Ref. Ares(2025)5103535 - 26/06/2025

SafeTeam

D5.2: Digital Assistant Implementation and Regulatory Considerations

Document number	D5.2
Document title	Digital Assistant Implementation and Regulatory Considerations
Version	2.1
Work package	WP5
Edition date	23.06.2025
Responsible unit	CAAi
Dissemination level	PU
Project acronym	SafeTeam
Grant	101069877
Call	Safe, Resilient Transport and Smart Mobility services for passengers and goods (HORIZON-CL5-2021-D6-01)
Topic	HORIZON-CL5-2021-D6-01-13: Safe automation and human factors in aviation – intelligent integration and assistance

This project has been co-funded by the European Union under Grant Agreement 101069877

Co-funded by the European Union

© SafeTeam Consortium.

SafeTeam Consortium

innaxis	Innaxis (INX)
AESA AGENCIA ESTATAL DE SEGURIDAD AÉREA	Agencia Estatal de Seguridad Aérea (AESA)
TUM	Technische Universität München (TUM)
DataBeacon	DataBeacon
ONERA THE FRENCH AEROSPACE LAB	ONERA
RI. SE	Rise Research Institutes of Sweden AB (RISE)
PEGASUS AIRLINES	PEGASUS HAVA TASIMACILIGI ANONIM SIRKETI (PEGASUS)
UK Civil Aviation Authority International	CAA INTERNATIONAL LIMITED (CAAi)

Document change record

Version	Date	Status
0.1	08/05/2025	Draft Submission
1.0	16/06/2025	Amendments Addressing Consortium Feedback
2.0	23/06/2025	Amendments Addressing Technical Univeristy of Munich Comments
2.1	23/06/2025	Final version approved for submission

Table of Contents

1 Scope

The aviation industry is increasingly exploring the potential of machine learning (ML) and concepts of Artificial Intelligence (AI) to enhance flight safety. Use Cases such as predicting unstable approaches during aircraft landings, digital assistance for area ATC and automatic risk assessment for drones are concepts being explored.

Project SafeTeam is a consortium of organisations across Europe working together to develop these Use Cases as early adopters to understand the challenges of implementation for more automated digital assistance. SafeTeam also progresses on the human factor's aspects on the use of digital assistants for aviation, including a deeper understanding on the technology and processes that will facilitate the adoption of more autonomous tools with integration into operations and enhancing human cognitive abilities. The project investigates approval and certification issues, concentrating on aspects related to the human ability to operate more sophisticated automated tools to develop the explainability of larger concepts of AI operations.

The objectives of SafeTeam in the development of these Use Cases aim to address the following objectives:

- Facilitate a human-centric approach to automation and its integration into a wide spectrum of air traffic operations.
- Propose methodologies for the assessment and monitoring of the system performance, with special focus on safety and resilience, to enable seamless Human-Machine cooperation.
- Progress in the development of Digital Assistants for aviation operations in support of human performance for all development, testing, validation and verification phases.
- Define understanding of the definition of the potential regulatory and certification requirements for automation tools to address safety critical elements.

The objective of Work Package 5 (WP5) was to increase the regulatory preparedness of the digital assistants by researching their successful certification as AI-based products, including predictive capabilities. The probabilistic nature of the predictions will trigger the introduction of new liability models, regulations, processes, and risk assessment methodologies.

To commence work on WP5, an initial kick off workshop was held virtually on the 3rd of February 2023. Eight people from the Consortium (CAAi, AESA and INNAXIS) and three representatives from the European Union Aviation Safety Agency (EASA) took part in the event. It focused on the planned scope of WP5 including a discussion of the proposed three Use Case studies and the input required from the European aviation regulators in the emerging technology.

The three Use Cases were selected as applications of digital assistants, each of which is centered around a specific solution relying on artificial intelligence techniques. Each digital assistant was carefully selected based on the relevance to the study, considering airborne and ground operations, as well as the suitability for introducing higher levels of automation, impact, benefits and potential to bring the solution to market.

