures how flexible the system is in adapting to different airspace structures, weather conditions, or emergency scenarios.
- *Cybersecurity incident rate.* Measures the number of cybersecurity incidents (e.g., breaches, intrusions) affecting AI-driven systems. As AI systems become more integrated into critical infrastructure, robust cybersecurity measures are essential to prevent system compromise.

- *Number of cybersecurity breaches.* Tracks the number of successful cyberattacks or breaches that compromise AI-based systems.
- *System vulnerability mitigation.* Measures how well the system proactively identifies and patches vulnerabilities, ensuring data protection and privacy.

Accountability. This topic concerns the effectiveness of the certification approach's accountability framework. A robust approach should clearly define a framework that assigns clear responsibilities and obligations to stakeholders throughout the civil aviation value chain. This framework should be designed to incentivize adoption of certification measures and ensure ongoing compliance with established safety and security standards. The evaluation should assess the level of discretion granted to stakeholders in implementing the framework. It's crucial to strike a balance between flexibility and ensuring a consistent level of safety across the industry. Furthermore, the evaluation should identify the primary entities held accountable for adherence to the framework and explore how accountability is distributed across the value chain. A well-defined approach will explicitly delineate accountability for different stakeholders involved in the design, development, operation, and maintenance of aviation systems.

Associated categories are:

- Compliance and regulatory standards: To ensure that the system adheres to industry-wide safety, security, and operational standards, as well as ethical guidelines, and to ensure that automated systems are certified for aviation use according to established safety protocols, they must comply with national and international aviation regulatory standards. Compliance ensures the system can operate legally and safely across different jurisdictions and airspaces. The adoption of AI-based systems must respect global and local regulatory frameworks, and there must be accountability mechanisms for detecting and reporting non-compliance, including for ethical violations. Inadequate training and skills may also determine a breach of regulatory standards. Training should ensure that operators are qualified to handle automation and manual overrides in line with regulatory standards.
- Just culture: Just culture rests on three pillars of justice—substantive (fair and legitimate rules), procedural (unbiased and transparent processes), and restorative (repairing relationships after incidents). Together, these principles promote trust, transparency, and open reporting, all essential for cultivating a robust safety culture. Research (Dekker & Breakey (2016); Cromie & Bott (2016); Kirwan (2024)), shows that contextual understanding influences disciplinary actions, with more lenient responses observed as additional information is provided. Emphasising restorative justice encourages organisations to focus on learning over punishment, fostering collaboration and improvement while reducing incident recurrence by addressing systemic causes rather than assigning individual blame. These insights help shape KPIs tailored to AI and automation in aviation, including trust and reporting metrics, which track operator confidence in AI systems, assess human-AI collaboration, and evaluate incident reporting frequency and quality. In particular, the contribution of just culture sees the drafting of compliance indicators that measure procedural fairness and operator participation in rule formation, including metrics for learning and accountability that assess post-incident analyses and restorative actions, ensuring AI and automation drive continuous improvement rather than punitive responses.

Examples of KPIs for this topic include:

- *List of entities held accountable*. Lists the primary entities held accountable for adherence to the framework and explores how accountability is distributed across the value chain.
- *Regulatory compliance rate*. Measures the percentage of automated systems that meet international safety and operational standards (e.g., ICAO, EASA, FAA regulations).
- *Audit and inspection pass rate*. Assesses how often the system successfully passes regulatory audits and inspections related to automation technology.
- *Just culture assessment.* Measures the level of just culture in an organisation.
- *Just incident investigation practices.* Ensure automation systems enable fair reporting and corrective actions aligned with just culture principles (so that incidents are addressed holistically, allowing operators to report issues free of fear of unfair consequences).
- *Definition of collaborative accountability.* Check whether collaborative accountability approaches have been implemented, which define the (shared) responsibility for safety of stakeholders.

Environmental protection. Assesses the certification approach's capacity to support the reduction of air travel's environmental footprint. An effective approach should address key environmental concerns associated with air travel, including mitigating climate change through CO2 emission reduction strategies, minimising aircraft noise pollution, and safeguarding local air quality around airports. International organisations establish environmental standards that member states translate into national regulations. This evaluation focuses on how effectively the certification approach fosters the adoption, consideration, or implementation of these established environmental standards.

Examples of KPIs for this topic include:

- *Fuel consumption reduction*. Measures the reduction in fuel use due to more efficient flight paths and fewer delays, directly impacting emissions.
- *CO2 emissions reduction*. Evaluates how automation contributes to lowering aviation's carbon footprint by optimising flight routes and minimising idle times.
- *Noise pollution reduction*. Measures any decrease in noise levels around airports due to optimised approach and departure procedures enabled by automation.
- *AI ecosystem impact.* Assesses the environmental impact of the ML energy use and system hardware development for AI-based systems.

Public oversight. Measures the extent of democratic control over the organisations, procedures, and enforcement mechanisms associated with the certification approach. It acknowledges the inherent tension between delegating certification activities and duties to private entities or non-traditional public bodies (across member states) and the need for effective public oversight. The evaluation considers concepts like "thirdness" (independence from industry or government) and potential biases within the oversight structure. Furthermore, it assesses the level of public participation in the certification process and transparency surrounding the certified products (technologies, systems etc.). A well-designed approach should ensure that public interest is served through robust oversight mechanisms and opportunities for public engagement.

Examples of KPIs for this topic include:

- *Public participation*. Measures the level of public participation in the certification process.

Efficiency. This criterion evaluates the overall efficiency of the certification process facilitated by the approach. This includes assessing the expected total completion time for technology certification. A well-designed approach should strike a balance between fostering innovation and establishing clear regulatory frameworks. It should ensure a level of rigour necessary to maintain safety without unnecessarily hindering the pace of technological advancement and production.

The term efficiency also relates to the aviation system itself. To ensure that automated systems contribute to a sustainable aviation future, they must be designed such that the overall operation is energy efficient, without sacrificing performance, to ensure long-term environmental sustainability and cost-effectiveness in ATM. Automation aims to contribute to the overall efficiency of air traffic management, improving performance in areas like Capacity (whether automation can help increase the capacity of the airspace and airports to handle more flights safely), Flight efficiency (optimising flight routes, altitudes, and speeds to reduce fuel consumption, delays, and environmental impacts), and Real-time data processing (the ability of the automation system to process vast amounts of real-time data from aircraft, weather systems, and ground infrastructure to optimise decision-making).

Examples of KPIs for this topic include:

- *Air traffic throughput.* Measures the volume of aircraft that can be safely managed by the automated system in a given period, often compared to pre-automation benchmarks.
- *Reduced delays.* Tracks the reduction in average departure, en route, and arrival delays due to automated systems improving flow management.
- *Capacity utilisation.* Monitors how well the airspace and airport capacity are used with the help of automated systems, maximising throughput while maintaining safety.
- *Maintenance and upgrade costs of automation systems*. Evaluates the ongoing costs required to maintain and upgrade automated systems versus their expected benefits.
- *Return on investment (ROI)*. Measures the overall economic return from deploying advanced automation systems in terms of cost savings, reduced delays, and improved efficiency.

Technical complexity. Evaluates the level of knowledge and experience necessary to understand the certification approach, utilise it correctly, and interpret its results. This includes the explainability of the approach, ensuring transparency and clarity in its application. Additionally, the evaluation considers the complexity of tools required to utilise the approach. An ideal approach would be accessible to a reasonable range of experts within the field, utilising tools that are efficient and do not necessitate excessive computational resources. Flexibility of the approach for applying it to new technologies, emerging air traffic management concepts (e.g., drones, urban air mobility), or changes in regulatory requirements is also an asset.

Examples of KPIs for this topic include:

- *Technical explainability.* How well can the functioning of the AI-based system be explained and analysed by technical experts?
- *V&V flexibility.* How well can the approaches for validation and verification be applied to a broad variety of technologies and operational concepts?

Human Factors. This criterion evaluates how effectively the certification approach considers human factors in interaction with advanced automation. A well-designed approach should account for the various ways humans will interact with the technology, encompassing considerations like human

oversight and human-AI teaming strategies. The evaluation should assess how well the approach facilitates the development of comprehensive training programs for personnel. These programs should equip personnel with the necessary skills to effectively collaborate with advanced automation, while fostering a strong safety culture. This includes promoting practices that discourage overreliance on the system, encourage the reporting of issues, and emphasise situational awareness.

Associated categories are:

- Human performance and workload: Advanced automation should enhance human performance, not hinder it. The certification process evaluates the integration of automation with human operators, focusing on Workload management (ensuring automation reduces, not increases, the workload for pilots and controllers), Situational awareness (ensuring that automation provides useful, timely, and clear information to human operators), Trust in automation (Balancing trust so operators neither overly rely on nor distrust automation, particularly in high-stress situations), and Training requirements (evaluating the extent of training needed for operators to effectively work with the automation). Operators must have access to well-structured and clear data about flight environments, with warnings of potential misinterpretation prominently highlighted in system interfaces. This can also be extended to the cases of liabilities possibly arising from inadequate human oversight: systems must keep operators informed of all relevant flight data, including emergencies.
- Human-Machine Interfaces: To ensure that operators can easily understand and interact with the system, reducing the risk of human error especially during high-traffic situations, the interface must be intuitive, user-friendly, and allow operators to monitor and control their flight(s) effectively without increasing cognitive workload and fatigue. The system must be capable of seamless interaction with the operators, providing clear, explainable decisions and alerts. Trust and collaboration between automated systems and human operators are crucial, especially during complex or emergency scenarios. With regard to liability related to human-system interaction (see section 3.3) the systems must feature intuitive and easy-to-navigate interfaces, with visual clarity and functionality that do not overload operators. The system must offer clear decision-making aids, with explanations for each automated recommendation and they shall be designed to respond to human inputs or commands within a predefined response time, and any delays must be logged and reported for performance evaluation.

Examples of KPIs for this topic include:

- *Controller workload.* Assesses how automation impacts the workload of air traffic controllers (e.g., through task delegation like conflict resolution or trajectory management).
- *Operator fatigue.* Measures the mental and the physical fatigue of operators in advanced automation operations.
- *Recovery capacity.* Measures the ability of operators to recover from failures, mistakes, and other threats.
- *Situation awareness.* Measures how well human operators can maintain an accurate understanding of the flight and traffic situations when interacting with automated tools.
- *Training and adaptation time.* Evaluates the amount of time required for operators to learn and adapt to new automated systems.
- *Operator alertness.* Measures operator alertness and engagement in semi-automated systems.

- *Misinterpretation errors.* Measures the frequency and causes of misinterpretations by operators of automated systems.
- *Human central agency.* Check whether AI supports rather than supplants human decision-making.

Data governance. Assesses the certification approach's capacity to establish robust data governance practices. Effective data governance ensures the accuracy, safety, usability, and accessibility of data used within advanced automation systems for civil aviation. This encompasses defining clear protocols for data access control, specifying who can access what data under specific conditions. The approach should also address data storage and usage practices, ensuring data integrity and adherence to relevant regulations.

Examples of KPIs for this topic include:

- *Data quality score.* A score or rating that represents the accuracy, consistency, timeliness, completeness, and reliability of the data.

3.5 Development of KPI for each of the EASA Objectives

The aim of this section is to develop KPIs for each of the Objectives proposed by EASA in section C of the Concept Paper (EASA, 2024b). This reference notes that these objectives are to be considered a first set, that aim to anticipate future EASA guidance and/or requirements to be complied with by safety-related ML applications. They apply to any AI-based system (defined by EASA as a system incorporating one or more ML models), and are intended for use in safety-related applications or for applications related to environmental protection covered by the Basic Regulation, in particular for the following domains:

- Initial and continuing airworthiness, applying to systems or equipment required for type certification or by operating rules, or whose improper functioning would reduce safety (systems or equipment contributing to failure conditions Catastrophic, Hazardous, Major or Minor);
- Air operations, applying to systems, equipment or functions intended to support, complement, or replace tasks performed by aircrew or other operations personnel (examples may be information acquisition, information analysis, decision-making, action implementation and monitoring of outputs);
- ATM/ANS, applying to equipment intended to support, complement or replace end-user tasks (examples may be information acquisition, information analysis, decision-making and action implementation) delivering ATS or non-ATS;
- Maintenance, applying to systems supporting scheduling and performance of tasks intended to timely detect or prevent unsafe conditions (airworthiness limitation section (ALS) inspections, certification maintenance requirements (CMRs), safety category tasks) or tasks which could create unsafe conditions if improperly performed (‘critical maintenance tasks’);
- Training, applying to systems used for monitoring the training efficiency or for supporting the organisational management system, in terms of both compliance and safety;
- Aerodromes, applying to systems that automate key aspects of aerodrome operational services, such as the identification of foreign object debris, the monitoring of bird activities, and the detection of UAS around/at the aerodrome;

- Environmental protection, applying to systems or equipment affecting the environmental characteristics of products.

As such, in this report, the Objectives are interpreted as PBRs for AI-based systems.

Using the material in sections 3.1-3.4 as guideline, for each of the Objectives, the report provides one or more KPIs, i.e. indicators that can be used by the applicant to measure whether the Objective has been satisfied. For each KPI also one or more Milestones have been identified that can be seen as targets or sub-targets for the KPI towards satisfying the Objective. The results are presented in Appendix D.

Some example KPIs are provided below. Please refer to Appendix D for details.

Ref.	Subject	Example KPI
C2.1	Characterization (CO/CL)	For each end user, the list of goals that are intended to be performed in interaction with the AI-based system. Document that describes how end users' inputs have been collected and accounted for in the development of the AI-based system.
C2.2	Safety assessment (SA)	Identification of data that needs to be recorded for the purpose of supporting the continuous safety assessment.
C2.3	Information and security (IS)	List of information security risks with an impact on safety. The effectiveness of the security controls introduced to mitigate the identified AI/ML-specific information security risks to an acceptable level.
C2.4	Ethics-based assessment (ET)	Assessment of the creation or reinforcement of unfair bias in the AI-based system, regarding both the data sets and the trained models, including an assessment of impact of the unfair bias on performance and safety.
C3.1	Learning assurance (DA, DM, LM, IMP, CM, QA, RU, SU)	Capturisation of the requirements on data to be pre-processed and engineered for the inference model in development and for the operations. Definition and documentation of pre-processing operations on the collected data in preparation of the model training. Assessment of the bias-variance trade-off in the model family selection.
C3.2	Development and post-ops AI explainability (EXP)	Identification of the methods at AI/ML item and/or output level satisfying the specified AI explainability needs.

C4.1	AI operational explainability (EXP)	Definition of relevant explainability regarding the appropriateness of the decision / action as expected.
C4.2	Human-AI teaming (HF)	Assessment of the ability of the AI-based system design to propose alternative solutions and support its positions.
C4.3	Modality of interaction and style of interface (HF)	An assessment of the ability to combine or adapt the interaction modalities depending on the characteristics of the task, the operational event and/or the operational environment.
C4.4	Error management (HF)	An assessment of the likelihood of design-related end-user errors in the design of the AI-based system.
C4.5	Failure management (HF)	An assessment of the ability to diagnose the failure and present the pertinent information to the end user.
C5	AI safety risk mitigation (SRM)	Assessment of the need for an additional dedicated layer of protection to mitigate the residual risks to an acceptable level.
C6	Organisation (ORG)	Auditability of the safety-related AI-based systems.

Table 9. EASA AI Objectives. Example KPIs

It is stressed that:

- The Objectives identified in (EASA, 2024b) are not finalised and are still being developed further by EASA. For example, EASA states that the applicability of their guidelines is limited as follows:
    - covering Level 1 and Level 2 AI applications, but not covering yet Level 3 AI applications;
    - covering supervised learning or unsupervised learning, but not other types of learning such as reinforcement learning;
    - covering offline learning processes where the model is 'frozen' at the time of approval, but not online learning processes.
- The KPI and Milestones identified In Appendix D are a first set and should be further developed as well.

Despite this disclaimer, the HUCAN team believes the set of KPI and Milestones can be used in support of the further development of a holistic approach for the certification of advanced automation, including AI-based systems, as anticipated in the remainder of the HUCAN project.

4 Concluding remarks and recommendations

In support of the holistic certification approach for AI-based systems and advanced automation that is developed in the HUCAN project, this report identified requirements and performance indicators in association with the developing guidance for AI-based systems by EASA (EASA, 2024b). In line with the holistic views of the ethics guidelines for trustworthy AI (High-level Expert Group on AI, 2019) and the AI Act of the European Union (Regulation 2024/1689), EASA's developing guidance material encapsulates a broad perspective on key performance areas that should be addressed by objectives and means of compliance for advanced automation and AI-based systems. In particular, depending on the level of automation, up to 142 objectives were identified for a wide range of areas, encompassing safety, security, ethics, explainability, human-AI teaming, AI assurance, and organisational aspects. The range of objectives will be extended in future EASA concept papers, as the scope has now mostly been on supervised machine learning, yet excluding AI techniques such as reinforcement learning, logic- and knowledge-based approaches, and hybrid AI, which can all support advanced automation in aviation and ATM.