Coordination of WP5 between CAAi, AESA, and the consortium partners was conducted online via regular conversations and correspondence, supported by a formal review meeting between CAAi, Onera, and the Technische Universität München (TUM), held virtually on the 18th October 2024 and a meeting held on the 13th February 2025 at the premises of TUM. A further review meeting between CAAi, Onera, and the Rise Research Institutes of Sweden AB (RISE), was held virtually on the 10th April 2025.

1.1 Case studies review approach

The review of the three case studies was conducted through a combination of direct engagement, structured discussion, and collaborative analysis with both UK CAA and the SafeTeam Consortium.

For Use Case 1, the UK CAA undertook a virtual workshop with Innaxis and ONERA to explore their proposed operating model, the underlying concept, and associated challenges. This session enabled a method to formulate initial regulatory questions, particularly in relation to air route optimisation and the potential complexities arising from airspace modernisation. These discussions informed the identification of regulatory considerations and supported the development of an initial regulatory response.

For Use Case 2, the UK CAA further engaged in an in-person session with TUM at their facilities which included a simulator demonstration. During this session, the UK CAA and TUM worked through a range of test scenarios, datasets and, modelling frameworks, using them to explore practical challenges and operational complexities for the unstable approach tool embedded to the system. This hands-on approach provided valuable insights and deepened our understanding of the regulatory implications specific to how maturity and reliability of such technologies should be initially addressed.

For Use Case 3, the research and formal response were conducted by AESA, the Spanish civil aviation authority, representing the European perspective. Given that this use case draws on the EASA SORA framework and is primarily applicable within the European regulatory environment, it was agreed that AESA would lead on the regulatory assessment.

The UK CAA and AESA collaborated throughout the development of all three use cases, contributing to a harmonised and coordinated approach to the overall project. The regulatory assessment has run in parallel to the digital assistants developments and its respective validation exercises. Accordingly, some of the recommendations provided have been incorporated to the developments. The final results of these exercises and the latest developments of the digital assistants are reported in D4.1, D4.2 and D4.3 for each of the use cases.

2 Use Case 1 Analysis – Digital Assistance for Area ATC (en-route)

Use Case 1 presents a digital system to enhance decision making processes for Air Traffic Control (ATC) when in the scenario of airborne interactions en-route. Development of such a tool aimed to incorporate human factors expertise according to the targeted level of autonomy and corresponding controller-machine tasks distribution. Determination of safety critical levels are proposed to be on a case-by-case basis depending on the complexity of the Use Case. The consortium aimed to carry out validation exercises to monitor ATC operators on various human factors with and without the tool i.e., workload, situational awareness and safety indicators that may determine the ability or competency of the operator.

2.1 Key Considerations

In WP5.2, the UK CAA presents considerations regarding the Use Case based on the current level of maturity reach by the consortium. From a regulatory perspective, these considerations include ensuring the technology meets has accounted for safety and reliability standards that would be required by an aviation regulator, addressing potential cybersecurity themes, and achieving seamless integration with existing ATC systems. Additionally, considerations made for Use Case 1 must also account for the complexities of unlocking routes to early certifications, which are critical for the lawful deployment and operation of the technology. This assessment was developed in parallel to the results final development and validation exercises. This gave SafeTeam Digital Assistants the opportunity to incorporate some upgrades to mitigate some of the regulatory issues raised.

CONSIDERATIONS	EXPOSITION
REGULATORY APPROVAL FOR DECISION MAKING TOOLS	NAA's currently have no dedicated certification pathway specifically for advisory-only complexity forecasting tools in the pre-tactical phase. As responsibility of decision making currently is outlined in operations manuals with nominated personnel, a line of accountability is already shown. This does create skewed boundary where operational deployment could face delays unless risk-based approaches and equivalence arguments are clearly articulated. Certification is more of a priority for NAA's once the decision-making tool exceeds advisory and becomes primary in safety critical elements of operations.
AUTOMATION BIAS AND OVER-RELIANCE	As expressed in SafeTeam: D3.1 , repeated use of advisory tools by operators, especially under high workload environments may lead to automation bias. There needs to be further examination of formal monitoring mechanisms for human-tool interactions. Consortium can further develop this understanding by looking into such models such as the NASA Task Load Index which may enhance understanding of unintended human factors risks.
UNCERTAINTY LEVELS AND TRANSPARENCY	From the review of SafeTeam's maturity level of Use Case 1, there is no mechanism to monitor model drift. This means forecast reliability could degrade without detection, especially in dynamic or more complex regions of airspace. Furthermore, the system seems to not present confidence intervals or metrics to determine levels of uncertainty. This can limit transparency and controller trust, making a tool such as this difficult to mature out of advisory phase of testing.
ATC HARMONISATION	As the automation tool is developed independently, there is a lack of formal understanding on shared accountability between tool developers, ANSPs, and operational ATCs. This introduces a liability gap. If a decision informed by a digital assistant contributes to an incident, clear responsibility boundaries are not yet defined. From a regulatory perspective,

	currently liability falls on the nominated ATC who make decisions on conflicts; this may yet change on how the assistant plays a part in contributing to the decision taken.
COMPLEXITY ACROSS FLIGHT INFORMATION REGIONS (FIRS)	The system does not currently show or account for variability in procedures across regions (e.g., London vs. Scottish FIRs), nor does it differentiate for varying levels of controller experience. This could present a barrier to adoption at an equal pace across ATC and safe scaling across operational environments as more usage is introduced.
CYBERSECURITY	Although the tool runs in a shadow environment and has restricted access, its future integration into operational environments or live data streams will necessitate robust cyber assurance. This includes protection against adversarial manipulation, data integrity monitoring, and resilience against hallucinated forecasts that could mislead pre-tactical planning.
COHESION TO FUTURE ATM CONCEPTS	While the tool indicates scalability to support DAM, UTM, and 4D trajectory management, its integration with these systems is not yet validated. Alignment with NAA strategies such as the UK CAA Airspace Modernisation concept may require further review. There is also a need to review the international interoperability standards and cross-border safety case coordination for commercial operations.

2.2 Regulatory Frameworks and Applicability

REGULATORY THEMES	EXPOSITION
GENERAL STANDARDS AND RECOMMENDED PRACTICES (SARPS)	A general review of how to integrate digital tools for enroute optimisation can be found in Regulation (EU) 2017/373 . This regulation governs the provision of air navigation services (ANS) and sets out the guidance for: Safety management systems (SMS) for ANSPs. Human factors integration and system usability. Software assurance and risk classification. Interoperability and systems safety assurance. Currently the regulation shows that if the technology was to be deployed operationally, the Air Navigation Service Provider (ANSP) should be responsible for validating and approving the assistant under their existing SMS.
HUMAN FACTORS UK REGULATION	UK CAA guidance under CAP 670 (ATS Safety Requirements) and CAP 1377 (ATM Automation: Guidance on Human-Technology Integration) should be referenced for: Defining human-in-the-loop validation procedures. Evaluating controller workload, trust, and interface design Preventing automation bias or over-reliance.

PERFOMANCE BASED OVERSIGHT (PBO) AND SAFETY PERFOMANCE INDICATORS (SPI).	SafeTeam's assistant contributes to safety by potentially reducing controller overload through advisory tactical interventions. From a regulatory standpoint, the tool's focuses on SPIs such as: • Number of controller interventions, • Conflict detection resolution time • Sector overload Under these circumstances, the UK CAA encourages the use of the PBO concept where tool-derived SPIs are continuously monitored and fed into key safety reviews. These reviews not only highlight safety critical challenges to adopt but, show a mature development of safety critical thinking of users and developers. This can provide assurance for fair and responsible use of digital assistance in more complex scenarios.
CYBERSECURITY STANDARDS AND ASSURANCE	All digital tools must consider alignment of cyber procedures for greater assurance. There is guidance on how cybersecurity oversight is enforced into aviation in CAP 1753. Guidance of nominated cyber roles and rules of engagement for Aerodromes will tackle the oversight required for more digital assistants implemented in the sector.
AIRSPACE CHANGE PROCESS	Given the tools architecture and compatibility with DAM and 4D trajectory, future maturity must also be able to adapt to airspace modernisation changes, The outlined procedures of CAP 1616 shows how changes to airspace structures should be taken into account for how well digital assistance can adapt to dynamic complexities within the airspace.