In this report a systematic analysis was made of the ways that the objectives of (EASA, 2024b) may be addressed by the use cases of the HUCAN project (HUCAN, 2024c). This was done by analysing the relevance of each objective for the technology, human-machine interactions and operational context of each use case. It was found that in the range of 65% to 86% of the objectives are relevant for the use cases. Not relevant objectives are most prominent in the learning assurance topic. Interestingly, there are also objectives that are defined out of the scope for particular levels of automation in (EASA, 2024b), but that are considered relevant for use cases, such as particular objectives for ethics and human factors. Furthermore, the applicability of each relevant objective was assessed for the technology readiness level of each use case. Here it was found that in the range of 6% to 37% of the overall sets of objectives are applicable for the TRL of the use case. These more limited percentages indicate that other technological and operational examples may need to be considered in the evaluation of the methods towards certification in HUCAN.

The integration of advanced automation in complex sociotechnical systems and the possibly adaptive performance of AI-based systems imply a need to shift from prescriptive requirements to performance-based approaches for certification and safety management. Quality criteria for performance indicators and a holistic overview of KPIs for advanced automation were presented in this report. These provide a basis for the selection of suitable KPIs in the HUCAN use cases, as well as for the KPIs in the methods that will be developed in HUCAN D4.4. Furthermore, a detailed list of initial KPIs and associated suitability criteria (milestones) were defined for all objectives of (EASA, 2024b). These KPIs can be further detailed for applicable objectives in the use cases during the validation study in HUCAN D4.3.

In conclusion, the HUCAN holistic certification approach for AI-based systems and advanced automation will address multiple key performance areas (KPAs) in cycles for design, development and evaluation of advanced automation and AI-based systems for a range of levels of automation. Herein, the maturity of the advanced automation concepts and supporting technology are increasing and their readiness levels are evaluated for a holistic scope of KPAs. In coordination with stakeholders, requirements addressing the various KPAs are updated and detailed as the designers, developers, evaluators and other stakeholders achieve better understanding of the performance of the overall

system and the impact on the KPAs. As such, it will support the development and certification of trustworthy advanced automation and supporting AI technology in support of approval by certifying authorities at the highest TRLs. Figure 3 gives an illustrative sketch of the HUCAN holistic approach, which will be further developed in HUCAN D4.4.

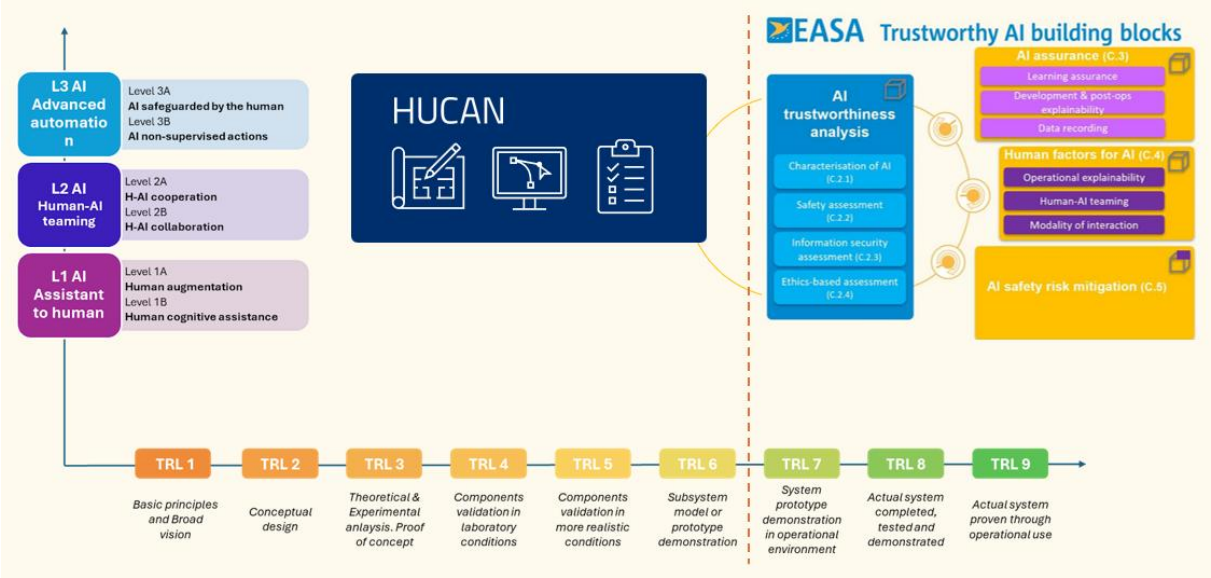

Figure 3: Illustrative sketch of HUCAN holistic approach for increasing TRLs in automation concepts with various LOAs in preparation of approval by certifying authorities

5 References

Cromie, Sam, and Franziska Bott (2016). "Just culture's "line in the sand" is a shifting one; an empirical investigation of culpability determination." Safety science 86 (2016): 258-272.

Dekker, Sidney WA, and Hugh Breakey (2016). "'Just culture:' Improving safety by achieving substantive, procedural and restorative justice." Safety science 85 (2016): 187-193.

EASA (2023). ARTIFICIAL INTELLIGENCE ROADMAP 2.0 Human-centric approach to AI in aviation, May 2023

EASA (2024a). Machine Learning Application Approval (MLEAP) final report, authored by MLEAP consortium for EASA, May 2024

EASA (2024b). EASA Artificial Intelligence Concept Paper Issue 2. Guidance for Level 1 & 2 machine-learning applications, April 2024

EASA (2024c). EASA AI Days. Presentation. Day 1 - July 2nd, 2024. Slide no. 32

EU COM 2018/237. Communication from the Commission: Artificial Intelligence in Europe.

High-level Expert Group on AI (2019). *Ethics guidelines for trustworthy AI.* European Commission, 8 April 2019

HUCAN (2024a). D3.1 - Certification methods and automation benefits issues and challenges, 29 February 2024.

HUCAN (2024b). D3.2 - Innovative approaches to approval and certification, 31 August 2024.

HUCAN (2024c). D4.1 - Case studies introduction: level of automation analysis and certification issues, 28 August 2024.

ICAO (2018) Doc 9859, Safety Management Manual.

Kirwan, Barry (2024). "The Impact of Artificial Intelligence on Future Aviation Safety Culture." Future Transportation 4.2 (2024): 349-379.

Regulation (EU) 2018/1139 of the European Parliament and of the Council of 4 July 2018 on common rules in the field of civil aviation and establishing a European Union Aviation Safety Agency, and amending Regulations (EC) No 2111/2005, (EC) No 1008/2008, (EU) No 996/2010, (EU) No 376/2014 and Directives 2014/30/EU of the European Parliament and of the Council, and repealing Regulations (EC) No 552/2004 and (EC) No 216/2008 of the European Parliament and of the Council and Council Regulation (EEC) No 3922/91

Regulation (EU) 2024/1689 of the European Parliament and of the council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)

SESAR JU (2024). European ATM Master Plan Stakeholder consultation workshop pre-read material, 22-23 April 2024. https://sesarju.eu/sites/default/files/documents/events/ATM%20MP%20workshop%20preread%20material_2024.04.08_FINAL.pdf

6 List of acronyms

Acronym	Description
AI	Artificial Intelligence
ANSP	Air Navigation Service Providers
ARGOS	Dynamic Allocation of Traffic between ATCO and System
ATC	Air Traffic Control
ATCO	Air Traffic Controller
ATM	Air Traffic Management
CDR	Conflict Detection and Resolution
CISP	Common Information Service Providers
CL	Collaboration
CM	Configuration Management
CO	Cooperation
ConOps	Concept of Operations
COTS	Commercial-Off-The-Shelf
CSCW	Computer Supported Cooperative Work
DA	Development Assurance
DA	Digital Assistant
DAR	Dynamic Airspace Reconfiguration
DM	Data Management
DQR	Data Quality Requirement
EASA	European Union Aviation Safety Agency
ET	Ethics
EU	European Union
EXP	Explainability
GDPR	General Data Protection Regulation
HAIRM	human-AI resource management
HF	Human Factors
HUCAN	Holistic Unified Certification Approach for Novel systems based on advanced automation
ICAO	International Civil Aviation Organization
IMP	Implementation
IS	Information Security

KPA	Key Performance Area
KPI	Key Performance Indicator
LM	Learning process Management
ML	Machine Learning
MOC	Means of Compliance
MUAC	Maastricht Upper Area Control
N/A	Not Applicable
OD	Operational Domain
ODD	Operational Design Domain
OoD	Out of Distribution
ORG	Organisation
PBR	Performance-Based Requirement
QA	Quality Assurance
R&D	Research & Development
ROI	Return on Investment
RU	Reuse of AI/ML models
SA	Safety Assessment
SESAR	Single European Sky ATM Research
SPI	Safety Performance Indicator
SRM	Safety Risk Mitigation
SU	Surrogate Modelling
TRL	Target Readiness Level
UAS	Unmanned Aircraft System
UC	Use Case
U-space	Unmanned airspace
USSP	U-Space Service Provider

Appendix A: Objectives EASA Concept Paper

This appendix feeds into sections 2.2, 2.6, Appendix B and Appendix D, and lists all Objectives from Section C (AI Trustworthiness guidelines) in the EASA Concept Paper with guidance for level 1&2 ML applications (EASA, 2024b), together with the Levels of Automation (LoA) for which the objectives are applicable, and the Means of Compliance anticipated by EASA. Here, Cx.y refers to the subsection in (EASA, 2024b).

Objectives in White are relevant for all Levels of Automation (1A-2B), Objectives in Green are relevant for 1B-2B, Objectives in Yellow are relevant for 2A-2B, and Objectives in Blue are relevant for 2B only.

C2. Trustworthiness analysis		
LoA	**EASA Objectives**	**Anticipated MOC**
C2.1(CO/CL). Characterisation and classification of the AI application		
1A - 2B	Obj.CO-01: The applicant should identify the list of end users that are intended to interact with the AI-based system, together with their roles, their responsibilities (including indication of the level of teaming with the AI-based system, i.e. none, cooperation, collaboration) and expected expertise (including assumptions made on the level of training, qualification and skills).	-
1A - 2B	Obj.CO-02: For each end user, the applicant should identify which goals and associated high-level tasks are intended to be performed in interaction with the AI-based system.	Ant. MOC CO-02
1A - 2B	Obj.CO-03: The applicant should determine the AI-based system taking into account domain-specific definitions of 'system'.	Ant.MOC CO-03
1A - 2B	Obj.CO-04: The applicant should define and document the ConOps for the AI-based system, including the task allocation pattern between the end user(s) and the AI-based system. A focus should be put on the definition of the OD and on the capture of specific operational limitations and assumptions.	Ant.MOC CO-04* *Dependencies: Obj.CO-01 Obj.CO-02
1A - 2B	Obj.CO-05: The applicant should document how end users' inputs are collected and accounted for in the development of the AI-based system.	Ant.MOC CO-05
1A - 2B	Obj.CO-06: The applicant should perform a functional analysis of the system, as well as a functional decomposition and allocation down to the lowest level.	Ant.MOC CO-06
1A - 2B	Obj.CL-01: The applicant should classify the AI-based system, based on the levels presented [by EASA], with adequate justifications.	Ant.MOC CL-01-1* *Dependencies: Obj.CO-02

		Ant.MOC CL-01-2
C2.2(SA). Safety assessment of ML Applications		
1A - 2B	Obj.SA-01: The applicant should perform a safety (support) assessment for all AI-based (sub)systems, identifying and addressing specificities introduced by AI/ML usage.	Ant.MOC-SA-01-1 Ant.MOC-SA-01-2 Ant.MOC-SA-01-3 Ant.MOC-SA-01-4 Ant.MOC-SA-01-5* *Dependencies: Objs.LA Ant.MOC-SA-01-6 Ant.MOC-SA-01-7 Ant.MOC-SA-01-7 Ant.MOC-SA-01-8 Ant.MOC-SA-01-9
1A - 2B	Obj.SA-02: The applicant should identify which data needs to be recorded for the purpose of supporting the continuous safety assessment.	Ant.MOC-SA-02 *Dependencies: Ant.MOC EXP-04-2
1A - 2B	Obj.SA-03: In preparation of the continuous safety assessment, the applicant should define metrics, target values, thresholds and evaluation periods to guarantee that design assumptions hold.	Ant.MOC SA-03
C2.3(IS). Information security risks management		
1A - 2B	Obj.IS-01: For each AI-based (sub)system and its data sets, the applicant should identify those information security risks with an impact on safety, identifying and addressing specific threats introduced by AI/ML usage.	Ant.MOC IS-01
1A - 2B	Obj.IS-02: The applicant should document a mitigation approach to address the identified AI/ML-specific information security risk.	Ant.MOC IS-02
1A - 2B	Obj. IS-03: The applicant should validate and verify the effectiveness of the security controls introduced to mitigate the identified AI/ML-specific information security risks to an acceptable level.	Ant.MOC IS-03
C2.4(ET). Ethics-based assessment		
2A – 2B	Obj.ET-01: The applicant should perform an ethics-based trustworthiness assessment for any AI-based system developed using ML techniques or incorporating ML models.	-

2A – 2B	Obj.ET-02: The applicant should ensure that the AI-based system bears no risk of creating overreliance, attachment, stimulating addictive behaviour, or manipulating the end user's behaviour.	Ant.MOC ET-02 *Dependencies: Obj.ET-01 Obj.IMP-09
2A – 2B	Obj.ET-03: The applicant should comply with national and EU data protection regulations (e.g. GDPR), i.e. involve their Data Protection Officer, consult with their National Data Protection Authority, etc.	Ant.MOC ET-03
2A – 2B	Obj.ET-04: The applicant should ensure that the creation or reinforcement of unfair bias in the AI-based system, regarding both the data sets and the trained models, is avoided, as far as such unfair bias could have a negative impact on performance and safety.	Ant.MOC ET-04
2A – 2B	Obj.ET-05: The applicant should ensure that end users are made aware of the fact that they interact with an AI-based system, and, if applicable, whether some personal data is recorded by the system.	Ant.MOC ET-05
2A – 2B	Obj.ET-06: The applicant should perform an environmental impact analysis, identifying and assessing potential negative impacts of the AI-based system on the environment and human health throughout its life cycle (development, deployment, use, end of life), and define measures to reduce or mitigate these impacts.	Ant.MOC ET-06 *Dependencies: Obj.ET-01
2A – 2B	Obj.ET-07: The applicant should identify the need for new skills for users and end users to interact with and operate the AI-based system, and mitigate possible training gaps	Ant.MOC ET-07 *Dependencies: Obj.ET-01 Prov.ORG-07 Prov.ORG-08
2A – 2B	Obj.ET-08: The applicant should perform an assessment of the risk of de-skilling of the users and end users and mitigate the identified risk through a training needs analysis and a consequent training activity	Ant.MOC ET-08 *Dependencies: Obj.ET-01 Prov.ORG-07 Prov.ORG-08
C3. AI Assurance		
LoA	**EASA Objectives**	**Anticipated MOC**
C3.1(DA). Learning assurance		
1A-2B	Obj.DA-01: The applicant should describe the proposed learning assurance process, taking into account each of the steps described in Sections C.3.1.2 to C.3.1.14, as well as the	Ant.MOC DA-01

	interface and compatibility with development assurance processes.	
1A-2B	Obj.DA-02: Based on (sub)system requirements allocated to the AI/ML constituent, the applicant should capture the following minimum for the AI/ML constituent requirements: — safety requirements allocated to the AI/ML constituent (e.g. performance, reliability, resilience); — information security requirements allocated to the AI/ML constituent; — functional requirements allocated to the AI/ML constituent; — operational requirements allocated to the AI/ML constituent, including AI/ML constituent ODD monitoring and performance monitoring (to support related objectives in Section C.3.2.6), detection of OoD input data and data-recording requirements (to support objectives in Section C.3.2.7); — other non-functional requirements allocated to the AI/ML constituent (e.g. scalability); and — interface requirements.	*Dependencies: Obj.CO-04
1A-2B	Obj.DA-03: The applicant should define the set of parameters pertaining to the AI/ML constituent ODD, and trace them to the corresponding parameters pertaining to the OD when applicable.	Ant.MOC DA-03 *Dependencies: Obj.CO-04
1A-2B	Obj.DA-04: The applicant should capture the DQRs for all data required for training, testing, and verification of the AI/ML constituent, including but not limited to: — the data relevance to support the intended use; — the ability to determine the origin of the data; — the requirements related to the annotation process; — the format, accuracy and resolution of the data; — the traceability of the data from their origin to their final operation through the whole pipeline of operations; — the mechanisms ensuring that the data will not be corrupted while stored, processed, or transmitted over a communication network; — the completeness and representativeness of the data sets; and — the level of independence between the training, validation and test data sets.	Ant.MOC DA-04