3 Use Case 2 Analysis – Unstable Approach Prediction for Aircraft Flight Decks

3.1 Key Findings

This case study aims at providing predictions to the crew in the cockpit when there is high probability that the flight becomes unstable.

The design principles, distribution of automation roles, and validation exercises in SafeTeam proposes that the use case can achieve TRL6 (see Annex I). Meanwhile, the performance during these exercises will be monitored and assessed for human factors considerations. Aspects like the concrete threshold, or how much heads up needs to be given to crew for predictions to be useful, will be defined in collaboration with operational experts (pilots and safety managers) and human factors professionals, as well as the appropriate interface defining how the information needs to be showed.

The ML model under review was developed to predict unstable approaches using flight data monitoring (FDM) data from airlines such as Iberia and Pegasus. The model focuses on the Airbus A320 fleet and incorporates data such as engine performance, airspeed, pressure, and gearbox metrics, alongside equations of motion. Sensor models specific to the A320 were also integrated. All aircraft systems are implanted in Simulink, while the ML model is a python implementation. The interface between the components is managed using UDP. During each approach, aircraft performance data simulated by the aircraft is collected from 10-4 miles to the runway threshold. Backward computing techniques were employed to model predictions, and the system provides textual notifications to pilots when an unstable approach is predicted.

Unstable approaches are a significant contributor to runway incidents, and their early detection could mitigate risks. This deliverable provides an in-depth analysis of a use case where ML can be applied to predict unstable approaches, based on a review of the technology and its current state. The high-level analysis concludes that while the technology shows promise, it remains in its discovery phase, with significant gaps in operational applicability, human factors, and certification readiness. Additionally, the use of bowtie appropriate risk assessments in this context still; requires a better understanding.

3.2 Operational Limitations:

The model was developed based on a limited operational design concept, with data from specific airports and airlines. However, it is important to note that while data is limited, approaches of aircraft are similar such as majority of airports using a 3-degree glide slope. This raises questions about its applicability to broader operational environments as operational complexities of different airlines and airports may heavily skew results. A broader data pool from more world regions could potentially scale up the accuracy of the system. The operational design concept at the time of development is narrow, focusing on a subset of the A320 fleet. For project SafeTeam, the model only focusses on the A320 fleet, therefore any further applications beyond this would requires further research about the model's applicability to other aircraft types, airlines, and airports with different operational procedures. For instance, approach speeds vary significantly depending on aircraft weight, weather conditions, and airport elevation. The model's current dataset does not account for these variables comprehensively, necessitating further research to enhance its robustness should there be a further demand to expand to more aircraft types.

3.3 Human Factors:

The ML system provides textual notifications to pilots when an unstable approach is predicted. However, it does not offer recommendations or corrective actions, limiting its utility. Pilots in focus groups have expressed that they typically recognise unstable approaches and know how to respond, suggesting that the system may not be necessary. This raises questions about the system's value proposition and whether it complements or conflicts with pilot decision-making. From initial focus groups with pilots, it is indicated that the unstable approaches are typically known to pilots without assistant tools, with corrective actions easily actioned. This suggests an assistant tool for unstable predictions would be a nice to have but would not contribute to initial pilot decisions. Any adjustments required to the developed tool from discussion with TUM indicate changes would be more focused on corrective actions e.g., speed brakes, gears and flap configurations as this is a functionality the ML model does not include. Conclusions to work package 4.2 will provide a furthermore insights to the potential outcomes of pilots once a comprehensive simulation exercise is conducted.

Human factors play a critical role in the system's effectiveness. For instance, pilot confidence in the system's predictions is crucial for its adoption. If pilots perceive the system as unreliable or unnecessary, they may disregard its notifications, rendering it ineffective. Furthermore, the system's impact on pilot workload and decision-making under pressure requires thorough investigation. Simulator studies involving pilots from diverse backgrounds and experience levels are needed to assess these factors.

From a regulatory perspective, the UK CAA would prefer tools considering human factors elements to follow the ICAO five human performance principles. There principles outlined in the ICAO Doc 10151 state:

- Principle 1: People's performance is shaped by their capabilities and limitations.
- Principle 2: People interpret situations differently and perform in ways that make sense to them.
- Principle 3: People adapt to meet the demands of a complex and dynamic work environment.
- Principle 4: People assess risks and make trade-offs.
- Principle 5: People's performance is influenced by working with other people, technology, and the environment.