1A-2B	Obj.DA-05: The applicant should capture the requirements on data to be pre-processed and engineered for the inference model in development and for the operations.	-
1A-2B	Obj.DA-06: The applicant should describe a preliminary AI/ML constituent architecture, to serve as reference for related safety (support) assessment and learning assurance objectives.	-
1A-2B	Obj.DA-07: The applicant should validate each of the requirements captured under Objectives DA-02, DA-03, DA-04, DA-05 and the architecture captured under Objective DA-06.	Ant.MOC DA-07 *Dependencies: Obj.DA-02 Obj.DA-03 Obj.DA-04 Obj.DA-05 Obj.DA-06
1A-2B	Obj.DA-08: The applicant should document evidence that all derived requirements generated through the learning assurance processes have been provided to the (sub)system processes, including the safety (support) assessment.	*Dependencies: Obj. DA-03 Obj. DA-04 Obj. DA-05 Obj. LM-01 Obj. LM-02 Obj. LM-04 Obj. IMP-01
1A-2B	Obj.DA-09: The applicant should document evidence of the validation of the derived requirements, and of the determination of any impact on the safety (support) assessment and (sub)system requirements.	*Dependencies: Obj. DA-03 Obj. DA-04 Obj. DA-05 Obj. LM-01 Obj. LM-02 Obj. LM-04 Obj. IMP-01
1A-2B	Obj.DA-10: Each of the captured AI/ML constituent requirements should be verified.	
C3.1(DM). Data management		
1A-2B	Obj.DM-01: The applicant should identify data sources and collect data in accordance with the defined ODD, while ensuring satisfaction of the defined DQRs, in order to drive the selection of the training, validation and test data sets.	
1A-2B	Obj.DM-02-SL: Once data sources are collected and labelled, the applicant should ensure that the annotated or labelled data	*Dependencies: Obj.DA-04

	in the data set satisfies the DQRs captured under Objective DA-04.	
1A-2B	Obj.DM-03: The applicant should define the data preparation operations to properly address the captured requirements (including DQRs).	
1A-2B	Obj.DM-04: The applicant should define and document pre-processing operations on the collected data in preparation of the model training.	Ant.MOC DM-04
1A-2B	Obj.DM-05: When applicable, the applicant should define and document the transformations to the pre-processed data from the specified input space into features which are effective for the performance of the selected learning algorithm.	Ant.MOC DM-05-1 Ant.MOC DM-05-2 Ant.MOC DM-05-3
1A-2B	Obj.DM-06: The applicant should distribute the data into three separate data sets which meet the specified DQRs in terms of independence (as per Objective DA-04): — the training data set and validation data set, used during the model training; — the test data set used during the learning process verification, and the inference model verification.	*Dependencies: Obj.DA-04 Obj.DA-07
1A-2B	Obj.DM-02-UL:Once data sources are collected and the test data set labelled, the applicant should ensure that the annotated or labelled data in this test data set satisfies the DQRs captured under Objective DA-04.	*Dependencies: Obj.DA-04
1A-2B	Obj.DM-07: The applicant should ensure verification of the data, as appropriate, throughout the data management process so that the data management requirements (including the DQRs) are addressed.	Ant.MOC DM-07-1 Ant.MOC DM-07-2 Ant.MOC DM-07-3 Ant.MOC DM-07-4 Ant.MOC DM-07-5
1A-2B	Obj.DM-08: The applicant should perform a data verification step to confirm the appropriateness of the defined ODD and of the data sets used for the training, validation and verification of the ML model.	Ant.MOC DM-08 *Dependencies: Obj.EXP-02 Obj.EXP-03
C3.1(LM). Learning process management		
1A-2B	Obj.LM-01: The applicant should describe the ML model architecture.	Ant.MOC LM-01
1A-2B	Obj.LM-02: The applicant should capture the requirements pertaining to the learning management and training processes, including but not limited to: — model family and model selection;	Ant.MOC LM-02

	— learning algorithm(s) selection; — explainability capabilities of the selected model; — activation functions; — cost/loss function selection describing the link to the performance metrics; — model bias and variance metrics and acceptable levels (only in supervised learning); — model robustness and stability metrics and acceptable levels; — training environment (hardware and software) identification; — model parameters initialisation strategy; — hyper-parameters and parameters identification and setting; — expected performance with training, validation and test data sets.	
1A-2B	Obj.LM-03: The applicant should document the credit sought from the training environment and qualify the environment accordingly.	
1A-2B	Obj.LM-04: The applicant should provide quantifiable generalisation bounds.	Ant.MOC LM-04
1A-2B	Obj.LM-05: The applicant should document the result of the model training.	Ant.MOC LM-05 *Dependencies: Obj.SA-01
1A-2B	Obj.LM-06: The applicant should document any model optimisation that may affect the model behaviour (e.g. pruning, quantisation) and assess their impact on the model behaviour or performance.	Ant.MOC LM-06
1A-2B	Obj.LM-07-SL: The applicant should account for the bias-variance trade-off in the model family selection and should provide evidence of the reproducibility of the model training process.	Ant.MOC LM-07-SL
1A-2B	Obj.LM-08: The applicant should ensure that the estimated bias and variance of the selected model meet the associated learning process management requirements.	Ant.MOC LM-08 *Dependencies: Obj.DM-02-UL
1A-2B	Obj.LM-09: The applicant should perform an evaluation of the performance of the trained model based on the test data set and document the result of the model verification.	Ant.MOC LM-09 *Dependencies: Obj.SA-01 Obj.LM-04

1A-2B	Obj.LM-10: The applicant should perform requirements-based verification of the trained model behaviour.	Ant.MOC LM-10 *Dependencies: Obj.LM-02 Obj.DA-02
1A-2B	Obj.LM-11: The applicant should provide an analysis on the stability of the learning algorithms.	Ant.MOC LM-11
1A-2B	Obj.LM-12: The applicant should perform and document the verification of the stability of the trained model, covering the whole AI/ML constituent ODD.	Ant.MOC LM-12
1A-2B	Obj.LM-13: The applicant should perform and document the verification of the robustness of the trained model in adverse conditions.	Ant.MOC LM-13
1A-2B	Obj.LM-14: The applicant should verify the anticipated generalisation bounds using the test data set.	Ant.MOC LM-14 *Dependencies: Obj.LM-04
1A-2B	Obj.LM-15: The applicant should capture the description of the resulting ML model.	
1A-2B	Obj.LM-16: The applicant should confirm that the trained model verification activities are complete.	Ant.MOC LM-16
C3.1(IMP). Model implementation		
1A-2B	Obj.IMP-01: The applicant should capture the requirements pertaining to the ML model implementation process.	Ant.MOC IMP-01
1A-2B	Obj.IMP-02: The applicant should validate the model description captured under Objective LM-15 as well as each of the requirements captured under Objective IMP-01.	*Dependencies: Obj.LM-15 Obj.IMP-01
1A-2B	Obj.IMP-03: The applicant should document evidence that all derived requirements generated through the model implementation process have been provided to the (sub)system processes, including the safety (support) assessment.	
1A-2B	Obj.IMP-04: Any post-training model transformation (conversion, optimisation) should be identified and validated for its impact on the model behaviour and performance, and the environment (i.e. software tools and hardware) necessary to perform model transformation should be identified.	Ant.MOC IMP-04-1 Ant.MOC IMP-04-2 *Dependencies: Obj.LM-06 Obj.IMP-01
1A-2B	Obj.IMP-05: The applicant should plan and execute appropriate development assurance processes to develop the inference model into software and/or hardware items.	Ant.MOC IMP-05

1A-2B	Obj.IMP-06: The applicant should verify that any transformation (conversion, optimisation, inference model development) performed during the trained model implementation step has not adversely altered the defined model properties.	Ant.MOC IMP-06 *Dependencies: Obj.IMP-01
1A-2B	Obj.IMP-07: The differences between the software and hardware of the platform used for model training and those used for the inference model verification should be identified and assessed for their possible impact on the inference model behaviour and performance.	Ant.MOC IMP-07
1A-2B	Obj.IMP-08: The applicant should perform an evaluation of the performance of the inference model based on the test data set and document the result of the model verification.	Ant.MOC IMP-08 *Dependencies: Obj.SA-01 Obj.LM-09
1A-2B	Obj.IMP-09: The applicant should perform and document the verification of the stability of the inference model.	Ant.MOC IMP-09
1A-2B	Obj.IMP-10: The applicant should perform and document the verification of the robustness of the inference model in adverse conditions.	Ant.MOC IMP-10
1A-2B	Obj.IMP-11: The applicant should perform requirements-based verification of the inference model behaviour when integrated into the AI/ML constituent.	Ant.MOC IMP-11 *Dependencies: Obj.IMP-01 Obj.DA-02 Obj.DM-02-UL
1A-2B	Obj.IMP-12: The applicant should confirm that the AI/ML constituent verification activities are complete.	Ant.MOC IMP-12
C3.1(CM). Configuration management		
1A-2B	Obj.CM-01: The applicant should apply all configuration management principles to the AI/ML constituent life-cycle data, including but not limited to: — identification of configuration items; — versioning; — baselining; — change control; — reproducibility; — problem reporting; — archiving and retrieval, and retention period.	Ant.MOC CM-01
C3.1(QA). Quality and process assurance		

1A-2B	Obj.QA-01: The applicant should ensure that quality/process assurance principles are applied to the development of the AI-based system, with the required independence level.	
C3.1(RU). Reuse of AI/ML models		
1A-2B	Obj.RU-01: The applicant should perform an impact assessment of the reuse of a trained ML model before incorporating the model into an AI/ML constituent. The impact assessment should consider: — alignment and compatibility of the intended behaviours of the ML models; — alignment and compatibility of the ODDs; — compatibility of the performance of the reused ML model with the performance requirements expected for the new application; — availability of adequate technical documentation (e.g. equivalent documentation depending on the required assurance level); — possible licensing or legal restrictions on the reused ML model (more particularly in the case of COTS ML models); and — evaluation of the required development level.	Ant.MOC RU-01 *Dependencies: Obj.DA-01
1A-2B	Obj.RU-02: The applicant should perform a functional analysis of the COTS ML model to confirm its adequacy to the requirements and architecture of the AI/ML constituent.	*Dependencies: Obj.DA-02
1A-2B	Obj.RU-03: The applicant should perform an analysis of the unused functions of the COTS ML model, and prepare the deactivation of these unused functions.	*Dependencies: Obj.DA-03 Obj.DA-04 Obj.DA-05 Obj.DA-10 Obj.DM-01 Obj.DM-05 Obj.DM-06 Obj.DM-07 Obj.LM-01 Obj.LM-02 Obj.LM-03 Obj.LM-08 Obj.LM-09 Obj.LM-10 Obj.LM-11

		Obj.LM-12 Obj.LM-15 Obj.IMP-01 Obj.IMP-05 Obj.IMP-06 Obj.IMP-11 Obj.CM-01 Obj.QA-01 Obj.EXP-03
C3.1(SU). Surrogate modelling		
1A-2B	Obj.SU-01: The applicant should capture the accuracy and fidelity of the reference model in order to support the verification of the accuracy of the surrogate model.	
1A-2B	Obj.SU-02: the applicant should identify, document and mitigate the additional sources of uncertainties linked with the use of a surrogate model.	
C3.2(EXP). Development and post-ops AI explainability		
1A-2B	Obj.EXP-01: The applicant should identify the list of stakeholders, other than end users, that need explainability of the AI-based system at any stage of its life cycle, together with their roles, their responsibilities and their expected expertise (including assumptions made on the level of training, qualification and skills).	*Dependencies: Obj.CO-01
1A-2B	Obj.EXP-02: For each of these stakeholders (or groups of stakeholders), the applicant should characterise the need for explainability to be provided, which is necessary to support the development and learning assurance processes.	Ant.MOC EXP-02
1A-2B	Obj.EXP-03: The applicant should identify and document the methods at AI/ML item and/or output level satisfying the specified AI explainability needs.	
1A-2B	Obj.EXP-04: The applicant should design the AI-based system with the ability to deliver an indication of the level of confidence in the AI/ML constituent output, based on actual measurements or on quantification of the level of uncertainty.	
1A-2B	Obj.EXP-05: The applicant should design the AI-based system with the ability to monitor that its inputs are within the specified ODD boundaries (both in terms of input parameter range and distribution) in which the AI/ML constituent performance is guaranteed.	

1A-2B	Obj.EXP-06: The applicant should design the AI-based system with the ability to monitor that its outputs are within the specified operational AI/ML constituent performance boundaries.	
1A-2B	Obj.EXP-07: The applicant should design the AI-based system with the ability to monitor that the AI/ML constituent outputs (per Objective EXP-04) are within the specified operational level of confidence.	Ant.MOC EXP-07 *Dependencies: Obj.EXP-04
1A-2B	Obj.EXP-08: The applicant should ensure that the output of the specified monitoring per the previous three objectives are in the list of data to be recorded per MOC EXP-09-2.	*Dependencies: Ant.MOC EXP-09-2
1A-2B	Obj.EXP-09: The applicant should provide the means to record operational data that is necessary to explain, post operations, the behaviour of the AI-based system and its interactions with the end user, as well as the means to retrieve this data.	Ant.MOC EXP-09-1 Ant.MOC EXP-09-2 Ant.MOC EXP-09-3 Ant.MOC EXP-09-4 Ant.MOC EXP-09-5
C4. Human factors for AI		
LoA	**EASA Objectives**	**Anticipated MOC**
C4.1(EXP). AI operational explainability		
1B-2B	Obj.EXP-10: For each output of the AI-based system relevant to task(s) (per Objective CO-02), the applicant should characterise the need for explainability.	*Dependencies: Obj.EXP-03 Obj.CO-02
1B-2B	Obj.EXP-11: The applicant should ensure that the AI-based system presents explanations to the end user in a clear and unambiguous form.	Ant.MOC EXP-11
1B-2B	Obj.EXP-12: The applicant should define relevant explainability so that the receiver of the information can use the explanation to assess the appropriateness of the decision / action as expected.	Ant.MOC EXP-12
1B-2B	Obj.EXP-13: The applicant should define the level of abstraction of the explanations, taking into account the characteristics of the task, the situation, the level of expertise of the end user and the general trust given to the system.	Ant.MOC EXP-13
1B-2B	Obj.EXP-14: Where a customisation capability is available, the end user should be able to customise the level of abstraction as part of the operational explainability.	Ant.MOC EXP-14
1B-2B	Obj.EXP-15: The applicant should define the timing when the explainability will be available to the end user taking into	Ant.MOC EXP-15/16

	account the time criticality of the situation, the needs of the end user, and the operational impact.	
1B-2B	Obj.EXP-16: The applicant should design the AI-based system so as to enable the end user to get upon request explanation or additional details on the explanation when needed.	Ant.MOC EXP-15/16
1B-2B	Obj.EXP-17: For each output relevant to the task(s), the applicant should ensure the validity of the specified explanation.	
1A–2B	Obj.EXP-18: The training and instructions available for the end user should include procedures for handling possible outputs of the ODD monitoring and output confidence monitoring.	
1A–2B	Obj.EXP-19: Information concerning unsafe AI-based system operating conditions should be provided to the end user to enable them to take appropriate corrective action in a timely manner.	
C4.2(HF). Human-AI teaming		
2A-2B	Obj.HF-01: The applicant should design the AI-based system with the ability to build its own individual situation representation.	Ant.MOC HF-01
2A-2B	Obj.HF-02: The applicant should design the AI-based system with the ability to reinforce the end-user individual situation awareness.	Ant.MOC HF-02
2B only	Obj.HF-03: The applicant should design the AI-based system with the ability to enable and support a shared situation awareness.	Ant.MOC HF-03
2A-2B	Obj.HF-04: If a decision is taken by the AI-based system that requires validation based on procedures, the applicant should design the AI-based system with the ability to request a cross-check validation from the end user.	Ant.MOC HF-04
2A-2B	Obj.HF-05: For complex situations under normal operations, the applicant should design the AI-based system with the ability to identify a suboptimal strategy and propose through argumentation an improved solution.	Ant.MOC HF-05
2A-2B	Corollary Obj.HF-05: The applicant should design the AI-based system with the ability to process and act upon a proposal rejection from the end user.	
2B only	Obj.HF-06: For complex situations under abnormal operations, the applicant should design the AI-based system with the ability to identify the problem, share the diagnosis including the root cause, the resolution strategy and the anticipated operational consequences.	Ant.MOC HF-06 *Dependencies: Obj.HF-05

2B only	Corollary Obj.HF-06: The applicant should design the AI-based system with the ability to process and act upon arguments shared by the end user.	
2B only	Obj.HF-07: The applicant should design the AI-based system with the ability to detect poor decision-making by the end user in a time-critical situation, alert and assist the end user.	Ant.MOC HF-07
2B only	Obj.HF-08: The applicant should design the AI-based system with the ability to propose alternative solutions and support its positions.	Ant.MOC HF-08
2B only	Obj.HF-09: The applicant should design the AI-based system with the ability to modify and/or to accept the modification of task allocation pattern (instantaneous/short-term).	Ant.MOC HF-09
C4.3(HF). Modality of interaction and style of interface		
2A-2B	Obj. HF-10: If spoken natural language is used, the applicant should design the AI-based system with the ability to process end-user requests, responses and reactions, and provide an indication of acknowledgement of the user's intentions.	Ant.MOC HF-10
2B only	Obj.HF-11: If spoken natural language is used, the applicant should design the AI-based system with the ability to notify the end user that he or she possibly misunderstood the information.	Ant.MOC HF-11
2B only	Obj.HF-12: If spoken natural language is used, the applicant should design the AI-based system with the ability to identify through the end user responses or his or her action that there was a possible misinterpretation from the end user.	Ant.MOC HF-12
2B only	Obj.HF-13: In case of confirmed misunderstanding or misinterpretation of spoken natural language, the applicant should design the AI-based system with the ability to resolve the issue.	Ant.MOC HF-13
2A-2B	Obj.HF-14: If spoken natural language is used, the applicant should design the AI-based system with the ability to not interfere with other communications or activities at the end user's side.	Ant.MOC HF-14
2B only	Obj.HF-15: If spoken natural language is used, the applicant should design the AI-based system with the ability to provide information regarding the associated AI-based system capabilities and limitations.	Ant.MOC HF-15
2A-2B	Obj.HF-16: If spoken procedural language is used, the applicant should design the syntax of the spoken procedural language so that it can be learned and applied easily by the end user.	