In the context of SafeTeam, further studies are required to be done against the ICAO principles. At the current time, due to the systems level of maturity, more development on its use case would need to be defined before this exercise is conducted.

4 Use Case 3 Analysis – Drone Risk Assessment Digital Assistant

4.1 Key Findings

Use Case 3 explores a Digital Assistant for pre-tactical UAS air risk assessment, designed to support operators conducting missions in the 'Specific' category under the SORA methodology. It provides quantitative air risk predictions, enabling operators to move beyond qualitative SORA estimations by calculating mid-air collision (MAC) probabilities using historical air traffic data.

The tool ingests data from cooperative surveillance sources (ADS-B, FLARM, Mode S) and generates Air Risk Maps, identifying areas of high or uncertain air traffic density. It enables mission planners to adjust trajectories, timing, or frequency of operations to ensure compliance with Target Level of Safety (TLOS) thresholds defined by SORA methodology.

This use case contributes to evolving regulatory practices by proposing a data-driven method to characterize the airspace and, in the future, justify TMPR, which are currently at the discretion of the applicant in the SORA framework.

4.2 Operational Limitations

Operational limitations are the boundaries within which the tool must operate to maintain safety and compliance with regulations or technical standards. These limitations, usually referred as operational domain or operational design domain, are related with performance and robustness.

Dependent on the Availability of Data

The system's effectiveness depends on the availability and accuracy of surveillance data. Airspace areas not covered by ADS-B/FLARM receivers result in "blue cells" (unknown risk), limiting confidence and operational coverage.

No detection of non-cooperative aircraft

The tool does not account for untracked aircraft (e.g., without transponders), which poses inherent limitations in dense GA environments. As a result, the tool has a limited application to airspace with co-operative aircraft, given that transponders already are active for geolocation purposes.

Geographic specificity:

The models are based on datasets from France (e.g., Marseille–Avignon), and as a result its generalization may require retraining or validation for deployment in other airspaces.

No tactical integration yet:

The tool currently supports pre-tactical planning only; real-time risk monitoring and U-space interoperability are not yet implemented.

4.3 Human Factors

The Digital Assistant includes explanatory modules that help users understand risk assessments, such as MAC probabilities, occupancy rates, and proposed alternative routes. The next development step involves user trials to assess trust, interpretability, and cognitive workload, consistent with the ICAO Human Performance principles (Doc 10151) and EASA's human-centric automation goals as well as OSOs 19/20 of the SORA methodology.

A key concern is avoiding automation over-reliance, particularly when operators interpret visual risk maps (e.g., misreading low-data zones as low-risk). Explainability of outcomes and transparent communication of system limitations (e.g., "unknown risk" areas) are also crucial.

4.4 EU Regulatory Framework and Applicability

Regulatory Area	Relevance to Use Case 3
Implementing Regulation (EU) 2019/947	The tool is intended for operations in the 'Specific' category and supports SORA-based risk assessments, particularly for Air Risk Class (ARC) justification.
JARUS SORA (v2.5 and transition to 3.0)	Enables quantitative air risk estimation to justify ARC reductions, in alignment with SORA 2.5 guidance. Closely aligned with ongoing development of SORA 3.0, which will formalise quantitative air risk methodologies.
Acceptable Means of Compliance (AMC) & Guidance Material (GM)	Currently, ARC reductions are accepted if properly justified. This tool contributes a repeatable, data-based methodology that could form the basis for future AMC for ARC downgrading.
Digital Assistants and AI in Aviation	Although not yet regulated as a standalone tool, the assistant could fall under software assurance requirements (e.g., EASA Part-UAS, Part-ATM/ANS for ground systems if integrated with U-space).
Human Factors Requirements	The tool's future certification or acceptance should demonstrate compliance with HF integration, per EASA Rulebook, including design for transparency, workload minimisation, and trust-building.
U-space and UTM Integration (Regulation (EU) 2021/664)	Future alignment with U-space services (especially Strategic Deconfliction and Conformance Monitoring) would enhance operational value and compliance.
EU Artificial Intelligence Act (Regulation (EU) 2024/1689)	This tool qualifies as a high-risk AI system under the EU AI Act, as it supports safety-critical decision-making in aviation. Compliance must include: risk management, data governance, transparency, human oversight, and post-market monitoring. Alignment with Annex III (critical infrastructure/transport) and Annex IV (technical documentation) is required prior to large-scale deployment.