2A-2B	Obj.HF-17: If gesture language is used, the applicant should design the gesture language syntax so that it is intuitively associated with the command that it is supposed to trigger.	Ant.MOC HF-17
2A-2B	Obj.HF-18: If gesture language is used, the applicant should design the AI-based system with the ability to disregard non-intentional gestures.	Ant.MOC HF-18
2B only	Obj.HF-19: If gesture language is used, the applicant should design the AI-based system with the ability to recognise the end-user intention.	
2B only	Obj.HF-20: If gesture language is used, the applicant should design the AI-based system with the ability to acknowledge the end-user intention with appropriate feedback.	Ant.MOC HF-20
2A-2B	Obj.HF-21: If spoken natural language is used, the applicant should design the AI-based system so that this modality can be deactivated for the benefit of other modalities.	Ant.MOC HF-21
2B only	Obj.HF-22: If spoken (natural or procedural) language is used, the applicant should design the AI-based system with the ability to assess the performance of the dialogue.	
2B only	Obj.HF-23: If spoken (natural or procedural) language is used, the applicant should design the AI-based system with the ability to transition between spoken natural language and spoken procedural language, depending on the performance of the dialogue, the context of the situation and the characteristics of the task.	Ant.MOC HF-23
2B only	Obj.HF-24: The applicant should design the AI-based system with the ability to combine or adapt the interaction modalities depending on the characteristics of the task, the operational event and/or the operational environment.	Ant.MOC HF-24
2B only	Obj.HF-25: The applicant should design the AI-based system with the ability to automatically adapt the modality of interaction to the end-user states, the situation, the context and/or the perceived end user's preferences.	Ant-MOC HF-25
C4.4(HF). Error management		
2A-2B	Obj.HF-26: The applicant should design the AI-based system to minimise the likelihood of design-related end-user errors.	Ant.MOC HF-26
2A-2B	Obj.HF-27: The applicant should design the AI-based system to minimise the likelihood of HAIRM-related errors.	Ant.MOC HF-27
2A-2B	Obj.HF-28: The applicant should design the AI-based system to be tolerant to end-user errors.	Ant.MOC HF-28 *Dependencies: Obj.HF-25

		Obj.HF-26 Obj-HF-27
2A-2B	Obj.HF-29: The applicant should design the AI-based system so that in case the end user makes an error while interacting with the AI-based system, the opportunities exist to detect the error.	Ant.MOC HF-29
2A-2B	Obj.HF-30: The applicant should design the AI-based system so that once an error is detected, the AI-based system should provide efficient means to inform the end user.	
C4.5(HF). Failure management		
2B only	Obj.HF-31: The applicant should design the system to be able to diagnose the failure and present the pertinent information to the end user.	Ant.MOC HF-31
2B only	Obj.HF-32: The applicant should design the system to be able to propose a solution to the failure to the end user.	Ant.MOC HF-32
2B only	Obj.HF-33: The applicant should design the system to be able to support the end user in the implementation of the solution.	Ant.MOC HF-33
2B only	Obj.HF-34: The applicant should design the system to provide the end user with the information that logs of system failures are kept for subsequent analysis.	Ant.MOC HF-34
C5. AI safety risk mitigation		
LoA	**EASA Objectives**	**Anticipated MOC**
C5(SRM). AI safety risk mitigation concept and top-level objectives		
1A-2B	Obj.SRM-01: Once activities associated with all other building blocks are defined, the applicant should determine whether the coverage of the objectives associated with the explainability and learning assurance building blocks is sufficient or whether an additional dedicated layer of protection, called hereafter safety risk mitigation, would be necessary to mitigate the residual risks to an acceptable level.	Ant.MOC SRM-01
1A-2B	Obj.SRM-02: The applicant should establish safety risk mitigation means as identified in Objective SRM-01.	Ant.MOC SRM-02 *Dependencies: Obj.SRM-01
C6. Organisations		
LoA	**EASA Objectives**	**Anticipated MOC**
C6.1(ORG). High level provisions and anticipated AMC		

1A-2B	Prov.ORG-01: The organisation should review its processes and adapt them to the introduction of AI technology.	
1A-2B	Prov.ORG-02: In preparation of the Commission Delegated Regulation (EU) 2022/1645 and Commission Implementing Regulation (EU) 2023/203 applicability, the organisation should continuously assess the information security risks related to the design, production and operation phases of an AI/ML application.	Ant AMC ORG-02
1A-2B	Prov.ORG-03: Implement a data-driven 'AI continuous safety assessment' process based on operational data and in-service events.	Ant.AMC ORG-03 *Dependencies: Obj.EXP-09
1A-2B	Prov.ORG-04: The organisation should establish means (e.g. processes) to continuously assess ethics-based aspects for the trustworthiness of an AI-based system with the same scope as for Objective ET-01.	Ant.AMC ORG-04 *Dependencies: Obj.ET-01
1A-2B	Prov.ORG-05: The organisation should adapt the continuous risk management process to accommodate the specificities of AI, including interaction with all relevant stakeholders.	Ant.AMC ORG-05
1A-2B	Prov.ORG-06: The organisation should ensure that the safety-related AI-based systems are auditable by internal and external parties, including especially the approving authorities.	
C6.2(ORG). Competence considerations		
1A-2B	Prov.ORG-07: The organisation should adapt the training processes to accommodate the specificities of AI, including interaction with all relevant stakeholders (users and end users).	Ant.AMC ORG-07 *Dependencies: Prov.ORG-06 Prov.ORG-07
1A-2B	Prov.ORG-08: The organisations operating the AI-based systems should ensure that end users' licensing and certificates account for the specificities of AI, including interaction with all relevant stakeholders.	

Table 10. EASA AI Roadmap 2.0. Concept Paper Issue 2. Objectives and Anticipated Means of Compliance

Appendix B: Application of Objectives to Use Cases

This appendix feeds into section 2.6, by assessing for each of the Solutions in Use Cases UC1 - UC4 whether (yes = 1, no = 0) the Objective proposed by EASA is relevant (R) to the realisation of the final solution, taking into account the concept defined so far and the expected level of automation, and whether the Objective is applicable (A) in the development phase, at the current TRL, with the aim of evaluating if some issues can and should be addressed during the development process to progressively align the solution with certification requirements.

According to the EASA Concept Paper, the Objectives in White are relevant for all Levels of Automation (1A-2B), the Objectives in Green cells are relevant for 1B-2B, the Objectives in Yellow cells are relevant for 2A-2B, and the Objectives in Blue cells are relevant for 2B only. The EASA Concept Paper does not consider 3A or 3B yet. At columns R: zeros in red font are considered relevant according to EASA at that LoA, but considered not relevant for the use case; ones in green font are considered not relevant according to EASA at that LoA, but considered relevant for the use case.

The results are a preliminary evaluation, which could be updated in future work.

EASA Objectives	UC1 Sim TRL2 LoA 1A		UC1 Opt TRL2 LoA 1B		UC2 TRL3 LoA 2A		UC3 TRL1 LoA 1B		UC4 L3 TRL4 LoA 1A		UC4 L5 TRL4 LoA 2B		UC4 L8 TRL4 LoA 3A	
	A	R	A	R	A	R	A	R	A	R	A	R	A	R
C2.1(CO/CL). Characterisation and classification of the AI application														
Obj.CO-01	0	1	1	1	1	1	1	1	1	1	1	1	1	1
Obj.CO-02	0	1	1	1	1	1	1	1	1	1	1	1	1	1
Obj.CO-03	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Obj.CO-04	0	1	1	1	0	1	1	1	1	1	1	1	1	1
Obj.CO-05	0	1	1	1	0	0	1	1	1	1	1	1	1	1
Obj.CO-06	1	1	1	1	0	1	1	1	1	1	1	1	1	1
Obj.CL-01	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Sum	3	7	7	7	4	6	7	7	7	7	7	7	7	7
C2.2(SA). Safety assessment of ML Applications														
Obj.SA-01	0	1	1	1	0	1	1	1	1	1	1	1	1	1
Obj.SA-02	0	1	1	1	0	1		1		1		1		1
Obj.SA-03	0	1	1	1	0	1		1		1		1		1

EASA Objectives	UC1 Sim TRL2 LoA 1A		UC1 Opt TRL2 LoA 1B		UC2 TRL3 LoA 2A		UC3 TRL1 LoA 1B		UC4 L3 TRL4 LoA 1A		UC4 L5 TRL4 LoA 2B		UC4 L8 TRL4 LoA 3A	
	A	R	A	R	A	R	A	R	A	R	A	R	A	R
Sum	0	3	3	3	0	3	1	3	1	3	1	3	1	3
C2.3(IS). Information security risks management														
Obj.IS-01	0	1	0	1	0	1		1	1	1	1	1	1	1
Obj.IS-02	0	1	0	1	0	1		1		1		1		1
Obj.IS-03	0	1	0	1	0	1		1		1		1		1
Sum	0	3	0	3	0	3	0	3	1	3	1	3	1	3
C2.4(ET). Ethics-based assessment														
Obj.ET-01	0	0	0	0	0	1				0		0		0
Obj.ET-02	0	0	0	0	0	1				1		1		1
Obj.ET-03	1	1	1	1	0	1				1		1		1
Obj.ET-04	0	0	0	0	0	1				1		1		1
Obj.ET-05	0	0	0	0	0	1				0		0		0
Obj.ET-06	0	0	1	1	0	1				1		1		1
Obj.ET-07	1	1	1	1	0	1				1		1		1
Obj.ET-08	0	0	0	0	0	1				1		1		1
Sum	2	2	3	3	0	8	0	0	0	6	0	6	0	6
C3.1(DA). Learning assurance														
Obj.DA-01	0	0	0	0	1	1		0						
Obj.DA-02	0	1	0	1	0	1		1						
Obj.DA-03	0	1	0	1	0	1		1						
Obj.DA-04	0	1	0	0	0	1		1						
Obj.DA-05	0	1	0	0	0	1		1						
Obj.DA-06	0	1	0	0	1	1		1						

EASA Objectives	UC1 Sim TRL2 LoA 1A		UC1 Opt TRL2 LoA 1B		UC2 TRL3 LoA 2A		UC3 TRL1 LoA 1B		UC4 L3 TRL4 LoA 1A		UC4 L5 TRL4 LoA 2B		UC4 L8 TRL4 LoA 3A	
	A	R	A	R	A	R	A	R	A	R	A	R	A	R
Obj.DA-07	0	1	0	0	0	1		1						
Obj.DA-08	0	1	0	0	0	1		1						
Obj.DA-09	0	1	0	0	0	1		1						
Obj.DA-10	0	1	0	0	0	1		1						
Sum	0	9	0	2	2	10	0	9	0	0	0	0	0	0
C3.1(DM). Data management														
Obj.DM-01	0	1	0	1	0	0		1						
Obj.DM-02	0	1	0	1	0	0		1						
Obj.DM-03	0	1	0	1	0	1		0						
Obj.DM-04	0	1	0	0	0	1		1						
Obj.DM-05	0	1	0	0	0	0		0						
Obj.DM-06	0	1	0	0	0	0		1						
Obj.DM-02-UL	0	0	0	0	0	0		0						
Obj.DM-07	0	1	0	0	0	1		1						
Obj.DM-08	0	1	0	1	1	1		1						
Sum	0	8	0	4	1	4	0	6	0	0	0	0	0	0
C3.1(LM). Learning process management														
Obj.LM-01	0	0	0	0	1	1		0						
Obj.LM-02	0	0	0	0	1	1		0						
Obj.LM-03	0	1	0	1	0	1		1						
Obj.LM-04	0	1	0	1	1	1		1						
Obj.LM-05	0	0	0	0	0	1		0						

EASA Objectives	UC1 Sim TRL2 LoA 1A		UC1 Opt TRL2 LoA 1B		UC2 TRL3 LoA 2A		UC3 TRL1 LoA 1B		UC4 L3 TRL4 LoA 1A		UC4 L5 TRL4 LoA 2B		UC4 L8 TRL4 LoA 3A	
	A	R	A	R	A	R	A	R	A	R	A	R	A	R
Obj.LM-06	1	1	1	1	0	1		0						
Obj.LM-07-SL	0	1	0	1	0	1		1						
Obj.LM-08	0	0	0	1	0	1		0						
Obj.LM-09	0	1	0	1	0	1		1						
Obj.LM-10	0	1	0	1	0	1		1						
Obj.LM-11	0	0	0	1	0	1		1						
Obj.LM-12	0	1	0	1	0	1		1						
Obj.LM-13	0	1	0	1	1	1		1						
Obj.LM-14	0	1	0	1	0	0		1						
Obj.LM-15	0	1	0	1	0	0		1						
Obj.LM-16	0	1	0	1	0	1		1						
Sum	1	11	1	13	4	14	0	11	0	0	0	0	0	0
C3.1(IMP). Model implementation														
Obj.IMP-01	0	1	0	0	0	1		1						
Obj.IMP-02	0	1	0	0	0	0		1						
Obj.IMP-03	1	1	1	1	0	1		1						
Obj.IMP-04	0	1	0	0	0	0		1						
Obj.IMP-05	1	1	1	1	0	1		1						
Obj.IMP-06	0	1	0	0	0	1		1						
Obj.IMP-07	0	0	0	0	0	1		0						
Obj.IMP-08	1	1	1	1	1	1		1						
Obj.IMP-09	1	1	1	1	1	1		1						
Obj.IMP-10	0	1	0	0	1	1		1						

EASA Objectives	UC1 Sim TRL2 LoA 1A		UC1 Opt TRL2 LoA 1B		UC2 TRL3 LoA 2A		UC3 TRL1 LoA 1B		UC4 L3 TRL4 LoA 1A		UC4 L5 TRL4 LoA 2B		UC4 L8 TRL4 LoA 3A	
	A	R	A	R	A	R	A	R	A	R	A	R	A	R
Obj.IMP-11	1	1	1	1	1	1		1						
Obj.IMP-12	1	1	1	1	1	1		1						
Sum	6	11	6	6	5	10	0	11	0	0	0	0	0	0
C3.1(CM). Configuration management														
Obj.CM-01	1	1	1	1	0	1		1						
Sum	1	1	1	1	0	1	0	1	0	0	0	0	0	0
C3.1(QA). Quality and process assurance														
Obj.QA-01	1	1	1	1	0	1		1						
Sum	1	1	1	1	0	1	0	1	0	0	0	0	0	0
C3.1(RU). Reuse of AI/ML models														
Obj.RU-01	0	1	0	1	0	0		1						
Obj.RU-02	0	1	0	1	0	0		1						
Obj.RU-03	0	1	0	1	0	0		1						
Sum	0	3	0	3	0	0	0	3	0	0	0	0	0	0
C3.1(SU). Surrogate modelling														
Obj.SU-01	0	0	0	0	0	0		0						
Obj.SU-02	0	0	0	0	0	0		0						
Sum	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C3.2(EXP). Development and post-ops AI explainability														
Obj.EXP-01	0	0	0	1	0	1		1		1		1		1
Obj.EXP-02	0	0	0	1	0	1		1		1		1		1
Obj.EXP-03	0	0	0	1	0	1		1		1		1		1
Obj.EXP-04	0	1	0	1	0	1		1		1		1		1

EASA Objectives	UC1 Sim TRL2 LoA 1A		UC1 Opt TRL2 LoA 1B		UC2 TRL3 LoA 2A		UC3 TRL1 LoA 1B		UC4 L3 TRL4 LoA 1A		UC4 L5 TRL4 LoA 2B		UC4 L8 TRL4 LoA 3A	
	A	R	A	R	A	R	A	R	A	R	A	R	A	R
Obj.EXP-05	1	1	1	1	1	1		1		1		1		1
Obj.EXP-06	1	1	1	1	1	1		1		1		1		1
Obj.EXP-07	1	1	1	1	1	1		1		1		1		1
Obj.EXP-08	1	1	1	1	1	1		1		1		1		1
Obj.EXP-09	0	0	0	1	0	1		1		1		1		1
Sum	4	5	4	9	4	9	0	9	0	9	0	9	0	9
C4.1(EXP). AI operational explainability														
Obj.EXP-10	0	0	0	1	0	1				1		1		1
Obj.EXP-11	0	0	0	1	0	1		1		1		1		1
Obj.EXP-12	0	0	0	1	0	1		1		1		1		1
Obj.EXP-13	0	0	0	1	0	1				1		1		1
Obj.EXP-14	0	0	0	0	0	1				1		1		1
Obj.EXP-15	0	0	0	0	0	1		1		1		1		1
Obj.EXP-16	0	0	0	1	0	1		1		1		1		1
Obj.EXP-17	0	0	0	1	0	1		1		1		1		1
Obj.EXP-18	0	1	1	1	0	1				1		1		1
Obj.EXP-19	0	1	0	1	0	1		1		1		1		1
Sum	0	2	1	8	0	10	0	6	0	10	0	10	0	10
C4.2(HF). Human-AI teaming														
Obj.HF-01	0	0	0	0	1	1		0		0		1		1
Obj.HF-02	0	0	0	0	0	0		0		1		1		1
Obj.HF-03	0	0	0	0	0	0		0		0		1		1
Obj.HF-04	0	0	0	0	1	1		0		0		1		1

EASA Objectives	UC1 Sim TRL2 LoA 1A		UC1 Opt TRL2 LoA 1B		UC2 TRL3 LoA 2A		UC3 TRL1 LoA 1B		UC4 L3 TRL4 LoA 1A		UC4 L5 TRL4 LoA 2B		UC4 L8 TRL4 LoA 3A	
	A	R	A	R	A	R	A	R	A	R	A	R	A	R
Obj.HF-05	0	0	0	0	0	0		0		1		1		1
Cor.Obj.HF-05	0	0	0	0	0	0		0		1		1		1
Obj.HF-06	0	0	0	0	0	0		0		0		1		1
Cor.Obj.HF-06	0	0	0	0				0		0		0		1
Obj.HF-07	0	0	0	0				0		0		1		1
Obj.HF-08	0	0	1	1				0		1		1		1
Obj.HF-09	0	0	0	0				0		1		1		1
Sum	0	0	1	1	2	2	0	0	0	5	0	10	0	11
C4.3(HF). Modality of interaction and style of interface														
Obj.HF-10	0	0	0	0				0	0	0	0	0	0	0
Obj.HF-11	0	0	0	0				0	0	0	0	0	0	0
Obj.HF-12	0	0	0	0				0	0	0	0	0	0	0
Obj.HF-13	0	0	0	0				0	0	0	0	0	0	0
Obj.HF-14	0	0	0	0				0	0	0	0	0	0	0
Obj.HF-15	0	0	0	0				0	0	0	0	0	0	0
Obj.HF-16	0	0	0	0				0	0	0	0	0	0	0
Obj.HF-17	0	0	0	0				0	0	0	0	0	0	0
Obj.HF-18	0	0	0	0				0	0	0	0	0	0	0
Obj.HF-19	0	0	0	0				0	0	0	0	0	0	0
Obj.HF-20	0	0	0	0				0	0	0	0	0	0	0
Obj.HF-21	0	0	0	0				0	0	0	0	0	0	0
Obj.HF-22	0	0	0	0				0	0	0	0	0	0	0

EASA Objectives	UC1 Sim TRL2 LoA 1A		UC1 Opt TRL2 LoA 1B		UC2 TRL3 LoA 2A		UC3 TRL1 LoA 1B		UC4 L3 TRL4 LoA 1A		UC4 L5 TRL4 LoA 2B		UC4 L8 TRL4 LoA 3A	
	A	R	A	R	A	R	A	R	A	R	A	R	A	R
Obj.HF-23	0	0	0	0				0	0	0	0	0	0	0
Obj.HF-24	0	0	0	0				0		1		1		1
Obj.HF-25	0	0	0	0				0		1		1		1
Sum	0	0	0	0	0	0	0	0	0	2	0	2	0	2
C4.4(HF). Error management														
Obj.HF-26	0	0	0	1				1		1		1		1
Obj.HF-27	0	0	0	0				0		0		0		0
Obj.HF-28	0	0	0	1				0		0		1		1
Obj.HF-29	0	0	0	1				1		1		1		1
Obj.HF-30	0	0	0	1				1		0		1		1
Sum	0	0	0	4	0	0	0	3	0	2	0	4	0	4
C4.5(HF). Failure management														
Obj.HF-31	0	0	1	1				1		1		1		1
Obj.HF-32	0	0	0	0				0		1		1		1
Obj.HF-33	0	0	0	0				0		1		1	1	1
Obj.HF-34	1	1	1	1				1		1		1		1
Sum	1	1	2	2	0	0	0	2	0	4	0	4	1	4
C5(SRM). AI safety risk mitigation concept and top-level objectives														
Obj.SRM-01	0	1	0	1				1		1		1		1
Obj.SRM-02	0	1	0	1				1		1		1		1
Sum	0	2	0	2	0	0	0	2	0	2	0	2	0	2
C6.1(ORG). High level provisions and anticipated AMC														
Prov.ORG-01	1	1	1	1				1		1		1		1

EASA Objectives	UC1 Sim TRL2 LoA 1A		UC1 Opt TRL2 LoA 1B		UC2 TRL3 LoA 2A		UC3 TRL1 LoA 1B		UC4 L3 TRL4 LoA 1A		UC4 L5 TRL4 LoA 2B		UC4 L8 TRL4 LoA 3A	
	A	R	A	R	A	R	A	R	A	R	A	R	A	R
Prov.ORG-02	1	1	1	1				1		1		1		1
Prov.ORG-03	0	1	1	1				1		1		1		1
Prov.ORG-04	0	1	0	1				1		1		1		1
Prov.ORG-05	0	1	1	1				1		1		1		1
Prov.ORG-06	1	1	1	1				1		1		1		1
C6.2(ORG). Competence considerations														
Prov.ORG-07	0	1	0	1				1		1		1		1
Prov.ORG-08	1	1	1	1				1		1		1		1
Sum	4	8	6	8	0	0	0	8	0	8	0	8	0	8

Table 11. HUCAN UCs. Relevance and applicability assessments

Appendix C: Liability and Human Factor analysis relevant for PBRs and KPIs

This appendix feeds into section 3.3 and identifies aspects of liability and human factors to be taken into account for the development of PBRs and KPIs. For each of 8 cases, the analysis identifies relevant stakeholders that are potentially liable, as the bearers of responsibility and accountable entities, including their interrelationships. For each stakeholder, it includes mitigation strategies, to support the research of solutions in an effort to tackle the complications highlighted with this analysis, and directly supporting the drafting of PBRs and KPIs.

i. Loss of System Control

Higher levels of automation in ATM systems may lead to operators becoming overly reliant on automated systems, which can hinder their ability to regain control in emergency situations.

Stakeholders Potentially Liable:

Operators (Individuals): liability may arise if operators fail to intervene when necessary due to complacency or lack of familiarity with manual controls. In particular, operators may be liable if they fail to regain control during system failures due to complacency or insufficient training. This includes situations where operators become overly reliant on automation. Operators that had not received adequate training in manual overrides may incur in liability for not having responded effectively during system failures. Moreover, the organisation employing the operators may also incur liability if it provides insufficient or inadequate training on manual overrides and emergency procedures. In cases where operators are unprepared to respond effectively during system failures, liability may be assigned to management for failing to equip staff with necessary skills.

Mitigation Strategy > operators should undergo regular, rigorous training on manual control procedures, especially in high-risk, highly automated systems. Periodic simulations of emergency scenarios should be conducted to ensure operators maintain familiarity with manual controls and intervention techniques. The end-user organisation is responsible for implementing and enforcing these training programs to support operator readiness and system safety

System Developers (Organisations and Entities): developers can be liable if their systems lack sufficient manual override capabilities or fail to provide clear and intuitive guidance on regaining control during emergencies. This includes designing interfaces that do not effectively alert operators when manual control is needed. The absence of a robust manual intervention process by these entities may significantly contribute to operational risks and therefore potentially determine liabilities.

Mitigation Strategy > developers should incorporate comprehensive manual override systems into automated solutions. These systems should be tested extensively in real-world simulations to ensure safe and risks-minimised usability. Clear guidance, alerts and system prompts must be included in the interface to aid operators in taking over manual control swiftly and effectively during system malfunctions.

Maintenance Organizations (Organizations): Maintenance responsibilities often fall under the departments of end-user organisations, such as airlines or Air Navigation Service Providers (ANSPs). These organisations are accountable for ensuring that both automated and manual control systems are operational. Negligence in maintaining these systems could result in liability if system failures occur, particularly if maintenance oversights lead to malfunctions that operators are unable to correct in time.

Mitigation Strategy > Maintenance organisations must implement stringent, routine maintenance checks on all control systems, including regular assessments of AI models. Maintenance logs should be reviewed periodically to ensure no issues are overlooked. Furthermore, end-user organisations must collaborate with system developers to establish clear guidelines for maintaining AI-based systems and ensure the continuous operational integrity of both manual and automated components. This collaboration will help address the nuances of maintenance responsibilities and enhance overall safety and reliability in ATM operations.

ii. Human-Computer Interaction

At medium or advanced levels of automation, the advanced automation itself can complicate human-computer interactions, resulting in misunderstandings or errors during critical decision-making processes. Operators struggling with the interface may lead to incorrect responses to alerts.

Stakeholders Potentially Liable:

Operators (Individuals): operators might misinterpret system alerts or recommendations due to a poorly designed interface, leading to operational failures. Unclear data presentation may result in significant operator errors. Such misunderstandings can lead to operational failures or safety incidents. If operators fail to respond correctly to critical alerts because the interface does not clearly convey necessary information, their liability may be influenced by several factors. Operators are expected to effectively use the tools and systems provided to them. If they misinterpret alerts due to poor design, they may still be liable for failing to act appropriately, especially if they did not seek clarification or assistance when faced with unclear information. Operators' liability in instances of human-computer interaction failures is contingent upon the clarity of the interface, the adequacy of training provided, and the expectation of reasonable competence in navigating the system. This highlights the importance of well-designed human-computer interfaces that facilitate effective decision-making and minimise the risk of errors during critical operations.

Mitigation Strategy > operators should receive comprehensive training on system interfaces and decision-making tools in automated environments. Training should emphasise how to interpret alerts and recommendations accurately, even under stressful conditions. Simulations and hands-on exercises should replicate real-world scenarios to help operators better understand system responses and interactions. Additionally, regular feedback loops between operators and system developers should be established, ensuring that any challenges in interacting with the system are addressed promptly.

System Developers (Organizations): developers could be held liable if their interfaces fail to provide clear and intuitive communication between the system and operators.

Mitigation Strategy > system developers must prioritise human factors engineering in the design of interfaces, focusing on simplicity, clarity, and usability. Usability testing should be a mandatory phase in the development process, with real operators providing input to ensure that the system's communication and feedback mechanisms are intuitive. Incorporating adaptive interfaces, which adjust to the operator's level of expertise or current workload, can further reduce the likelihood of misinterpretation. Developers should also introduce customizable settings, enabling operators to configure the interface according to their preferences, without compromising safety.

Maintenance Organizations (Organizations): maintenance organisations may be liable if software updates or system enhancements negatively affect the usability of the human-machine interface, resulting in operator miscommunication or confusion. Ignoring feedback from operators regarding interface issues after system upgrades may contribute to incidents and consequently to liabilities.

Mitigation Strategy > maintenance organisations must ensure that every software update or system upgrade is tested rigorously for any potential impact on human-computer interaction. Feedback from operators should be continuously gathered and incorporated into future updates. Maintenance teams should work closely with developers to address interface challenges identified in the field, ensuring that any usability issues are rectified promptly. Additionally, post-update usability audits should be carried out, and new interface features should be introduced gradually, with adequate training provided to operators before they are deployed.

iii. Lack of Information or Data Misinterpretation

At all levels of automation, also low, automated systems may present information in ways that can be easily misinterpreted by operators, especially when under pressure.

Stakeholders Potentially Liable:

Operators (Individuals)*:* operators are liable if they fail to properly interpret critical data, which can lead to serious operational mistakes. High levels of automation can exacerbate this risk if the system data becomes too complex or ambiguous, implying a shift of liability towards the developer. If the data presentation is ambiguous or overly complex, liability may shift towards the developers or system designers. They have a responsibility to ensure that information is presented clearly and understandably, especially in high-pressure scenarios where quick decision-making is essential. While operators retain a level of responsibility for interpreting critical data accurately, a significant degree of liability can shift toward developers if the ambiguity or complexity of the information stems from poor design or lack of clarity (see section below).

Mitigation Strategy > operators should receive thorough training focused on data interpretation and system responses. User-friendly interfaces with clearly presented data, along with real-time alerts, may reduce the cognitive load on operators. Frequent simulations should be implemented to refine operator responses to complex data scenarios.

System Developers (Organizations): Developers may face liability if their systems do not provide clear, contextualised data, leading to misinterpretation by operators. Inadequate data presentation may contribute to decision-making errors. If the data presentation is ambiguous or overly complex, liability may shift towards the developers or system designers. They have a responsibility to ensure that

information is presented clearly and understandably, especially in high-pressure scenarios where quick decision-making is essential. Factors influencing this shift in liability may include the following cases: i) if the system's design does not facilitate easy understanding of the information being presented, operators may have grounds to argue that the developers are at least partially responsible for any resulting errors. Clear and intuitive interfaces are critical to enabling effective human-computer interaction; ii) developers are expected to adhere to established industry standards for data presentation and usability. If they fail to do so, it may strengthen the case for liability on their part. This includes ensuring that alerts, warnings, and critical data are designed to minimise the risk of misinterpretation; iii) the adequacy of training provided to operators on how to interpret and respond to data can also play a role in determining liability. If developers fail to provide comprehensive documentation or support for their systems, this could contribute to an operator's misinterpretation of the information.

Mitigation Strategy > developers should ensure that the human-machine interface (HMI) is intuitive and designed with human factors in adequate consideration. Critical information must be highlighted and made easily accessible to operators. Developing standardised display formats that simplify data presentation and reduce operator confusion can mitigate this risk.

Maintenance organisations*:* maintenance organisations are responsible for ensuring that data reporting systems function accurately and consistently reflect real-time operational conditions. If they fail to maintain or update **these** systems properly, leading to inaccuracies in data reporting **or lacking data**, they could be liable for any resulting operator errors. This liability is especially critical **at higher levels of automation or** when updates or system changes introduce new complexities or vulnerabilities in data interpretation.

***Mitigation Strategy* >** maintenance organisations should implement strict protocols for the regular testing, calibration, and updating of data reporting systems to ensure that they continue to provide accurate and real-time information. Comprehensive post-maintenance verification procedures should be put in place, particularly following system updates or patches, to confirm that data presentation remains reliable and free from errors. Additionally, they should establish a feedback mechanism where operators can report data interpretation issues, ensuring that potential problems are identified and addressed promptly. Regular audits of system performance should be carried out, focusing on data accuracy and system integrity, to proactively detect and mitigate potential failures.

iv. Regulatory Violation or Non-Compliance with Certification Standards

Increased level automation (medium or high) in air traffic management systems must adhere to regulatory standards; failure to do so can result in significant liability. Non-compliance may arise from inadequate implementation or oversight of the automated systems.

Stakeholders Potentially Liable:

Operators (Individuals): operators may be liable if they voluntarily engage in practices that violate regulations or fail to report issues, thereby jeopardising safety.

Mitigation Strategy > operators should undergo regular and thorough training on relevant regulatory frameworks and certification standards. This training should emphasise the impact of automation on

compliance, ensuring operators understand how to monitor and ensure the system remains within regulatory limits. Moreover, implement tools within the automated system that can assist operators in identifying potential non-compliance, for example by providing real-time alerts or prompts for potential regulatory breaches, operators are better equipped to respond proactively.

System Developers (Organisations and Entities): system developers may face liability if their systems do not meet regulatory standards or fail to secure necessary certifications, leading to system failures or safety breaches.

Mitigation Strategy > developers should maintain close communication with regulatory authorities throughout the development lifecycle to ensure compliance with evolving regulatory and certification standards. Implementing a compliance checklist for each phase of development ensures that all necessary certifications and approvals are obtained before deployment.