In addition, some efforts of addressing how to apply these regulatory requirements to real applications have been carried out by EASA through the EASA AI Concept Paper (Issue 2) and the Machine Learning Application Approval (MLEAP).

4.5 EU Regulatory Gaps and Proposed Developments

To enable the widespread operational adoption of this Digital Assistant within the EU regulatory ecosystem, the following regulatory evolutions and enablers are recommended:

Define AMC/GM for Quantitative ARC Justification

EASA should develop an acceptable methodology for ARC reduction using historical traffic data, specifying acceptable data quality, observation windows, and statistical thresholds for MAC probability.

Minimum Data Quality Requirements

Establish clear minimum coverage and resolution criteria (e.g., FLARM receiver density, cell observation hours) to ensure reliability of "low risk" classification.

Harmonised Classification of Aircraft Densities

Create a harmonised mapping between aircraft density and risk categories, analogous to the population density thresholds used for Ground Risk assessment in SORA.

Formal Validation Pathway

Encourage integration of the tool into validation sandboxes under EASA's Innovation Partnership Programme or JARUS experimental working groups.

U-space Integration Pathway

Define how such forecasting tools can feed into U-space Common Information Services (CIS) or be recognised as a Strategic Air Risk Information Provider under Regulation (EU) 2021/664.

Liability and Role Definition

Clarify roles and responsibilities in cases where flight plans developed with support of digital tools result in incidents. The operator retains accountability today, but shared risk models may evolve.

Certification or Declaration Scheme

Consider developing a lightweight approval process for digital safety support tools.

Alignment with the EU Artificial Intelligence Act

As this tool performs safety-related risk forecasting in a regulated domain (aviation), it likely falls under the high-risk category in the EU AI Act (2024). Therefore, future deployments must:

1. Implement a robust risk management system.
2. Ensure transparent outputs, documentation of the logic behind them, and visibility of uncertainty or coverage limitations.
3. Define a human oversight strategy, particularly for "human-in-the-loop" decision models.
4. Establish a post-market monitoring plan and incident reporting pipeline.
5. Maintain full technical documentation.

5 AI Strategy and Certification Considerations

Certifying ML systems for aviation use presents unique challenges. Unlike traditional deterministic systems, ML models operate probabilistically, making it difficult to establish clear cause-and-effect relationships. This complicates the certification process, as aviation regulators must ensure that the system's predictions are reliable and safe under all operational conditions.

For example, the UK CAA AI strategy emphasises the need for trustworthy ML systems, however, building trust in this context is challenging. One approach is to compare pilot self-assessments of approach stability with ML predictions to identify discrepancies and improve the model's accuracy. Additionally, establishing a framework for ongoing monitoring and validation of the system's performance is essential to maintain trustworthiness. CAP 2970 presents the UK CAA stance on building trust in AI. Building trust follows developing ML systems with scrutiny on five topics a system should be considering.