Maintenance Organizations (Organisations and Entities): maintenance organisations are also liable if they fail to ensure that automated systems remain compliant with regulatory standards post-deployment. Negligence in maintenance practices, including not keeping up with system updates or failing to perform required inspections, can lead to violations that compromise safety.

Mitigation Strategy > maintenance organisations must establish and adhere to rigorous maintenance protocols that include regular audits and inspections of both automated and manual systems to ensure ongoing compliance with regulatory standards. They should also implement a structured program for tracking and documenting maintenance activities and any regulatory changes that may affect compliance. Additionally, providing training for maintenance personnel on the importance of regulatory compliance, including understanding how automated systems can evolve and require updates, is crucial. Collaboration with system developers to stay informed about updates to regulatory standards is also essential for maintaining compliance.

v. Over-Reliance on Automation

At higher levels of automation, operators may become complacent, relying too heavily on automated systems without adequate situational awareness.

Stakeholders Potentially Liable:

Operators (Individuals)*:* Operators can become overly dependent on automation, reducing their situational awareness and ability to intervene when necessary. If they fail to act because of over-reliance on automation, they could be held liable for errors or incidents.

Mitigation Strategy > operators must undergo regular training that emphasises manual intervention and situational awareness in highly automated environments. Training programs should include simulations of system failures that require operators to override automation and take control. Operators should also receive education on system limitations, reinforcing the importance of their active role even when automation is functioning properly. Monitoring tools that alert operators when they are disengaging too much from oversight can help maintain their attentiveness.

System Developers (Organizations): developers may face liability if the design of their systems encourages excessive reliance on automation without clearly communicating the limitations or failure conditions of the automated features.

Mitigation Strategy > developers should design systems that encourage human oversight and actively communicate system limitations to operators. This could involve incorporating alerts or notifications that remind operators of the automation's boundaries, or systems that periodically require manual inputs to keep operators engaged. Developers should also design interfaces that provide sufficient and timely information to operators about system performance and potential issues. Regular human-machine interaction testing should be conducted to ensure that the system supports rather than undermines operator engagement.

Maintenance Organizations (Organizations): maintenance teams must ensure that automation systems are periodically reviewed and assessed for reliability. Maintenance organisations may be liable if they fail to address issues in automated systems that contribute to over-reliance, such as delayed updates or unaddressed system warnings that cause operators to over-trust the system.

Mitigation Strategy > maintenance organisations should implement thorough and frequent reviews of automated systems to identify and fix issues that could contribute to operator complacency. This includes regular software updates, patches to enhance system transparency, and addressing known bugs or reliability concerns that may give operators a false sense of security (though on-board certified software usually is not updated as this would require a new pass through the certification process). Additionally, maintenance teams should collaborate with developers to ensure that system diagnostics and failure warnings are correctly calibrated and actively communicated to operators. Feedback loops between operators, developers, and maintenance teams can help identify potential over-reliance risks and ensure that systems maintain a balance between automation and operator oversight.

vi. Human Oversight

Increased automation (medium or high) may result in reduced human oversight of critical systems, leading to complacency and errors. In situations where operators are not actively monitoring automated systems, significant risks can arise, especially during unexpected events. For example, if an automated alert is triggered but the operator fails to respond due to distraction or overconfidence in the automation, an incident may occur.

Stakeholders Potentially Liable:

Operators (Individuals): operators may be liable if they fail to maintain vigilance and oversight of automated systems. Liability could arise if they fail to maintain adequate vigilance over automated systems, especially when critical alerts or situations arise. If operators are distracted or overconfident in the automation and miss important warnings, they can be held responsible for incidents.

Mitigation Strategy > operators should undergo regular training that emphasises the importance of continuous monitoring, even in highly automated environments. Training sessions should include drills that simulate situations where manual intervention is required to address system alerts. Additionally, establishing protocols for periodic checks of the automated system's status (e.g., via dashboards or

system alerts) can help operators stay engaged. Technology solutions like particular alert to operators when their vigilance drops can also assist in ensuring active oversight.

System Developers (Organizations): developers may be held responsible if their systems do not promote or require sufficient human oversight. Inadequate design or communication that encourages passivity can contribute to liability. If systems fail to require adequate human oversight or do not effectively alert operators of critical events, developers could be implicated in incidents caused by operator inaction.

Mitigation Strategy > developers should ensure that systems are designed with human oversight as a priority. This includes integrating features that actively engage operators, such as requiring manual acknowledgment of critical alerts or periodic input to confirm operator attention. The interface should clearly differentiate between routine and critical alerts, ensuring that high-priority warnings are impossible to overlook. Developers could also introduce systems that detect when operators are disengaged (e.g., through eye-tracking or other biometrics) and re-engage them with prompts. Testing for human oversight in varying operational scenarios should be an integral part of system development.

Maintenance Organizations (Organizations): maintenance organisations are responsible for ensuring that alert systems and monitoring mechanisms continue to function properly over time. If maintenance does not update or repair systems to ensure their reliability, leading to operator inaction, they may be held partially liable for incidents.

Mitigation Strategy > maintenance organisations should implement a robust schedule of inspections and updates for systems, particularly focusing on the reliability of alert and monitoring tools. This includes ensuring that alert systems are not only operational but calibrated to trigger attention appropriately. Maintenance teams should also work closely with operators and developers to gather feedback on system performance, especially regarding human-machine interaction. By continuously refining alert and oversight mechanisms, maintenance teams can help foster effective human supervision and reduce the risk of operator complacency.

vii. Inadequate Training and Skill Gaps

At medium or high levels of automation, the management of ATM automated systems can reveal or create skill gaps among stakeholders, especially if they are not adequately trained to handle manual operations or interventions.

Stakeholders Potentially Liable:

Operators (Individuals): operators may be liable if they lack the necessary skills or training to intervene effectively during system failures. Unprepared operators, relying too heavily on automation, made critical errors when manual intervention was needed.

Mitigation Strategy > operators should undergo regular and comprehensive training programs that focus on both automated system management and manual intervention techniques. Training must include real-world simulations where operators practise taking control during automation failures. Continuous education through refresher courses, particularly on manual procedures and system

overrides, should be mandatory to ensure operators are always prepared to act effectively in critical situations. Competency tests should also be conducted periodically to identify skill gaps and address them proactively.

System Developers (Organizations): system developers may share liability if their systems are deployed without providing sufficient training materials or support. Systems that are highly automated but fail to educate operators on potential manual interventions or troubleshooting processes increase the likelihood of operator errors, as evidenced in several use cases.

Mitigation strategy > system developers should collaborate with operators and trainers to create detailed, accessible training programs for all levels of system automation. These programs should include interactive tutorials, manuals, and simulations that prepare users for both typical and emergency scenarios. The training should highlight system limitations and areas where manual intervention may be required. In addition, developers should ensure that training evolves alongside system updates, and that operators are informed about any new features or changes in the system.

Maintenance Organizations (Organizations): Maintenance teams should support ongoing training initiatives. If they fail to collaborate with training departments or neglect to provide input on maintenance-related skills, they may be implicated in liability claims arising from operator errors. Maintenance teams are responsible for ensuring that systems remain operable and that maintenance-related skills are part of the operators' training. If they fail to provide feedback or collaborate on updating training materials regarding maintenance protocols, they could share liability for incidents.

Mitigation Strategy > maintenance teams should work closely with training departments to ensure that operators are well-versed in system maintenance, troubleshooting, and emergency procedures. This includes providing detailed feedback on common technical failures and identifying areas where operator skills may need strengthening. Maintenance organisations should contribute to developing real-time operational guidelines and support operators with hands-on training in basic maintenance tasks, especially for manual system interventions. Furthermore, they should participate in post-incident reviews to ensure that any lessons learned about maintenance issues are incorporated into future training programs.

viii. Difficulties in the allocation of Responsibility or Unclear Responsibility During Automation Failures

At all levels of automation, when an automated system fails, confusion and difficulty can arise regarding who is the subject responsible for addressing the failure, particularly in high-stakes situations.

Stakeholders Potentially Liable:

Operators (Organisations/Individuals): operators may be held liable if they fail to take action during a system failure due to unclear protocols regarding responsibility. However, it is important to recognize that holding individual operators liable in such cases may not align with the principles of a just culture, which promotes accountability while encouraging open reporting and learning from errors without fear of retribution

Mitigation Strategy > clear, well-documented protocols should be established, specifying when and how operators are expected to act in case of automation failures. These protocols must be reinforced through training that covers role-specific responsibilities during system breakdowns, as well as decision-making procedures in ambiguous situations. Role-playing exercises and simulations should be regularly conducted to ensure operators feel confident in taking responsibility when required. Establishing clear lines of communication between operators and their supervisors, and other stakeholders can also mitigate hesitation. To foster a just culture, it is essential to focus on system improvements to minimise risks of individual operators' liabilities. Role-playing exercises and simulations should be regularly conducted to ensure operators feel confident in taking responsibility when required. Establishing clear lines of communication between operators, their supervisors, and other stakeholders can also mitigate hesitation and clarify responsibilities during critical incidents.

System Developers (Organizations): developers may face liability if they fail to create systems with clearly defined roles and responsibilities in case of system failure. Those responsible for developing the operational protocols (protocol developers) may also face liability if the protocols are unclear or inadequate. This includes both system developers and the end-user organisation that implements these protocols. Their responsibility lies in ensuring that operators have clear, actionable guidance during system failures.

Mitigation Strategy > system developers should design automation systems with explicit operational guidelines that detail the chain of responsibility during system breakdowns. These guidelines should be integrated into the system interface or be readily accessible during operations to avoid ambiguity. Developers should work closely with operators, regulators, and maintenance teams to ensure that these guidelines are clear and consistently enforced. In addition, developers should ensure that automated systems provide real-time prompts or alerts to remind operators of their roles and required actions during emergencies, reducing uncertainty. Clear, well-documented protocols should be established by responsible developers, specifying when and how operators are expected to act in case of automation failures. These protocols must be reinforced through training that covers role-specific responsibilities during system breakdowns, as well as decision-making procedures in ambiguous situations.

Maintenance Organizations (Organizations)*:* maintenance teams must ensure that responsibilities are clearly outlined and communicated to all stakeholders. If they do not reinforce these protocols, they may share liability in incidents stemming from miscommunication. Maintenance teams can share liability if they fail to communicate or reinforce clearly defined responsibility protocols, leading to operational confusion during system failures.

Mitigation Strategy > maintenance organisation must collaborate with system developers and operational managers to ensure that responsibility protocols are well-established and communicated to all stakeholders. Maintenance organisations should participate in regular reviews of failure protocols to ensure that these remain up-to-date and relevant. They should also provide feedback to operators and developers when system updates or maintenance procedures alter the operational responsibilities. Additionally, maintenance teams should help conduct refresher training for all stakeholders, focusing on updated procedures and how to respond during system malfunctions.

Appendix D: KPIs and Milestones for EASA Objectives

This appendix feeds into Section 3.5 by providing, for each of the Objectives identified in (EASA, 2024b) and listed in Appendix A, one or more potential KPIs, and associated Milestones. It is noted that the development of objectives by EASA as reported in (EASA, 2024b) is an ongoing process, such that the list of objectives is not finalised. The identification of potential KPIs and milestones has been done in the HUCAN project on the basis of anticipated MOCs in (EASA, 2024b) and our interpretation of relevant KPIs and milestones. They are a first set only and should be further developed.

C2. Trustworthiness analysis		
Objectives	**KPIs**	**Milestones**
C2.1(CO/CL). Characterisation and classification of the AI application		
Obj.CO-01	List of end users that are intended to interact with the AI-based system. Roles of end users that are intended to interact with the AI-based system. Responsibilities of end users that are intended to interact with the AI-based system (including indication of the level of teaming with the AI-based system). Expected expertise of end users (including assumptions made on the level of training, qualification and skills).	List of users is completed. List of users has been validated by independent means. Roles of users have been defined. Roles have been validated by independent means. Responsibilities of end users have been defined. Responsibilities have been validated by independent means. Expected expertise has been determined. Expected expertise has been validated by independent means.
Obj.CO-02	For each end user, the list of goals that are intended to be performed in interaction with the AI-based system. For each end user, the high-level tasks (associated with the goals) that are intended to be performed in interaction with the AI-based system.	List of goals is completed. The list of goals has been validated by independent means. The list of high level tasks relevant to the end users, in interaction with the AI-based system, has been defined and documented. The list has been validated by independent means.
Obj.CO-03	The domain-specific AI-based system.	The AI-based system that has been determined takes into account domain-specific definitions of 'system'. If relevant, the system has been decomposed into (AI-based) subsystem(s).
Obj.CO-04	The ConOps for the AI-based system, including the task allocation pattern	The ConOps for the AI-based system including the task allocation pattern has been documented.

	between the end user(s) and the AI-based system.	The ConOps has been validated by independent means. It has been shown that the focus is put on the definition of the OD and on the capture of specific operational limitations and assumptions.
Obj.CO-05	Document that describes how end users' inputs have been collected and accounted for in the development of the AI-based system.	The document has been completed. The document has been validated by independent means.
Obj.CO-06	A functional analysis of the system. A functional decomposition and allocation down to the lowest level.	A functional analysis of the system has been completed. The functional analysis has been reviewed and validated by independent means. A functional decomposition of the system has been completed. The decomposition shows which items are AI/ML, and which items are non AI/ML. The functional decomposition has been reviewed and validated by independent means.
Obj.CL-01	Classification of the AI-based system, based on the levels presented by EASA. Justification of the classification.	The AI-based system has been classified with justification. The classification takes into consideration the high-level task(s) that are allocated to the end user(s), in interaction with the AI-based system. The classification and justification have been reviewed and validated by independent means.
C2.2(SA). Safety assessment of ML Applications		
Obj.SA-01	A safety assessment for all AI-based (sub)systems, identifying and addressing specificities introduced by AI/ML usage. List of sources of uncertainties. List of varying conditions. Automation failure rate. Human-automation interaction failures. Number of cybersecurity breaches. System vulnerability mitigation.	A safety assessment has been performed for all AI-based (sub)systems. The safety assessment fits the specifics of the aviation domain in which the AI-based system is used, but takes a holistic approach. The safety assessment identifies and addresses specificities introduced by AI/ML usage. The safety assessment has been validated by independent means.

Obj.SA-02	Identification of data that needs to be recorded for the purpose of supporting the continuous safety assessment. Epistemic and aleatory uncertainties in the data.	The data has been identified. The identification is validated by independent means.
Obj.SA-03	List of design assumptions. Metrics, target values, thresholds and evaluation periods to guarantee that design assumptions hold.	The list of design assumptions has been documented. The design assumptions have been validated by independent means. Metrics, target values, thresholds and evaluation periods to guarantee that design assumptions hold have been identified. Metrics, target values, thresholds and evaluation periods have been validated by independent means.
C2.3(IS). Information security risks management		
Obj.IS-01	List of information security risks with an impact on safety.	A list of information security risks with an impact on safety has been determined. It has been validated that the identified information security risks address specific threats introduced by AI/ML usage.
Obj.IS-02	The mitigation approach to address the identified AI/ML-specific information security risks.	The mitigation approach to address the identified AI/ML-specific information security risks have been documented. The mitigation approach has been validated by independent means.
Obj.IS-03	The effectiveness of the security controls introduced to mitigate the identified AI/ML-specific information security risks to an acceptable level.	The effectiveness of the security controls introduced to mitigate the identified AI/ML-specific information security risks to an acceptable level have been validated and verified.
C2.4(ET). Ethics-based assessment		
Obj.ET-01	An ethics-based trustworthiness assessment for any AI-based system developed using ML techniques or incorporating ML models.	An ethics-based trustworthiness assessment has been completed for any AI-based system developed using ML techniques or incorporating ML models. This assessment has verified Transparency, Responsiveness, understandability, Sociability. The ethics-based trustworthiness assessment has been reviewed by independent means.

Obj.ET-02	An assessment of the risk of creating overreliance, attachment, stimulating addictive behaviour, or manipulating the end user's behaviour.	An assessment has been completed of the risk of creating overreliance, attachment, stimulating addictive behaviour, or manipulating the end user's behaviour. The assessment concludes that the AI-based system bears no risk of creating overreliance, attachment, stimulating addictive behaviour, or manipulating the end user's behaviour. The assessment and conclusions have been reviewed and validated by independent means.
Obj.ET-03	List of national and EU data protection regulations (e.g. GDPR).	The applicant has involved their Data Protection Officer. The applicant has consulted with their National Data Protection Authority. The applicant has verified compliance with national and EU data protection regulations (e.g. GDPR). The authorities have confirmed that the AI-based system complies with national and EU data protection regulations (e.g. GDPR).
Obj.ET-04	Assessment of the creation or reinforcement of unfair bias in the AI-based system, regarding both the data sets and the trained models, including an assessment of impact of the unfair bias on performance and safety.	The assessment has been completed. The assessment has been reviewed and validated by independent means. The assessment has shown that any unfair bias in the AI-based system regarding both the data sets and the trained models, that has impact on performance and safety, is avoided.
Obj.ET-05	Any means to make end users aware of the fact that they interact with an AI-based system, and, if applicable, whether some personal data is recorded by the system.	Means have been developed to make end users aware of the fact that they interact with an AI-based system, and, if applicable, whether some personal data is recorded by the system. These means have been implemented. An evaluation has shown that the end users are aware.
Obj.ET-06	An environmental impact analysis that identifies and assesses potential negative impacts of the AI-based system on the environment and human health throughout its life cycle (development, deployment, use, end of life).	The environmental impact analysis has been completed. The results have been reviewed and validated by independent means.