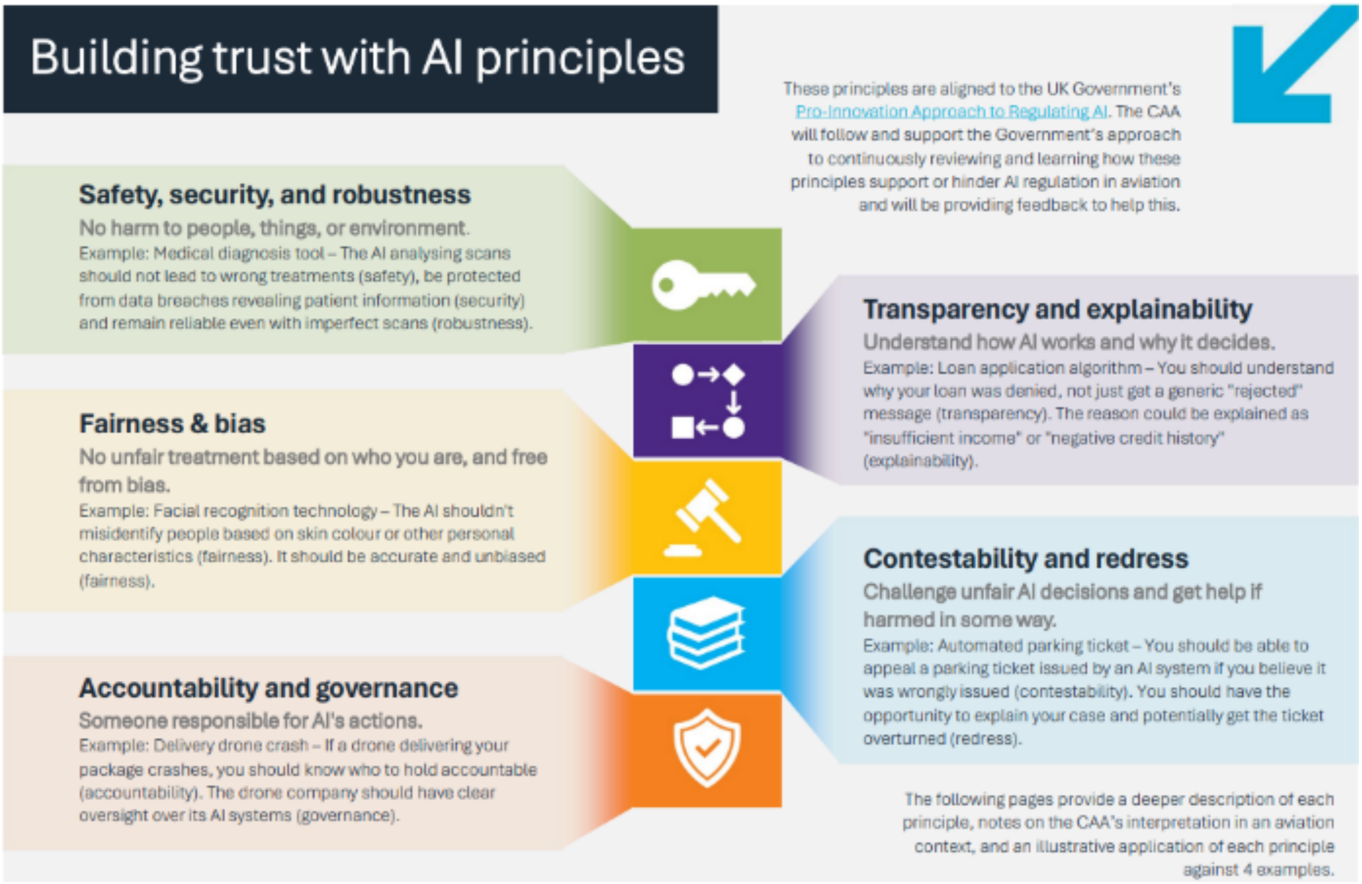

Figure 1: UK AI Strategy Trust Principles

5.1 Safety, Security and Robustness

Applications of AI should function appropriately in a secure, safe, and robust way in normal and foreseeable use, and in cases of misuse or other adverse conditions. Risks should be identified, assessed, and carefully managed, with an ability to analyse the system's lifecycle in response to an inquiry. Aviation already has a strong safety-first culture. Safety Management Systems are a systematic and proactive approach to managing safety risks. Introducing AI and automated systems to an aviation system will bring about new risks that will need to be captured by the Safety Management System (SMS). The ability to adequately describe the safety, security, and robustness at all stages of the lifecycle is linked closely with the transparency and explainability of the system. Safety, security, and robustness require assurance – clearly defined target levels of safety, security, and robustness, and methods to demonstrate that the system always maintains those levels.