	Measures to reduce or mitigate these impacts.	The results identify and assess potential negative impacts of the AI-based system on the environment and human health throughout its life cycle (development, deployment, use, end of life), and define measures to reduce or mitigate these impacts.
Obj.ET-07	The need for new skills for users and end users to interact with and operate the AI-based system. List of possible training gaps. List of mitigations of possible training gaps.	The identification of the need for new skills for users and end users to interact with and operate the AI-based system has been completed. Possible training gaps have been identified and mitigated. The results have been reviewed and validated by independent means.
Obj.ET-08	Assessment of the risk of de-skilling of the users and end users. A training needs analysis and a consequent training activity aiming to mitigate the identified risk.	An assessment has been completed of the risk of de-skilling of the users and end users. This assessment has identified risk mitigations through a training needs analysis and a consequent training activity. The results have been reviewed and validated through independent means. An evaluation has shown that the mitigations are effective.
C3. AI Assurance		
Objectives	**KPIs**	**Milestones**
C3.1(DA). Learning assurance		
Obj.DA-01	Description of the proposed learning assurance process, taking into account each of the steps described in Sections C.3.1.2 to C.3.1.14, as well as the interface and compatibility with development assurance processes.	The proposed learning assurance process has been described. The description has been reviewed by independent means. The review has validated that the description takes into account each of the steps described in Sections C.3.1.2 to C.3.1.14, as well as the interface and compatibility with development assurance processes.
Obj.DA-02	Capturisation of the following minimum for the AI/ML constituent requirements: — safety requirements allocated to the AI/ML constituent (e.g. performance, reliability, resilience);	The following minimum for the AI/ML constituent requirements have been captured: — safety requirements allocated to the AI/ML constituent (e.g. performance,

	— information security requirements allocated to the AI/ML constituent; — functional requirements allocated to the AI/ML constituent; — operational requirements allocated to the AI/ML constituent, including AI/ML constituent ODD monitoring and performance monitoring (to support related objectives in Section C.3.2.6), detection of OoD input data and data-recording requirements (to support objectives in Section C.3.2.7); — other non-functional requirements allocated to the AI/ML constituent (e.g. scalability); and — interface requirements.	reliability, resilience); — information security requirements allocated to the AI/ML constituent; — functional requirements allocated to the AI/ML constituent; — operational requirements allocated to the AI/ML constituent, including AI/ML constituent ODD monitoring and performance monitoring (to support related objectives in Section C.3.2.6), detection of OoD input data and data-recording requirements (to support objectives in Section C.3.2.7); — other non-functional requirements allocated to the AI/ML constituent (e.g. scalability); and — interface requirements. The capturisation has been reviewed and validated by independent means.
Obj.DA-03	Definition of the set of parameters pertaining to the AI/ML constituent ODD.	The set of parameters pertaining to the AI/ML constituent ODD has been defined and traced to the corresponding parameters pertaining to the OD when applicable. The results have been reviewed and validated by independent means.
Obj.DA-04	Capturisation of the DQRs for all data required for training, testing, and verification of the AI/ML constituent, including but not limited to: — the data relevance to support the intended use; — the ability to determine the origin of the data; — the requirements related to the annotation process; — the format, accuracy and resolution of the data; — the traceability of the data from their origin to their final operation through the whole pipeline of operations; — the mechanisms ensuring that the data will not be corrupted while stored, processed, or transmitted over a communication network; — the completeness and	The DQRs for all data required for training, testing, and verification of the AI/ML constituent have been captured, including but not limited to: — the data relevance to support the intended use; — the ability to determine the origin of the data; — the requirements related to the annotation process; — the format, accuracy and resolution of the data; — the traceability of the data from their origin to their final operation through the whole pipeline of operations; — the mechanisms ensuring that the data will not be corrupted while stored, processed, or transmitted over a communication network;

	representativeness of the data sets; and — the level of independence between the training, validation and test data sets.	— the completeness and representativeness of the data sets; and — the level of independence between the training, validation and test data sets. The capturisation has been reviewed and validated by independent means.
Obj.DA-05	Capturisation of the requirements on data to be pre-processed and engineered for the inference model in development and for the operations.	The requirements have been captured on data to be pre-processed and engineered for the inference model in development and for the operations. The capturisation has been reviewed and validated by independent means.
Obj.DA-06	Description of a preliminary AI/ML constituent architecture, to serve as reference for related safety (support) assessment and learning assurance objectives.	A preliminary AI/ML constituent architecture has been described, to serve as reference for related safety (support) assessment and learning assurance objectives. The description has been reviewed and validated by independent means. An evaluation has shown that the architecture serves as reference for related safety (support) assessment and learning assurance objectives.
Obj.DA-07	Validation of each of the requirements captured under Objectives DA-02, DA-03, DA-04, DA-05 and the architecture captured under Objective DA-06.	The validation has been completed. The result has been reviewed and validated by independent means. The result has shown that each of the requirements captured under Objectives DA-02, DA-03, DA-04, DA-05 and the architecture captured under Objective DA-06 are validated.
Obj.DA-08	Documented evidence that all derived requirements generated through the learning assurance processes have been provided to the (sub)system processes, including the safety (support) assessment.	The documented evidence has been completed. The evidence shows that all derived requirements generated through the learning assurance processes have been provided to the (sub)system processes, including the safety (support) assessment. The results have been reviewed and validated by independent means.
Obj.DA-09	Documented evidence of the validation of the derived requirements, and of the determination of any impact on the	The documented evidence has been completed. The evidence has shown the validation

	safety (support) assessment and (sub)system requirements.	of the derived requirements, and of the determination of any impact on the safety (support) assessment and (sub)system requirements. The results have been reviewed and validated by independent means.
Obj.DA-10	Verification of each of the captured AI/ML constituent requirements.	Each of the captured AI/ML constituent requirements has been verified. The results have been reviewed and validated by independent means.
C3.1(DM). Data management		
Obj.DM-01	Identification of data sources and data in accordance with the defined ODD.	An assessment has shown that data sources and data satisfy the defined DQRs, and drive the selection of the training, validation and test data sets. The results have been reviewed and validated by independent means.
Obj.DM-02-SL	The annotated or labelled data in the data set collected.	An assessment has shown that the annotated or labelled data in the data set satisfies the DQRs captured under Objective DA-04. The results have been reviewed and validated by independent means.
Obj.DM-03	Definition of the data preparation operations.	An assessment has shown that the data preparation operations properly address the captured requirements (including DQRs). The results have been reviewed and validated by independent means.
Obj.DM-04	Definition and documentation of pre-processing operations on the collected data in preparation of the model training.	Pre-processing operations on the collected data in preparation of the model training have been defined. The results have been reviewed and validated by independent means.
Obj.DM-05	Definition and documentation of the transformations to the pre-processed data from the specified input space into features.	The transformations to the pre-processed data from the specified input space have been defined and documented. The results have been reviewed and validated by independent means. The results show that the features are effective for the performance of the selected learning algorithm.
Obj.DM-06	Distribution of the data into three	The data has been distributed into

	separate data sets: — the training data set and validation data set, used during the model training; — the test data set used during the learning process verification, and the inference model verification.	three separate data sets: — the training data set and validation data set, used during the model training; — the test data set used during the learning process verification, and the inference model verification. The results have been reviewed and validated by independent means. The results are shown to meet the specified DQRs in terms of independence (as per Objective DA-04).
Obj.DM-02-UL	Assessment of the annotated or labelled data in the test data set.	The assessment has shown that the annotated or labelled data in the test data set satisfies the DQRs captured under Objective DA-04. The results have been reviewed and validated by independent means.
Obj.DM-07	Assessment of the data management process.	A data management process has been implemented. The process has been reviewed and validated by independent means. The process includes verification of the data, as appropriate, so that the data management requirements (including the DQRs) are addressed.
Obj.DM-08	Data verification step to confirm the appropriateness of the defined ODD and of the data sets used for the training, validation and verification of the ML model.	The data management includes a data verification step to confirm the appropriateness of the defined ODD and of the data sets used for the training, validation and verification of the ML model. An independent review has confirmed that the defined ODD and of the data sets used for the training, validation and verification of the ML model are appropriate.
C3.1(LM). Learning process management		
Obj.LM-01	Description of the ML model architecture.	The ML model architecture has been described. The ML model architecture has been reviewed and validated by independent means.
Obj.LM-02	The requirements pertaining to the learning management and training processes, including but not limited to: — model family and model selection;	The requirements have been captured. The capturisation has been reviewed and validated by independent means.

	— learning algorithm(s) selection; — explainability capabilities of the selected model; — activation functions; — cost/loss function selection describing the link to the performance metrics; — model bias and variance metrics and acceptable levels (only in supervised learning); — model robustness and stability metrics and acceptable levels; — training environment (hardware and software) identification; — model parameters initialisation strategy; — hyper-parameters and parameters identification and setting; — expected performance with training, validation and test data sets.	
Obj.LM-03	Documentation of the credit sought from the training environment and qualify the environment accordingly.	The document has been completed. The results have been reviewed and validated by independent means.
Obj.LM-04	Quantifiable generalisation bounds.	Quantifiable generalisation bounds have been provided. Quantifiable generalisation bounds have been validated by independent means.
Obj.LM-05	The result of the model training.	The results of the model training have been provided. The result of the model training has been validated by independent means.
Obj.LM-06	Document of any model optimisation that may affect the model behaviour (e.g. pruning, quantisation). Assessment of their impact on the model behaviour or performance.	There is a document of any model optimisation that may affect the model behaviour (e.g. pruning, quantisation). Each document has been validated by independent means. An assessment has been completed of their impact on the model behaviour or performance. The assessment has been validated by independent means.
Obj.LM-07-SL	Assessment of the bias-variance trade-off in the model family selection. Evidence of the reproducibility of the model training process.	The bias-variance trade-off in the model family selection has been assessed and accounted for. The evidence of the reproducibility of the model training process has been

		provided. The evidence has been reviewed and validated by independent means.
Obj.LM-08	Assessment of the estimated bias and variance of the selected model.	The assessment is completed of the estimated bias and variance of the selected model. The assessment has been reviewed and validated by independent means. The assessment has shown that the estimated bias and variance of the selected model meet the associated learning process management requirements.
Obj.LM-09	Evaluation of the performance of the trained model based on the test data set. Documentation of the result of the model verification.	The performance of the trained model based on the test data set has been evaluated and documented. The evaluation has been reviewed and validated by independent means.
Obj.LM-10	Requirements-based verification of the trained model behaviour.	A requirements-based verification of the trained model behaviour has been performed and documented. The requirements-based verification has been reviewed and validated by independent means.
Obj.LM-11	Analysis on the stability of the learning algorithms.	An analysis on the stability of the learning algorithms has been performed and documented. The analysis has been reviewed and validated by independent means.
Obj.LM-12	The verification of the stability of the trained model, covering the whole AI/ML constituent ODD.	The verification of the stability of the trained model has been performed and documented. The verification covers the whole AI/ML constituent ODD. The verification has been reviewed and validated by independent means.
Obj.LM-13	The verification of the robustness of the trained model in adverse conditions.	The verification of the robustness of the trained model in adverse conditions has been performed and documented. The verification has been reviewed and validated by independent means.
Obj.LM-14	The verification of the anticipated generalisation bounds using the test data set.	The verification of the anticipated generalisation bounds using the test data set has been performed and documented. The verification has been

		reviewed and validated by independent means.
Obj.LM-15	The description of the resulting ML model.	The description of the resulting ML model has been captured. The result has been reviewed and validated by independent means.
Obj.LM-16	Justification of completeness of the trained model verification activities.	The justification of completeness of the trained model verification activities has been documented. The result has been validated by independent means. The result confirms that the trained model verification activities are complete.
C3.1(IMP). Model implementation		
Obj.IMP-01	The requirements pertaining to the ML model implementation process.	The requirements pertaining to the ML model implementation process have been captured. The result has been validated by independent means.
Obj.IMP-02	Validation of the model description captured under Objective LM-15. Validation of each of the requirements captured under Objective IMP-01.	The validation of the model description captured under Objective LM-15 has been completed. The validation of each of the requirements captured under Objective IMP-01. The results have been reviewed and validated by independent means.
Obj.IMP-03	Documentation of evidence that all derived requirements generated through the model implementation process have been provided to the (sub)system processes, including the safety (support) assessment.	The documentation of evidence has been completed. The documentation has been reviewed and validated by independent means. The documentation has shown that all derived requirements generated through the model implementation process have been provided to the (sub)system processes, including the safety (support) assessment.
Obj.IMP-04	Assessment of impact of post-training model transformation (conversion, optimisation) on the model behaviour and performance. Identification of the environment (i.e. software tools and hardware) necessary to perform model transformation.	The impact assessment has been completed. The environment has been identified. The results have been reviewed and validated by independent means.
Obj.IMP-05	Appropriate development assurance processes to develop the inference	Appropriate development assurance processes have been planned to develop the inference model into

	model into software and/or hardware items.	software and/or hardware items. Appropriate development assurance processes have been executed to develop the inference model into software and/or hardware items. The results have been reviewed and validated by independent means.
Obj.IMP-06	Assessment of adverse alteration of the defined model properties regarding any transformation (conversion, optimisation, inference model development) performed during the trained model implementation step.	The assessment has been performed and documented. The assessment has been reviewed and validated by independent means. The assessment has shown that any transformation (conversion, optimisation, inference model development) performed during the trained model implementation step has not adversely altered the defined model properties.
Obj.IMP-07	Identification and assessment of possible impact on the inference model behaviour and performance of the differences between the software and hardware of the platform used for model training and those used for the inference model verification.	The assessment has been performed and documented. The assessment has been reviewed and validated by independent means.
Obj.IMP-08	Evaluation of the performance of the inference model based on the test data set and document the result of the model verification.	The evaluation has been performed and documented. The evaluation has been reviewed and validated by independent means.
Obj.IMP-09	Assessment of the stability of the inference model.	The assessment has been performed and documented. The assessment has been reviewed and validated by independent means. The assessment has verified the stability of the interference model.
Obj.IMP-10	Assessment of the robustness of the inference model in adverse conditions.	The assessment has been performed and documented. The assessment has been reviewed and validated by independent means.
Obj.IMP-11	Requirements-based verification of the inference model behaviour when integrated into the AI/ML constituent.	The requirements-based verification has been completed. The assessment has been reviewed and validated by independent means.
Obj.IMP-12	List of the AI/ML constituent verification activities.	The AI/ML constituent verification activities have been checked for completeness. The applicant has

		confirmed that the AI/ML constituent verification activities are complete.
C3.1(CM). Configuration management		
Obj.CM-01	Assessment of application of all configuration management principles to the AI/ML constituent life-cycle data, including but not limited to: — identification of configuration items; — versioning; — baselining; — change control; — reproducibility; — problem reporting; — archiving and retrieval, and retention period.	The assessment has been completed. The assessment has been reviewed and validated by independent means. The assessment has shown that all configuration management principles have been applied to the AI/ML constituent life-cycle data.
C3.1(QA). Quality and process assurance		
Obj.QA-01	Assessment of application of quality/process assurance principles to the development of the AI-based system, with the required independence level.	The assessment has been completed. The assessment has been reviewed and validated by independent means. The assessment has shown that quality/process assurance principles are applied to the development of the AI-based system, with the required independence level.
C3.1(RU). Reuse of AI/ML models		
Obj.RU-01	Impact assessment of the reuse of a trained ML model before incorporating the model into an AI/ML constituent, which considers: — alignment and compatibility of the intended behaviours of the ML models; — alignment and compatibility of the ODDs; — compatibility of the performance of the reused ML model with the performance requirements expected for the new application; — availability of adequate technical documentation (e.g. equivalent documentation depending on the required assurance level); — possible licensing or legal restrictions on the reused ML model (more particularly in the case of COTS ML	The impact assessment has been completed. The impact assessment has been reviewed and validated by independent means.

	models); and — evaluation of the required development level.	
Obj.RU-02	Functional analysis of the COTS ML model to confirm its adequacy to the requirements and architecture of the AI/ML constituent.	The functional analysis of the COTS ML model has been completed. The analysis has been reviewed and validated by independent means. The analysis confirms its adequacy to the requirements and architecture of the AI/ML constituent.
Obj.RU-03	Analysis of the unused functions of the COTS ML model.	The analysis of the unused functions of the COTS ML model has been completed. The analysis has been reviewed and validated by independent means. The deactivation of these unused functions has been prepared.
C3.1(SU). Surrogate modelling		
Obj.SU-01	Assessment of the accuracy and fidelity of the reference model.	The assessment of the accuracy and fidelity of the reference model has been captured. The assessment has been reviewed and validated by independent means. The assessment is shown to support the verification of the accuracy of the surrogate model.
Obj.SU-02	Identification of the additional sources of uncertainties linked with the use of a surrogate model.	The additional sources of uncertainties linked with the use of a surrogate model have been identified and documented. The additional sources of uncertainties linked with the use of a surrogate model have been mitigated. An assessment has shown that the mitigations are effective. The results have been reviewed and validated by independent means.
C3.2(EXP). Development and post-ops AI explainability		
Obj.EXP-01	List of stakeholders, other than end users, that need explainability of the AI-based system at any stage of its life cycle. Roles of these stakeholders. Responsibilities of these stakeholders. Expected expertise of these stakeholders (including assumptions made on the level of training,	List of stakeholders is completed. List of stakeholders has been validated by independent means. Roles of stakeholders have been defined. Roles have been validated by independent means. Responsibilities of stakeholders have been defined. Responsibilities have been validated by independent means.

	qualification and skills).	Expected expertise has been determined. Expected expertise has been validated by independent means.
Obj.EXP-02	Characterisation of the need for explainability for each of the stakeholders (or groups of stakeholders), which is necessary to support the development and learning assurance processes.	The need for explainability for each of the stakeholders (or groups of stakeholders) has been characterised. The results have been reviewed and validated by independent means.
Obj.EXP-03	Identification of the methods at AI/ML item and/or output level satisfying the specified AI explainability needs.	The methods have been identified and documented. The results have been reviewed and validated by independent means.
Obj.EXP-04	Assessment of the AI-based system's ability to deliver an indication of the level of confidence in the AI/ML constituent output, based on actual measurements or on quantification of the level of uncertainty.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system is able to deliver an indication of the level of confidence in the AI/ML constituent output, based on actual measurements or on quantification of the level of uncertainty.
Obj.EXP-05	Assessment of the AI-based system's ability to monitor that its inputs are within the specified ODD boundaries (both in terms of input parameter range and distribution) in which the AI/ML constituent performance is guaranteed.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system is able to monitor that its inputs are within the specified ODD boundaries (both in terms of input parameter range and distribution) in which the AI/ML constituent performance is guaranteed.
Obj.EXP-06	Assessment of the AI-based system's ability to monitor that its outputs are within the specified operational AI/ML constituent performance boundaries.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system is able to monitor that its outputs are within the specified operational AI/ML constituent performance boundaries.
Obj.EXP-07	Assessment of the AI-based system's ability to monitor that the AI/ML constituent outputs (per Objective EXP-04) are within the specified operational level of confidence.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system is able to monitor that the AI/ML constituent outputs (per Objective EXP-

		04) are within the specified operational level of confidence.
Obj.EXP-08	Verification of whether the output of the specified monitoring per the previous three objectives are in the list of data to be recorded per MOC EXP-09-2.	The verification has been completed. The results have been reviewed and validated by independent means. The results show that the output of the specified monitoring per the previous three objectives are in the list of data to be recorded per MOC EXP-09-2.
Obj.EXP-09	Verification of whether the means is provided to record operational data that is necessary to explain, post operations, the behaviour of the AI-based system and its interactions with the end user, as well as the means to retrieve this data.	The verification has been completed. The results have been reviewed and validated by independent means. The results show that the means is provided to record operational data that is necessary to explain, post operations, the behaviour of the AI-based system and its interactions with the end user, as well as the means to retrieve this data.

C4. Human factors for AI		
Objectives	**KPIs**	**Milestones**
C4.1(EXP). AI operational explainability		
Obj.EXP-10	Characterisation of the need for explainability for each output of the AI-based system relevant to task(s) (per Objective CO-02).	The need for explainability has been characterised for each output of the AI-based system relevant to task(s) (per Objective CO-02). The results have been reviewed and validated by independent means.
Obj.EXP-11	Assessment of the explanations presented to the end user by the AI-based system regarding clarity and ambiguity.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system presents explanations to the end user in a clear and unambiguous form.
Obj.EXP-12	Definition of relevant explainability regarding the appropriateness of the decision / action as expected.	The relevant explainability has been defined. The results have been reviewed and validated by independent means. The results have shown that the receiver of the information can use the explanation to assess the appropriateness of the decision / action as expected.

Obj.EXP-13	Definition of the level of abstraction of the explanations, taking into account the characteristics of the task, the situation, the level of expertise of the end user and the general trust given to the system.	The level of abstraction of the explanations has been defined. The results have been reviewed and validated by independent means.
Obj.EXP-14	Assessment of the end user's ability to customise the level of abstraction as part of the operational explainability, where a customisation capability is available.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the end user is able to customise the level of abstraction as part of the operational explainability.
Obj.EXP-15	Definition of the timing when the explainability will be available to the end user taking into account the time criticality of the situation, the needs of the end user, and the operational impact.	The timing has been defined. The results have been reviewed and validated by independent means.
Obj.EXP-16	Assessment of the ability of the end user to get upon request explanation or additional details on the explanation when needed.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the end user is able to get upon request explanation or additional details on the explanation when needed.
Obj.EXP-17	Assessment of the validity of the specified explanation for each output relevant to the task(s).	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the specified explanation for each output relevant to the task(s) is valid.
Obj.EXP-18	Analysis of the training and instructions available for the end user.	The analysis has been completed. The results have been reviewed and validated by independent means. The results show that the training and instructions available for the end user include procedures for handling possible outputs of the ODD monitoring and output confidence monitoring.
Obj.EXP-19	Analysis of the information provided to the end user concerning unsafe AI-based system operating conditions.	The analysis has been completed. The results have been reviewed and validated by independent means. The results show that the information concerning unsafe AI-based system

		operating conditions provided to the end user enables them to take appropriate corrective action in a timely manner.
C4.2(HF). Human-AI teaming		
Obj.HF-01	Assessment of the ability of the AI-based system design to build its own individual situation representation.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system designed is able to build its own individual situation representation.
Obj.HF-02	Assessment of the ability of the AI-based system design to reinforce the end-user individual situation awareness.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system designed is able to reinforce the end-user's individual situation awareness.
Obj.HF-03	Assessment of the ability of the AI-based system design to enable and support a shared situation awareness.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system designed is able to enable and support a shared situation awareness.
Obj.HF-04	Assessment of the ability of the AI-based system design to request a cross-check validation from the end user, if a decision is taken by the AI-based system that requires validation based on procedures.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system designed is able to request a cross-check validation from the end user, if a decision is taken by the AI-based system that requires validation based on procedures.
Obj.HF-05	Assessment of the ability of the AI-based system design to identify a suboptimal strategy and propose through argumentation an improved solution, for complex situations under normal operations.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system designed is able to identify a suboptimal strategy and propose through argumentation an improved solution, for complex situations under normal operations.
Cor.Obj.HF-05	Assessment of the ability of the AI-based system design to process and act	The assessment has been completed. The results have been reviewed and

	upon a proposal rejection from the end user.	validated by independent means. The results show that the AI-based system designed is able to process and act upon a proposal rejection from the end user.
Obj.HF-06	Assessment of the ability of the AI-based system design to identify the problem, share the diagnosis including the root cause, the resolution strategy and the anticipated operational consequences, for complex situations under abnormal operations.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system designed is able to identify the problem, share the diagnosis including the root cause, the resolution strategy and the anticipated operational consequences, for complex situations under abnormal operations.
Cor.Obj.HF-06	Assessment of the ability of the AI-based system design to process and act upon arguments shared by the end user.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system designed is able to process and act upon arguments shared by the end user.
Obj.HF-07	Assessment of the ability of the AI-based system design to detect poor decision-making by the end user in a time-critical situation, alert and assist the end user.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system designed is able to detect poor decision-making by the end user in a time-critical situation, alert and assist the end user.
Obj.HF-08	Assessment of the ability of the AI-based system design to propose alternative solutions and support its positions.	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system designed is able to propose alternative solutions and support its positions.
Obj.HF-09	Assessment of the ability of the AI-based system design to modify and/or to accept the modification of task allocation pattern (instantaneous/short-term).	The assessment has been completed. The results have been reviewed and validated by independent means. The results show that the AI-based system designed is able to modify and/or to accept the modification of task allocation pattern (instantaneous/short-term).

C4.3(HF). Modality of interaction and style of interface		
Obj.HF-10	KPI identification omitted for now: Not applicable to Use Cases.	Milestone identification omitted.
Obj.HF-11	KPI identification omitted for now: Not applicable to Use Cases.	Milestone identification omitted.
Obj.HF-12	KPI identification omitted for now: Not applicable to Use Cases.	Milestone identification omitted.
Obj.HF-13	KPI identification omitted for now: Not applicable to Use Cases.	Milestone identification omitted.
Obj.HF-14	KPI identification omitted for now: Not applicable to Use Cases.	Milestone identification omitted.
Obj.HF-15	KPI identification omitted for now: Not applicable to Use Cases.	Milestone identification omitted.
Obj.HF-16	KPI identification omitted for now: Not applicable to Use Cases.	Milestone identification omitted.
Obj.HF-17	KPI identification omitted for now: Not applicable to Use Cases.	Milestone identification omitted.
Obj.HF-18	KPI identification omitted for now: Not applicable to Use Cases.	Milestone identification omitted.
Obj.HF-19	KPI identification omitted for now: Not applicable to Use Cases.	Milestone identification omitted.
Obj.HF-20	KPI identification omitted for now: Not applicable to Use Cases.	Milestone identification omitted.
Obj.HF-21	KPI identification omitted for now: Not applicable to Use Cases.	Milestone identification omitted.
Obj.HF-22	KPI identification omitted for now: Not applicable to Use Cases.	Milestone identification omitted.
Obj.HF-23	KPI identification omitted for now: Not applicable to Use Cases.	Milestone identification omitted.
Obj.HF-24	An assessment of the ability to combine or adapt the interaction modalities depending on the characteristics of the task, the operational event and/or the operational environment.	The design of the AI-based system has been assessed regarding the ability to combine or adapt the interaction modalities depending on the characteristics of the task, the operational event and/or the operational environment. The assessment has been reviewed by independent means. The assessment has shown that the AI-based system

		design has sufficient ability to combine or adapt the interaction modalities depending on the characteristics of the task, the operational event and/or the operational environment.
Obj.HF-25	An assessment of the ability to automatically adapt the modality of interaction to the end-user states, the situation, the context and/or the perceived end user's preferences.	The design of the AI-based system has been assessed regarding the ability to automatically adapt the modality of interaction to the end-user states, the situation, the context and/or the perceived end user's preferences. The assessment has been reviewed by independent means. The assessment has shown that the AI-based system design has sufficient ability to automatically adapt the modality of interaction to the end-user states, the situation, the context and/or the perceived end user's preferences.
C4.4(HF). Error management		
Obj.HF-26	An assessment of the likelihood of design-related end-user errors in the design of the AI-based system.	The design of the AI-based system has been assessed regarding the likelihood of design-related end-user errors. The assessment has been reviewed by independent means. The assessment has shown that the likelihood of design-related end-user errors has been minimised.
Obj.HF-27	An assessment of the likelihood of HAIRM-related errors in the design of the AI-based system.	The design of the AI-based system has been assessed regarding the likelihood of HAIRM-related errors. The assessment has been reviewed by independent means. The assessment has shown that the likelihood of HAIRM-related errors has been minimised.
Obj.HF-28	An assessment of the tolerance to end-user errors in the design of the AI-based system.	The design of the AI-based system has been assessed regarding the tolerance to end-user errors. The assessment has been reviewed by independent means. The assessment has shown that the AI-based system is tolerant to end user errors.

Obj.HF-29	An assessment of the opportunities to detect errors by end user interacting with the AI-based system.	The design of the AI-based system has been assessed regarding opportunities to detect errors by end user interacting with the AI-based system. The assessment has been reviewed by independent means. The assessment has shown that the AI-based system design has sufficient opportunities to detect the error.
Obj.HF-30	An assessment of the means to inform the end user interacting with the AI-based system that an error has been detected.	The design of the AI-based system has been assessed regarding means to inform the end user interacting with the AI-based system that an error has been detected. The assessment has been reviewed by independent means. The assessment has shown that the AI-based system design has sufficient means to inform the end user interacting with the AI-based system that an error has been detected.
C4.5(HF). Failure management		
Obj.HF-31	An assessment of the ability to diagnose the failure and present the pertinent information to the end user.	The design of the AI-based system has been assessed regarding the ability to diagnose the failure and present the pertinent information to the end user. The assessment has been reviewed by independent means. The assessment has shown that the AI-based system design has sufficient ability to diagnose the failure and present the pertinent information to the end user.
Obj.HF-32	An assessment of the ability to propose a solution to the failure to the end user.	The design of the AI-based system has been assessed regarding the ability to propose a solution to the failure to the end user. The assessment has been reviewed by independent means. The assessment has shown that the AI-based system design has sufficient ability to propose a solution to the failure to the end user.
Obj.HF-33	An assessment of the ability to support the end user in the implementation of the solution.	The design of the AI-based system has been assessed regarding the ability to support the end user in the implementation of the solution. The

		assessment has been reviewed by independent means. The assessment has shown that the AI-based system design has sufficient ability to support the end user in the implementation of the solution.
Obj.HF-34	An assessment of the provision to the end user of the information that logs of system failures are kept for subsequent analysis.	The design of the AI-based system has been assessed regarding the provision to the end user of the information that logs of system failures are kept for subsequent analysis. The assessment has been reviewed by independent means. The assessment has shown that the AI-based system design accounts for the provision to the end user of the information that logs of system failures are kept for subsequent analysis.
C5. AI safety risk mitigation		
Objectives	**KPIs**	**Milestones**
C5(SRM). AI safety risk mitigation concept and top-level objectives		
Obj.SRM-01	Assessment of the coverage of the objectives associated with the explainability and learning assurance building blocks. Assessment of the need for an additional dedicated layer of protection to mitigate the residual risks to an acceptable level.	Both assessments have been completed. Both assessments have been reviewed by independent means.
Obj.SRM-02	Safety risk mitigation means as identified in Objective SRM-01.	Safety risk mitigation means as identified in Objective SRM-01 have been established. The results are reviewed and validated by independent means.
C6. Organisations		
Objectives	**KPIs**	**Milestones**
C6.1(ORG). High level provisions and anticipated AMC		
Prov.ORG-01	Review of organisation's processes.	The organisation has reviewed its processes. The organisation has adapted its processes to the introduction of AI technology. The

		results have been validated by independent means.
Prov.ORG-02	The information security risks related to the design, production and operation phases of an AI/ML application.	The information security risks related to the design, production and operation phases of an AI/ML application have been identified. The information security risks are continuously assessed. The assessment is reviewed by independent means.
Prov.ORG-03	A data-driven 'AI continuous safety assessment' process based on operational data and in-service events.	A data-driven 'AI continuous safety assessment' process has been implemented based on operational data and in-service events. The process is regularly evaluated and reviewed by independent means.
Prov.ORG-04	Processes to continuously assess ethics-based aspects for the trustworthiness of an AI-based system with the same scope as for Objective ET-01.	The organisation has established processes to continuously assess ethics-based aspects for the trustworthiness of an AI-based system with the same scope as for Objective ET-01. The processes are regularly evaluated and reviewed by independent means.
Prov.ORG-05	The specificities of AI, including interaction with all relevant stakeholders, as accommodated in the continuous risk management process.	The continuous risk management process is regularly adapted to accommodate the specificities of AI, including interaction with all relevant stakeholders. The process is regularly evaluated and reviewed by independent means.
Prov.ORG-06	Auditability of the safety-related AI-based systems.	An assessment has shown that the safety-related AI-based systems are auditable by internal and external parties, including especially the approving authorities. The assessment has been reviewed and validated by independent means.
C6.2(ORG). Competence considerations		
Prov.ORG-07	The specificities of AI, including interaction with all relevant stakeholders, as accommodated in the training processes.	The training processes are regularly adapted to accommodate the specificities of AI, including interaction with all relevant stakeholders. The processes are regularly evaluated and reviewed by independent means.

Prov.ORG-08	The specificities of AI, including interaction with all relevant stakeholders, as accounted for in end users' licensing and certificates.	The end users' licensing and certificates are regularly adapted to account for the specificities of AI, including interaction with all relevant stakeholders. The processes are regularly evaluated and reviewed by independent means.

Table 12: KPIs and Milestones for EASA Objectives