5.2 Transparency and Explainability

Organisations and individuals developing and deploying AI should clearly communicate when, how, and why it is used, and explain the system's decision-making process in an appropriate level of detail and timeliness that matches the risks posed by it. It should also be transparent to a human such that those decisions and outcomes can be traced and explained.

The key point to note here is the proportionality to risk. The degree of transparency and explainability is dependent on the complexity of the software. For example, machine learning techniques can develop software that is incomprehensible to an experienced software engineer. It may be proposed that if an aviation system which has a high degree of risk associated to it is not sufficiently explainable or transparent, it will not be approved for use. The level of acceptance will adapt with technology and skills maturity, as they develop to an extent where complex machine-learnt systems can be explained through novel means.

5.3 Fairness and Bias

AI should be created, deployed, and maintained in a way which complies with applicable regulations and laws, and must not discriminate against individuals or organisations, or somehow create unfair commercial outcomes.

Creation, deployment, and maintenance is intended to describe all possible stages of an AI system's lifecycle. Applicable regulations and laws are dependent on the context. These may range from technical regulations such as Air Traffic Management / Air Navigation Services (ATM/ANS) Regulations, through to the General Data Protection Regulation (GDPR). It is intended that a system is updated throughout its life to reflect the requirements of any applicable laws and regulations. This is like any other type of system that must remain compliant with evolving laws and regulations.

With the context of SafeTeam developments, more automated systems must comply with the Air Traffic Management (ATM) / ANSP regulations, as well as operating appropriately within the context of the UK/ European Rules of the Air. It operates in such a way as to not result in unfair monopolisation of the airspace by a single airline operator.

5.4 Contestability and Redress

Individuals and organisations should have clear routes to dispute harmful outcomes or decisions generated by AI/automated systems. Appropriate application of this principle will be dependent on the context.

The focus here is on the ability to contest an outcome, as opposed to a specific functional output. For example, a system which adjusts aircraft control surfaces every nanosecond essentially creates 1,000,000,000 functional outputs every second. It would be impractical to enable contestability for each output. However, if the outcome of the system was to pitch the aircraft up and turn left 35 degrees to avoid a collision with another aircraft, this outcome will likely need to be contestable, particularly in the case of an incident. There is a close relationship with many of the other principles, particularly transparency and explainability which enables contestability.

5.5 Accountability and Governance

Organisations should ensure the proper functioning of the AI system throughout its lifecycle and that it is created, operated, and maintained in accordance with applicable regulatory frameworks. This should be clearly demonstrated through their actions and decision-making process.

Where "fairness" is focused on the system, "accountability and governance" are aimed at the organisations involved. As such, there are organisational factors that affect the "proper functioning of the AI system" – roles, procedures, oversight, committees, and many more. In aviation, the "operator" is a term defined in law and determines the legal responsibilities of an individual or organisation with regards to governance, safety reporting, training, and much more. Similar terms are defined for other stakeholders in the system. The application of an AI system into any of these roles should not predispose the application of existing legal responsibilities.

6 Recommendations

Upon the review and provided input into Project SafeTeam to support early regulatory thinking on safety-critical elements, and it has been concluded that, given the level of maturity achieved, it would be premature to establish a regulatory position at this stage. A formal review by the relevant regulatory experts will need to take place closer to the point of live operational trials once greater system maturity has been achieved.

- Further Testing and Validation: To ensure robustness, additional testing should be carried out to validate Use Case 2's model's predictive accuracy across a broader spectrum of airport types, operational environments, and environmental conditions. Simulator based case studies should also be developed to evaluate the system's performance and effectiveness in more complex representative real-world scenarios.
- Human Factors Research: Further research is needed to examine how human factors such as pilot confidence, decision-making processes for ATC, and training influence the ML system's reliability. Additionally, exploring the relationship between pilot self-assessments of approach stability and ML predictions could provide valuable insights for enhancing trust in the system and improving operational integration.
- Review of themed regulations and guidance's: It is recommended that further work is conducted into regulatory guidance's. This will ensure alignment with safety, reliability, and regulatory standards that may be adopted with standards in mind for a form of alternate means of compliance (AMC). Engagement with standards bodies will be essential to develop appropriate procedures for responding to the use cases developed in project SafeTeam.

Annex I: TRL Level Scale: