EXPLORATORY RESEARCH

# Risk assessment of AISA

| | |
|---|---|
| **Deliverable ID:** | **D5.1** |
| **Dissemination Level:** | **PU** |
| **Project Acronym:** | **AISA** |
| **Grant:** | **892618** |
| **Call:** | **H2020-SESAR-2019-2** |
| **Topic:** | **SESAR-ER4-01-2019** |
| **Consortium Coordinator:** | **FTTS** |
| **Edition date:** | **10 May 2022** |
| **Edition:** | **00.01.00** |
| **Template Edition:** | **02.00.05** |

EUROPEAN PARTNERSHIP

Co-funded by the European Union

## Authoring & Approval

### Authors of the document

| Name / Beneficiary | Position / Title | Date |
|---|---|---|
| Javier A. Pérez Castán/UPM | Lecturer | 15 April 2022 |
| Luis Pérez Sanz/UPM | Senior Lecturer | 19 April 2022 |

### Reviewers internal to the project

| Name / Beneficiary | Position / Title | Date |
|---|---|---|
| Tomislav Radišić/FTTS | PC /Associate Professor | 3 May 2022 |
| Ruth E. Häusler Hermann/ZHAW | WP5 leader / Assistant Professor | 3 May 2022 |

### Approved for submission to the SJU By - Representatives of all beneficiaries involved in the project

| Name / Beneficiary | Position / Title | Date |
|---|---|---|
| Javier A. Pérez Castán/UPM | Lecturer | 6 May 2022 |
| Tomislav Radišić/FTTS | PC /Associate Professor | 6 May 2022 |
| Ruth E. Häusler Hermann/ZHAW | WP5 leader / Assistant Professor | 6 May 2022 |
| Roland Gurály/SLOT | General Manager | 6 May 2022 |
| Keiko Moebus/SKYGUIDE | Head of Human Factors | 6 May 2022 |
| Bernd Neumayr/JKU | University Assistant | 6 May 2022 |
| Thomas Feuerle/TUBS | Deputy Manager | 6 May 2022 |

### Rejected By - Representatives of beneficiaries involved in the project

| Name and/or Beneficiary | Position / Title | Date |
|---|---|---|

## Document History

| Edition | Date | Status | Name / Beneficiary | Justification |
|---|---|---|---|---|
| 00.00.01 | 17/03/2022 | Draft | Javier A. Pérez Castán | New document |
| 00.00.01 | 15/04/2022 | Draft | Javier A. Pérez Castán | Initial draft |
| 00.00.02 | 4/05/2022 | Draft | Javier A. Pérez Castán | Reviews from the consortium members included |
| 00.01.00 | 10/05/2022 | Final Draft | Javier A. Pérez Castán | Final draft |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

# AISA

## AI SITUATIONAL AWARENESS FOUNDATION FOR ADVANCING AUTOMATION

This deliverable is part of a project that has received funding from the SESAR Joint Undertaking under grant agreement No 892618 under European Union's Horizon 2020 research and innovation programme.


### Abstract

This deliverable presents task 5.2 on the risk assessment of AISA. Risk assessment focuses on performing a safety analysis by identifying hazards, analysing them, and their risks (based on probability and severity) and providing mitigation measures. This work analyses the whole system that covers novel technologies based on artificial intelligence. This risk assessment provides valuable information for the further development of the AISA system that could be applied as potential safety requirements.

Risk assessment has identified areas, families, systems, and functions that can be critical to limiting the development of AI systems. This implies that measures must be imposed to avoid the appearance of some risks or to mitigate the consequences. From the risk assessment, the primary risks of distributed human-machine SA are identified. Many mitigation measures proposed in this work are related to the implementation of risks as safety requirements during the design phase of the system.




EUROPEAN PARTNERSHIP

Co-funded by the European Union

## Purpose

This deliverable is based on the Work Package (WP) 5 Concept assessment and Task 5.2 Risk assessment of AISA. The primary goals of this deliverable are:

1. Deep look into the operational concept of AISA and their limitations taking into account the different horizons.
2. The constitution of a library based on the identification and analysis of hazards and risks for the AISA system.
3. The evaluation of risks based on the ICAO methodology and the proposal of mitigation measures.

## Intended Audience

There are two main groups of the intended audience:

- The AISA consortium.
- Experts from the related fields.

The development of risk assessment via AISA deliverable (AISA D.5.1) is important for the consortium as it performs the risk assessment of the AISA system in the framework of WP5. The document is also useful for external stakeholders, especially the following ones:

- Air Traffic Management (ATM) system developers who would like to understand how AI, and particularly ML and KG methods, could be integrated into ATM.
- ATM and safety experts conducting related research.

General safety and AI experts would like to see the possible use of AI in a new domain.

## Associated documentation

The document is linked to several AISA documents; here, only the most relevant ones are listed:

- AISA D2.1: Concept of Operations for AI Situational Awareness System.
- AISA D2.2: Requirements for automation of monitoring tasks via AI SA.
- AISA D3.1: 4D trajectory prediction module.
- AISA D3.2: Conflict Detection module.
- AISA D3.3: Air Traffic Complexity estimation module.
- AISA D4.1: Proof-of-concept KG system.
- AISA D4.2: KG-Prolog Mapper.
- AISA D4.3: Populated knowledge graph.



EUROPEAN PARTNERSHIP

Co-funded by the European Union

## Terminology

The following table lists the abbreviations used in this document.

| Abbreviation | Description |
|---|---|
| ADS-B | Automatic Dependent Surveillance-Broadcast |
| AI | Artificial Intelligence |
| AIP | Aeronautical Information Publication |
| AIRAC | Aeronautical Information Regulation And Control |
| AISA | Artificial Intelligence Situational Awareness |
| ANSP | Air Navigation Service Provider |
| ATC | Air Traffic Control |
| ATCOs | Air Traffic Control Officers |
| ATM | Air Traffic Management |
| ATSEP | Air Traffic Safety Electronics Personnel |
| CD | Conflict Detection |
| ConOps | Concept of Operations |
| CWP | Controller Working Position |
| EASA | European Union Aviation Safety Agency |
| EATMP | European Air Traffic Management Programme |
| ERASMUS | En Route ATM Soft Management Ultimate System |
| ESARR | EUROCONTROL Safety Regulatory Requirements |
| ETO | Estimated Time Over |
| FAA | Federal Aviation Administration |
| FMP | Flow Management Position |
| HEP | Human Error Probabilities |
| ICAO | International Civil Aviation Organization |
| KG | Knowledge Graph |
| ML | Machine Learning |





Co-funded by the European Union

| | |
|---|---|
| NAS | National Aviation Services |
| NM | Nautical Miles |
| PoC | Proof-Of-Concept |
| PRD | Prohibited, Restricted and Dangerous |
| ProLog | Programming in Logic |
| RNP1 | Required Navigation Performance 1 |
| SA | Situational Awareness |
| SESAR | Single European Sky ATM Research |
| SHACL | Shapes Constraint Language |
| SORA | Specific Operations Risk Assessment |
| SPARQL | SPARQL Protocol And RDF Query Language |
| SUP | Shift Supervisors |
| SW | Software |
| TRL-1 | Technology Readiness Level 1 |
| TSA | Team Situational Awareness |
| UAS | Unmanned Aircraft Systems |
| UML | Unified Modelling Language |
| UPM | Universidad Politécnica De Madrid |
| WP | Work Package |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

# Table of Contents





EUROPEAN PARTNERSHIP

Co-funded by the European Union

# List of Tables





EUROPEAN PARTNERSHIP

Co-funded by the European Union

# List of Figures

# 1 Introduction

The AISA project proposes building a foundation for automation by developing an intelligent situationally-aware system instead of automating isolated individual tasks. This system will initially be able to automate some of the monitoring tasks because machines cannot currently reach the same level of awareness as humans. However, as development progresses, it will be able to handle more complex tasks. AISA is built on the basis of different novel technologies based on Artificial Intelligence (AI), knowledge engineering, and Machine Learning (ML). Each of these technologies deals with a specific development of the AISA system. Previous work can be found in the deliverables accessible on the AISA website (https://aisa-project.eu/).

This deliverable presents the development of Task 5.2 on risk assessment. Risk assessment focuses on performing a safety analysis by identifying hazards, analysing them and their risk (based on probability and severity) and providing mitigation measures. Therefore, this risk assessment will provide crucial information for further development of the AISA system that could be applied as potential safety requirements. This work analyses the whole system, which covers specific technologies and the system as a whole.

This task is framed in WP5 that focuses on assessing the concept as defined in WP 2 and developed in WPs 3 and 4. Situational Awareness (SA) is compared between AI and air traffic controller (ATCO), and human performance is evaluated.

## 1.1 Literature review

EASA defines AI as 'a branch of computer science that aims to create intelligent machines' [1]. Artificial intelligence has become an essential part of the technology industry. It can be narrow, handling only one particular task, or strong, meaning a machine with the ability to apply intelligence to any problem. It is also important to highlight the differences between the safety assessment as a result and the risk assessment on safety as a process [2]:

- Safety assessment refers to the process developed to identify hazards and their consequences without analysing their impact on the system.
- Risk assessment refers to the evaluation of the severity and consequences of hazards identified in the system by applying a safety methodology. The goal is to identify and mitigate risks.

The literature review focuses on identifying previous work that could help the completion of this work. The first step that has been carried out is a review of the literature related to the application of risk assessments in ATM. In particular, the study has focused on works related to the introduction of Air Traffic Control (ATC) tools and the use of new technologies such as ML and knowledge engineering. Due to the particularity and novelty of the usage in the ATM domain of these technologies, the topics identified do not directly cover this particular problem.



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| Document | Brief description | Result |
|---|---|---|
| **Accident Risk Assessment for Advanced Air Traffic Management [3]** | The TOPAZ methodology is used to evaluate two en-route RNP1 traffic streams, flying in opposite directions. The aim is to learn how ATC influences accident risk and how far the nominal separation, between opposite RNP1 traffic streams, can be safely reduced. | It handles complex interactions between different ATM elements and is validated with a risk assessment exercise. |
| **Agent-Based Modelling of Hazards in ATM [4]** | A large number of hazards in the current and future ATM is modelled. Existing agent-based model constructs of the TOPAZ safety risk assessment methodology are compared against the hazards in the database. Then the same is done for a new model that has been developed for those hazards that are not modelled in the previous phases. | Many usual hazards are identified and evaluated in the ATM domain. Safety-relevant scenarios that deal with a wide spectrum of issues related to the environment.<br><br>Technical systems, human operators, organization of ATM, environmental conditions and others are also studied. |
| **Systematic accident risk assessment in air traffic by Monte Carlo Simulation [5]** | TOPAZ safety risk assessment is defined for a particular operation: aircraft departing from a runway, which is occasionally crossed by taxiing aircraft. The assessment focuses on the effectiveness of a runway incursion alert system that warns an ATCO, in reducing the safety risk for typical and reduced visibility conditions. | It analyses interactions between multiple agents (humans and systems) in advanced air traffic operations. It can be a good option for difficult scenarios where many agents interact. |
| **Change-Oriented Risk Management in Civil Aviation Operation: A Case Study in China Air Navigation Service Provider [6]** | The SCOHI model identifies hazards by integrating '5M' (mission-man-machine-management-environment) and hazard and operability (HAZOP) techniques specify changes and impacts in the surrounding environment.<br><br>The case study analyses the change in the ANSP system from conventional control to radar control operations. | The effectiveness and applicability of the SCOHI model are tested with a risk assessment. An ANSP controls strategy is provided to help control the risk and maintain an acceptable level of safety during changes of the system. |





Co-funded by the European Union

| | | |
|---|---|---|
| **Risk Assessment based on SORA Methodology for a UAS Media Production Application [7]** | As Unmanned Aircraft Systems (UAS) operations are subject to compliance with applicable regulations, a risk assessment was performed. This paper presents the application of the SORA methodology to an autonomous system for aerial cinematography with a small team of UAS. | The complete risk assessment for new elements in ATM as well as mitigation measures for its future integration into airspace operations are proposed. |
| **Machine learning in air traffic control [8]** | This work analyses the MALORCA project (Machine Learning of Speech Recognition Models for Controller Assistance). It aims to reduce development and maintenance costs for assistant-based speech recognition through ML rather than manual software programming. | AI is used to achieve automatic speech recognition. The solution automatically learns local acoustic and semantic patterns and controller models from radar and speech data recordings, which are then automatically introduced into the speech recognition software. |
| **Prediction of delay due to air traffic control using machine learning [9]** | The prediction of delays is presented for air traffic streams bounded for a congested airport. The weather forecast and the expected trajectory are investigated and a neural network is used in order to measure the exact average delay using pre-departure information. | The feasibility of artificial neural networks used to make prediction about delay is investigated. A single parameter that describes the traffic condition is also derived. |
| **Predicting flight routes with a Deep Neural Network in the operational Air Traffic Flow and Capacity Management system [10]** | Neural networks are used to predict existing routes and traffic. A deep neural network is trained on historical trajectories and a set of predictors to predict the most likely route. Through iterative training on newly recorded data, the neural network can keep up with changes. | It performs predictions based on ML and deep neural network and identifies hazards to be covered. |
| **Review of techniques to support the EATMP safety assessment methodology [11]** | EATMP aims to define the means for providing assurance that a ground air navigation system is safe for operational use. More than 500 techniques from nine different industries were collected and documented, and briefly described in this report. | It discusses the two keywords in this scope, safety assurance and ground ANSP. Consolidated results of the identification and selection of techniques and methods to support |





Co-funded by the European Union

| | | |
|---|---|---|
| | | the EATMP safety assessment method are shown. |
| **Air Traffic Control Tools Assessment [12]** | Various ATC tools are presented that can serve as an aid and assist ATCO, such as MTCD, AMAN, or DMAN. This article shows the main features of the tools, which should help ATCO reduce their workload and manage the increasing number of aircraft in flight. | It provides comprehensive and organized material, describing new tools and systems used by ATCOs. It proposes improvements for further research and development of ATC tools. |
| **Human error data collection as a precursor to the development of a human reliability assessment capability in air traffic management [13]** | A first step towards development of an ATM Human Reliability Analysis approach is taken by deriving some Human Error Probabilities (HEPs) in an ATM context. HEPs are collected by analysing the results of a real-time simulation involving ATCOs and pilots, with a focus on communication errors. | A simulation has been carried out where controllers gave instructions to pilots who needed to execute them. The aim was to find human errors, as it was a new 'spacing' instruction of the aircraft. |
| **Application of Artificial Intelligence in the National Airspace System- A primer [14]** | This study seeks to examine the opportunity to exploit basic applications of AI technology to the National Aviation Service (NAS) to improve aggregate operational performance and efficiency. It describes 10 capabilities that AI can be clustered in (ML, Big Data, …) and as well it mentions examples that solve problems or demonstrates potential solutions across the NAS Integrated. | This paper defines AI; the capabilities associated with AI; current use cases within the aviation ecosystem; and how to prepare for the insertion of AI into the NAS. |
| **The future of Air Traffic Control: Human operator and automation [15]** | This report is divided into 3 sections. The first focuses on the development of ATM systems from a human factor perspective; the second assesses future automation alternatives and the role of the human operator in ensuring safety and efficiency; and the third one concludes the importance of continuing developing automation for ATCOs. | The report identifies the criticality of the interaction between the automation and the ATCOs, and the automation and the pilot in the cockpit. |
| **Functional modelling for risk assessment of** | It presents the ERASMUS project, which proposes to reduce the number of air | The analysis and risk assessment of future |

| | | |
|---|---|---|
| **automation in a changing air traffic management environment [16]** | conflicts by minor adjustments to their speed. It develops the Functional Resonance Analysis Method (FRAM), to indicate and evaluate the effects and impact on controller and pilot work resulting from ERASMUS automation. | automation systems in ATM is shown and the interrelation between system components (controller-pilot). |
| **A Descriptive Classification of Causes of Data Quality Problems in Data Warehousing [17]** | The state-of-the-art purpose of the paper is to identify the reasons for data deficiencies, non-availability or reach ability problems at the different stages of data warehousing (data sources, data integration & data profiling, data staging and ETL, data warehouse modelling & schema design) and to formulate descriptive classification of the causes. | Useful for identifying hazards for the Knowledge Engineering area, in particular, with data quality issues. |
| **Knowledge Graph Quality Control: A Survey [18]** | This paper aims to present a comprehensive survey on the quality control of Knowledge Graphs (KGs). First, it defines six main evaluation dimensions of KG quality and investigates their correlations and differences. Second, quality control treatments during KG construction are introduced from the perspective of these dimensions of KG quality. Third, the quality enhancement of a constructed KG is described from various dimensions. | It shows how to evaluate the dimensions of KG quality, the quality control of the construction process, and the quality enhancement methods. It can be used for those areas of research related to the KG. |
| **Architecture and Quality in Data Warehouses: An Extended Repository Approach [19]** | A large number of quality aspects relevant for data warehousing cannot be expressed with metamodels. This paper makes two contributions towards solving these problems. Firstly, it enriches the meta-data about architectures by explicit enterprise models. Second, many very different mathematical techniques are being developed to measure or optimize certain aspects of data quality. | Acquire greater knowledge in the area of data warehouse and information integration. |

The main results of the literature review have been the identification of previous works that could serve as a basis for the risk assessment of AISA and to identify hazards and risks related with these novel technologies that can enrich the future hazard library.





Co-funded by the European Union

## 1.2 Current risk assessment methodologies

The purpose of a risk assessment methodology is generally to ensure that the proposed system is safe from a risk perspective. There are several methodologies that address risk assessment purposes. The most important was proposed by ICAO and FAA and EUROCONTROL adopted it. They are briefly explained in this section to show their guidelines. Most of the methodologies are quite similar and the purpose is the same: the analysis and management of the risks in a system. After analysing all of them, it is concluded that the ICAO methodology could be the best option to carry out this particular risk assessment. It is the one that allows for a better approach without the necessity of quantitative analysis.

### 1.2.1 ICAO

ICAO defines safety risk management as the process that includes the evaluation and mitigation of safety risks as a consequence of the hazard [2]. The methodology is divided into several steps. First, it starts with the identification of potential hazards within the equipment and procedures used. Once they are detected and analysed, the goal is to reduce their impact as much as possible. Two metrics are defined to characterise the risk: likelihood and severity. The likelihood evaluates the probability that the consequences of a particular risk could appear during service provision. Table 1 shows the five likelihood categories considered by ICAO (denoted with a numerical value of 1 to 5). The meaning of each level depends on the definition of the problem by the risk experts.

| *Likelihood* | *Meaning* | *Value* |
|---|---|---|
| Frequent | Likely to occur many times (has occurred frequently) | 5 |
| Occasional | Likely to occur sometimes (has occurred infrequently) | 4 |
| Remote | Unlikely to occur, but possible (has occurred rarely) | 3 |
| Improbable | Very unlikely to occur (not known to have occurred) | 2 |
| Extremely improbable | Almost inconceivable that the event will occur | 1 |

**Table 1. Likelihood for safety risks [2].**

The next step is similar to the previous, focusing on the severity of each hazard. Classification is indicated with letters (from A to E) as shown in Table 2.



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| *Severity* | *Meaning* | *Value* |
|---|---|---|
| Catastrophic | • Aircraft / equipment destroyed<br>• Multiple deaths | A |
| Hazardous | • A large reduction in safety margins, physical distress or a workload such that operational personnel cannot be relied upon to perform their tasks accurately or completely<br>• Serious injury<br>• Major equipment damage | B |
| Major | • A significant reduction in safety margins, a reduction in the ability of operational personnel to cope with adverse operating conditions as a result of an increase in workload or as a result of conditions impairing their efficiency<br>• Serious incident<br>• Injury to persons | C |
| Minor | • Nuisance<br>• Operating limitations<br>• Use of emergency procedures<br>• Minor incident | D |
| Negligible | • Few consequences | E |

**Table 2. Severity for safety risks [2]**

Once the risk is quantified based on the alphanumeric value (output from the combination of likelihood and severity), it can be referred to the tolerability matrix of Table 3.

| *Safety Risk* | | *Severity* | | | | |
|---|---|---|---|---|---|---|
| *Probability* | | *Catastrophic A* | *Hazardous B* | *Major C* | *Minor D* | *Negligible E* |
| Frequent | 5 | 5A | 5B | 5C | 5D | 5E |
| Occasional | 4 | 4A | 4B | 4C | 4D | 4E |
| Remote | 3 | 3A | 3B | 3C | 3D | 3E |
| Improbable | 2 | 2A | 2B | 2C | 2D | 2E |
| Extremely improbable | 1 | 1A | 1B | 1C | 1D | 1E |

**Table 3. Tolerability matrix for risk assessment [2].**

Depending on the alphanumeric value, the risk is located in one of the three areas: acceptable (green), tolerable (orange) and intolerable (red). Table 4 shows the recommended actions for each area.



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| *Safety Risk Index Range* | *Safety Risk Description* | *Recommended Action* |
|---|---|---|
| 5A, 5B, 5C, 4A, 4B, 3A | INTOLERABLE | Take immediate action to mitigate the risk or stop the activity. Perform priority safety risk mitigation to ensure additional or enhanced preventative controls are in place to bring down the safety risk index to tolerable. |
| 5D, 5E, 4C, 4D, 4E, 3B, 3C, 3D, 2A, 2B, 2C, 1A | TOLERABLE | Can be tolerated based on the safety risk mitigation. It may require management decision to accept the risk. |
| 3E, 2D, 2E, 1B, 1C, 1D, 1E | ACCEPTABLE | Acceptable as is. No further safety risk mitigation required. |

**Table 4. Recommended actions for risk assessment levels [1].**

These are the basis used in this work to perform the risk assessment. Taking into account the likelihood and severity values, the different risks will be evaluated and mitigation measures will be provided to reduce the risk level of hazards.

### 1.2.2 FAA

FAA develops a similar process to that of ICAO but with some differences. Figure 1 shows the steps proposed for the Safety Risk Management methodology.

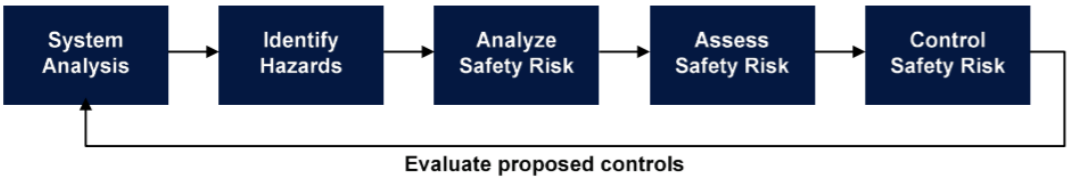


**Figure 1. 5 steps of the process by the FAA [20].**

First, a complete and accurate description of the system is required. It should provide information that serves as the basis for identifying and understanding hazards, as well as their causes and associated risks. When describing it, many aspects must be considered, such as human factors requirements or system functions, processes, procedures, performances, etc.

Once the system has been described, it is time to identify hazards. During this step, each hazard must be documented as well as its possible causes, the conditions under which hazards might be realized, and their corresponding effects. It should be noted that each hazard may have a different risk level in each possible system state, or even not exist in every system state. Therefore, it is important to consider all credible possibilities and all conditions that could cause or contribute to an accident. The third step deals with the assignment of severity and likelihood to each of the hazards identified in the previous step. The combination of both parameters is the risk.

Similarly to the ICAO methodology, the FAA also provides a guide in which generic severity and likelihood definitions are given to determine the value associated with each hazard. It is important to note that, compared with ICAO scales, likelihood and severity acquire different


EUROPEAN PARTNERSHIP
Co-funded by the European Union

values. Severity is evaluated by a numeric shape (in the ICAO methodology, it is evaluated with a character from A to E), while a character would be assigned to the likelihood of the hazard. Both ranges of values are shown in Figure 2 and Figure 3.

| Minimal<br>5 | Minor<br>4 | Major<br>3 | Hazardous<br>2 | Catastrophic<br>1 |
|---|---|---|---|---|
| Negligible safety effect | • Physical discomfort to persons<br>• Slight damage to aircraft/vehicle | • Physical distress or injuries to persons<br>• Substantial damage to aircraft/vehicle | Multiple serious injuries; fatal injury to a relatively small number of persons (one or two); or a hull loss without fatalities | Multiple fatalities (or fatality to all on board) usually with the loss of aircraft/ vehicle |

*Excludes vehicles, crew, and participants of commercial space flight.

**Figure 2. Severity definitions by FAA. [20]**

| | Qualitative | Quantitative – Time/Calendar-based Occurrences Domain-wide/System-wide |
|---|---|---|
| Frequent<br>A | Expected to occur routinely | Expected to occur more than 10 times per year |
| Probable<br>B | Expected to occur often | Expected to occur between one and 10 times per year |
| Remote<br>C | Expected to occur infrequently | Expected to occur one time every 1 to 3 years |
| Extremely Remote<br>D | Expected to occur rarely | Expected to occur one time every 3 to 10 years |
| Extremely Improbable<br>E | Unlikely to occur, but not impossible | Expected to occur less than once every 10 years |

**Figure 3. Likelihood Definitions for Commercial Operations Category by FAA. [20]**

Using the risk matrix shown in Figure 4, it is possible to determine the risk level. This matrix seems to be almost the same as the one shown in the ICAO methodology, just differing in the alphanumeric characters. However, it can be identified that FAA occupies a larger area in the low-risk hazards than ICAO. That means that the ICAO methodology is more restrictive than the FAA one.





Co-funded by the European Union

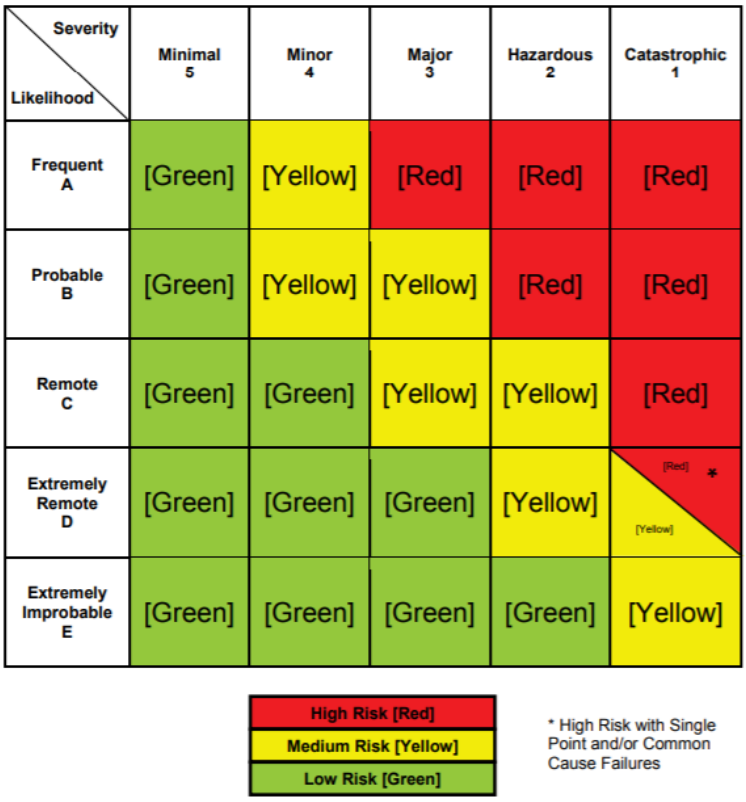


**Figure 4. FAA Risk Matrix [20].**

Finally, the fifth step consists of developing and managing options to deal with the risk. It is the risk mitigation phase, and many options are given to avoid or reduce the risks identified previously. This process should contain sufficient detail to allow the assessment of its impact on safety risk.

### 1.2.3 EUROCONTROL

EUROCONTROL also provides its own methodology not far from those seen in the literature [21]. The process does not differ from the previous ones, but the main differences are related to the severity and likelihood values shown in Figure 5 and Figure 6.





Co-funded by the European Union

| Severity Class | 1 [Most Severe] | 2 | 3 | 4 | 5 No safety effect [Least Severe] |
|---|---|---|---|---|---|
| Effect on Operations*) | Accidents | Serious incidents | Major incidents | Significant incidents | No immediate effect on safety |
| Examples of effects on operations include*): | ❑ one or more catastrophic accidents,<br>❑ one or more mid-air collisions<br>❑ one or more collisions on the ground between two aircraft<br>❑ one or more Controlled Flight Into Terrain<br>❑ total loss of flight control.<br><br>No independent source of recovery mechanism, such as surveillance or ATC and/or flight crew procedures can reasonably be expected to prevent the accident(s). | ❑ large reduction in separation (e.g., a separation of less than half the separation minima), without crew or ATC fully controlling the situation or able to recover from the situation.<br><br>❑ one or more aircraft deviating from their intended clearance, so that abrupt manoeuvre is required to avoid collision with another aircraft or with terrain (or when an avoidance action would be appropriate). | ❑ large reduction (e.g., a separation of less than half the separation minima) in separation with crew or ATC controlling the situation and able to recover from the situation.<br><br>❑ **minor** reduction (e.g., a separation of more than half the separation minima) in separation without crew or ATC fully controlling the situation, hence jeopardising the ability to recover from the situation (without the use of collision or terrain avoidance manoeuvres). | ❑ increasing workload of the air traffic controller or aircraft flight crew, or slightly degrading the functional capability of the enabling CNS system.<br><br>❑ **minor** reduction (e.g., a separation of more than half the separation minima) in separation with crew or ATC controlling the situation and fully able to recover from the situation. | No hazardous condition i.e. no immediate direct or indirect impact on the operations. |

**Figure 5. Severity Classification Scheme in ATM by EUROCONTROL. [21]**

| Severity Class | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Maximum tolerable probability (of ATM direct contribution ) | $1{,}55.10^{-8}$ Per Flight/Hour | To be included in a future revision of ESARR 4, once enough safety data have been collected according to ESARR 2♦. | To be included in a future revision of ESARR 4, once enough safety data have been collected according to ESARR 2♦. | To be included in a future revision of ESARR 4, once enough safety data have been collected according to ESARR 2♦. | To be included in a future revision of ESARR 4, once enough safety data have been collected according to ESARR 2♦. |

♦ *To be determined at national level based on past evidence on numbers of ATM related incidents.*

**Figure 6. Risk Classification Scheme in ATM by EUROCONTROL. [21]**

It provides another option to evaluate severity and likelihood. However, it barely matches what is sought in this work. First, severity values are more focused on aircraft separation and vulnerability of these minimum distances instead of referring to workload and other aspects that this analysis is trying to identify. Second, probability values exist only for class 1 while for the rest are not currently defined. Therefore, it is not possible to choose this classification as a valid method for this task, since it would not provide likelihood values for most hazards.



EUROPEAN PARTNERSHIP

Co-funded by the European Union

### 1.2.4 FAA/EUROCONTROL

Both FAA and EUROCONTROL have worked together to maintain and improve the effectiveness of safety assessment [22]. They have summarized both methodologies and give a seven-stage safety assessment process.

The process is as follows: After having a system description, the system must be analysed in order to see how it could impact, for the better and/or for worse, with respect to safety and what involves considering the scope of the assessment. There is also a need to learn how the system should behave, the nominal system, from which all possible hazards can be identified and added afterward into a risk model. The goal is to evaluate the risk to the proposed system or a possible change. In these types of risk methodology, an event or fault tree analysis is commonly used in which several hazards are consequences of a specific one. Regarding this project, this type of diagram has not been used, since many of the threats may be independent of each other. Once the previous steps are performed, the safety analyst determines how probable these failures are and how likely the system is to recover from such failures. All this builds up the total risk estimation for the system. Finally, mitigation measures are identified to reduce or even remove the different risks, later necessary to confirm that the actual risk is tolerable once the various solutions have been analysed. Therefore, conceptually, the methodology is the same as that of the ICAO.





Co-funded by the European Union

# 2 Methodology

The purpose of this risk assessment is to analyse the feasibility of the AISA concept by focusing on the risks. In addition, recommendations are identified for safe development and possible implementation. One of the achievements of this work is further elaboration of the AISA system by examining the operational concept and the requirements indicated in D2.1 2 [23] and D2.2 [24].

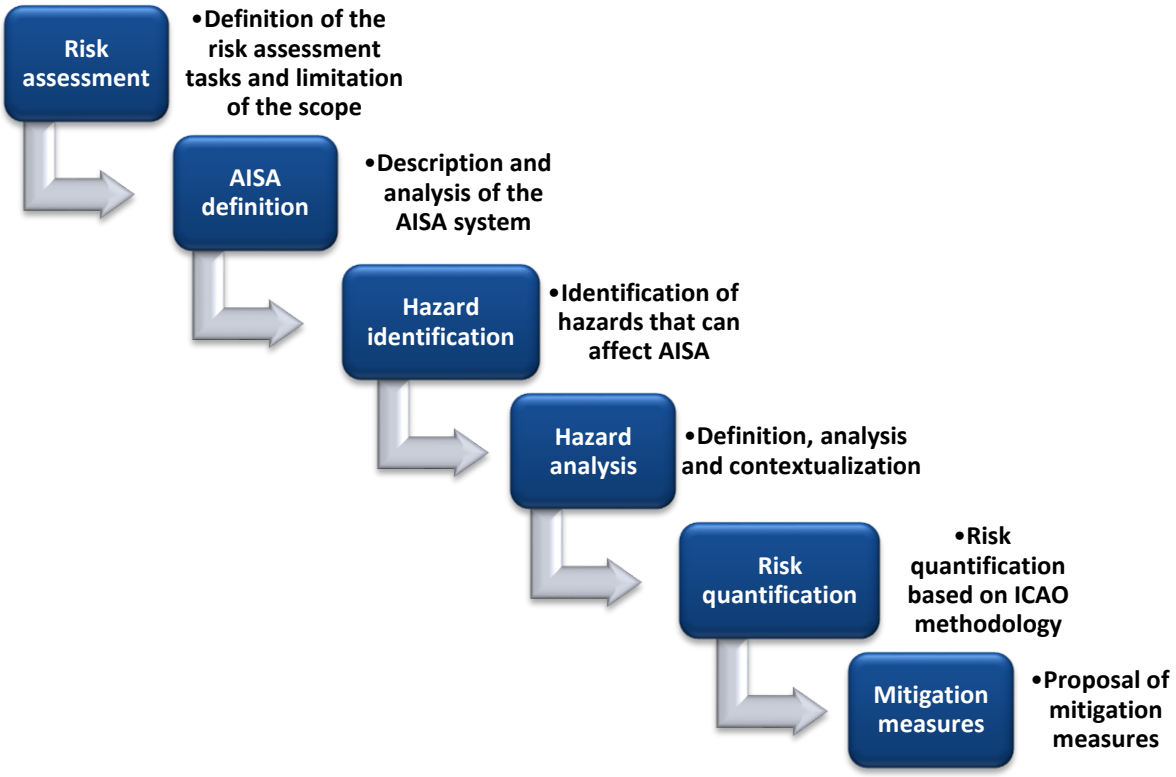


**Figure 7. Steps of the methodology.**

Figure 7 shows a diagram with the steps of the methodology used for the risk assessment.

1. Characterization of the risk-assessment scope: This is the first step of the risk assessment methodology because it limits the scope of the risk assessment by defining the tasks to be developed. In addition, it characterises the boundaries and assumptions considered for the risk assessment.

2. Definition of the AISA system: This step is similar to the definition of the nominal system. It details the scope of the AISA project and tries to summarise the operational concept of AISA from D2.1 and D2.2. The goal is to provide enough information to understand the risk assessment by a non-expert reader. Another goal is to explore the AISA system by increasing the knowledge of the AISA concept.



EUROPEAN PARTNERSHIP

Co-funded by the European Union

3. Hazard identification: The goal is to identify the hazards that can affect the AISA system based on the knowledge available at the time of the risk assessment. One of the crucial aspects is the integration of different concepts and technologies in the AISA system: knowledge engineering, machine learning, different data sources, and reasoning engineering among others. This step involved all AISA members to provide their expertise in hazard identification.

4. Hazard analysis: After identifying potential hazards, they must be analysed to understand how they can affect the AISA system. Hazards will be described and grouped by different areas and families to ease their classification. It is important to analyse the context and implications that can appear for each hazard.

5. Risk assessment: the risk assessment performs hazards analysis considering the likelihood and severity based on the ICAO methodology (see Section 1.2). This is a qualitative analysis that will provide information on each risk and the need to propose mitigation measures to reduce the risk. This analysis was split into two steps:

    a. Task 5.2 performed an initial risk assessment to quantify risks and propose mitigation measures for all hazards.

    b. UPM led the final risk assessment session in which experts from different fields and the AISA consortium participated. The goal was to analyse pre-identified risks by UPM based on AISA system hazards. Appendix B summarises the development of the final risk assessment session.

6. Proposition of mitigation measures: the last step of the risk assessment is to provide mitigation measures to reduce the risk of hazards. The goal is to provide mitigation measures for every hazard. After the proposal of mitigation measures, the risk is re-quantified. Some of the expected mitigation measures could be used in the future development of the AISA system as safety requirements.

## 2.1 Hazard library

One of the outcomes of this work is a library that compiles the hazards and risks for the AISA system. This library is expected to expand the knowledge of ML and knowledge engineering, as well as the introduction of this type of technology in ATM. The library is a document that contains the information obtained from this analysis to identify hazards, risks, and potential mitigation measures. This safety record is specified by:

- Identifier: code that uniquely identifies a risk.
- Family: the risks are grouped by different common areas.
- Hazard: name of the hazard.
- Risk description: brief and clear description of the risk.
- Operational context: clarification of the implications for the AISA system.
- Likelihood: number (1-5) associated to the risk.



EUROPEAN PARTNERSHIP

Co-funded by the European Union

- Severity: letter (A-E) associated to the risk.
- Mitigation measures: mitigation measures identified in order to reduce the risk of hazard and to increase the system safety.

## 2.2 Risk assessment scope

As already explained, one of the key steps in a risk assessment is to determine the limits of the risk assessment. Namely, it must be clarified what it is going to be done and what is beyond the scope.

The AISA system is analysed based on the Concept of Operations [23], which is divided into two horizons. On the one hand, there is an approach addressed to the future system: focusing on the interaction between the human and the machine, defining which information should be provided to whom and how the machine and the ATCO can share monitoring tasks depending on the workload (Future ConOps). On the other hand, there is a short-term solution to develop the technology and analyse the viability of its development (Project ConOps). Moreover, the limitations of the Proof-of-Concept (PoC) developed throughout this project are considered and are crucial to understanding the development and limitations of the AISA system.

The main limitations of this analysis are:

- AISA is a TRL-1 exploratory project showing initial conceptual directions for a possible future system only; therefore, the usual requirement setting methods are only partially relevant.
- The risk assessment focuses on PoC and Project ConOps, as well as information extracted from the Requirements document [2]. Future functions considered for AISA are not evaluated.
- A qualitative risk assessment is developed because most of the intelligence technologies considered in this work are currently in development and statistical data are not available.
- Risk assessment is limited to current knowledge of new AI technologies. As AI technologies evolve, a new risk could be added in future work.
- This risk assessment analyses an experimental phase with which is expected to obtain solutions to the main problems of the current PoC level and to suggest improvements for the further development of the system.
- The impact of automating some monitoring tasks by AISA on ATCO's workload is not considered. Other project tasks can provide information to the risk assessment based on their results, such as task 5.1 (which focuses on evaluating the comparison of SA between AI and ATCO) and task 5.3 (which evaluates the human performance in distributed Situational Awareness), both results are published in D5.2.
- The risk assessment encompasses the proposal of mitigation measures for individual risks. It is expected to re-quantify the likelihood and severity considering the mitigation measures proposed. However, the introduction of mitigation measures into the AISA



EUROPEAN PARTNERSHIP

Co-funded by the European Union

system is out of the scope. This limitation is imposed because the AISA system is developed at the PoC level.

- Modes of operation are not considered, based on the possible distribution of monitoring tasks between AISA and ATCO.
- The risk assessment focuses on an operational level; then, legal issues and cybersecurity risks are also out of the scope of this analysis.



EUROPEAN PARTNERSHIP

Co-funded by the European Union

# 3 Description of AISA system

This section shows some concepts and a brief description of the AISA system to understand it. More detailed information can be found in D2.1 and D2.2. The goal is to describe the base system considering two AISA levels (Future and Project). Future ConOps shows how the system should be in the future (horizon 2050), and Project ConOps develops the way the system should be in a short horizon (2035). In particular, the Proof-of-Concept (PoC) develops the solution planned at the project level restricted by temporary limitations. Therefore, this assumption is considered because it covers the development performed during this project.

Conceptually, AISA is a system made up of human actors (ATCO) and machine (AI) working together by sharing the same (or similar) SA and monitoring tasks. Sharing the SA between the ATCOs and the AI is denoted as Team SA (TSA) and allows one to reach the same conclusions for ATCOs and the AI. The goal is to achieve some level of AI automation and help ATCOs in different situations by providing information for the decision-making process. A differential factor is that AI helps ATCO labour, which implies that AISA can provide some reasoning to the ATCO depending on the information or solution provided.

Figure 8 shows the conceptual design of the AISA system:

1. The core of the system is the Knowledge Graph (KG), which receives, handles, and stores the information.
2. ML modules make predictions in the following areas currently: trajectory, conflict, and complexity. The future AISA system could use more ML modules.
3. The reasoning engine provides conclusions of the AISA system based on rule-based knowledge designed in conjunction with ATCOs.

Combining all of these elements, the system acquires the SA similarly to a human does.

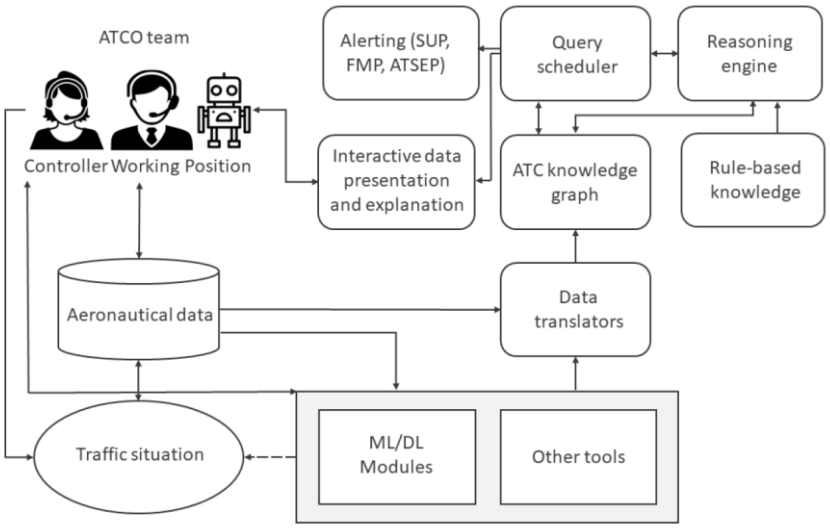


**Figure 8. Conceptual schema of AISA system [23].**





Co-funded by the European Union

Moreover, the analysis of the AISA system goes into the concept and requirements. The structure is based on specific questions that can help the reader understand the entire AISA system, the functionalities, and the performance.

## 3.1 AISA general overview

This section performs an analysis of the AISA system as a whole.

| | |
|---|---|
| **What is AISA?** | AISA consists of both human and machine (AI) actors that will work together as a team. By doing so, AI will be able to take over some monitoring tasks, warn the ATCO if it lost SA, and the ATCO will need to respond to the situation at hand, as well as support the ATCO's decision-making process by providing suggestions. The vision of the AISA project is for AI to serve as a platform. Other automated tools would be able to get data from the platform necessary for their operation; thus, the AI will have sufficient SA to automate certain ATCO monitoring tasks. Instead of automating isolated individual tasks, such as conflict detection or coordination, the system is a foundation for automation by developing an intelligent situationally-aware system**.** |
| **What does AISA do?** | The main goal of the system is to form an additional SA system that can build up SA similar to the SA of ATCOs by using the same data that ATCOs use. The system is then able to reason about the collected knowledge and present its findings to the ATCO team in a transparent manner for further evaluation by the ATCOs.<br><br>Having a good understanding of the traffic situation, in terms of artificial SA, means that the system is able to gather all the necessary data regarding the current traffic situation, turn it into knowledge, and then draw conclusions based on the knowledge gained. |
| **How does AISA work?** | In this project, the AISA architecture is based on two main parts: KG with reasoning engine and ML modules.<br><br>- The KG is used to store all the knowledge necessary to perform the monitoring tasks. The reasoning engine is used to reason about the facts stored in the KG. The Reasoning Engine in AISA is used to implement rules that cannot be described via the KG.<br>- ML modules perform those tasks that cannot be calculated directly or cannot be inferred from the existing knowledge in the knowledge graph, i.e. predictions and estimates.<br>Then, ML is used at a lower level to predict individual probabilistic events, whereas reasoning engine is used at a higher level to draw conclusions about the system state. By combining the reasoning engine with ML, it will be possible for AI to 'be aware' of the situation in a manner similar to a human, that is, AI will be able to assess complex |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| | |
|---|---|
| | interactions between objects, draw conclusions, explain the reasoning behind those conclusions, and predict future state of the system. |
| **What are the pre-requisites for AISA to operate?** | For AISA to be part of the air traffic future, the assumption is that most new advanced technologies described in the ATM Master Plan will be available, which is a necessary prerequisite considering the high automation of AISA, such as SWIM, AIXM, FIXM, CPDLC, etc. |
| **What are the AISA functions?** | Automation of monitoring tasks<br>Central coordination of tools/modules<br>Automation of gathering of missing information<br>Automated reporting<br>Awareness of the system state<br>Awareness of Team member's state |
| **What does SA mean in AISA?** | Situational awareness (SA) is the perception of environmental elements and events with respect to time or space, the comprehension of their meaning, and the projection of their future status. There are 3 types of SA:<br>1. Awareness of the traffic situation. To be aware of the traffic situation, in terms of artificial SA, means that the system is able to gather all necessary data regarding the current traffic situation, turn it into knowledge, and then draw conclusions based on the knowledge gained.<br>2. Awareness of its own state. Queries related to self-monitoring will be used to ensure that the system is operating nominally. These will allow AI to be aware that part, or whole, of the system is failing and to transfer the tasks back to the ATCOs.<br>3. Awareness of the states of the other team members. Workload, however, can be inferred on the basis of the traffic complexity with the idea that the more complex the traffic is, the higher the workload will be. Now, this can be estimated by trained ML systems for complexity assessment. This approach will be taken in AISA, with a dedicated ML module used to assess the complexity of the current air traffic situation. |
| **What does TSA mean in AISA system?** | Shared or TSA means that two or more people have a commonly understood mental image of what is happening and/or what is going to happen in the near future. Sharing the same TSA among ATCO team members and AI will allow the automated system to reach the same conclusions as ATCOs when confronted with the same problem and to be able to explain the reasoning behind those conclusions. Enabling human-machine SA requires that both entities have access to the same data. This would mean a connection must be formed between data for human use, language, and data for machine use, information. |
| **What does it imply that AISA** | The first point is that AISA can read and connect data in the same way as a human does. If we build the AISA system, this system provides a SA to a human. The information from AISA will be presented to the ATCO |





Co-funded by the European Union

| | |
|---|---|
| **performs SA tasks?** | and, therefore, some work will not be done by the ATC. In this way, the ATC could cover a larger geographical area or a large number of aircrafts. |
| **How does AISA acquire SA?** | By combining the reasoning engine with ML, we believe that AISA will be 'aware' of the situation in a manner similar to a human. AI will be able to assess complex interactions between objects, draw conclusions, explain the reasoning behind those conclusions, and predict future state of the system. Queries will be developed for each of the tasks that the system should be able to execute. By running these queries in short intervals, a continuous monitoring will be achieved. Queries will be used to achieve situational awareness and provide results to the ATCO through the Controller Working Position (CWP). |
| **What is the team ATC and its responsibilities?** | In AISA, the vision for future operations includes humans in the central role. At the core of the concept, there are controllers. Whether the ATCO team is made of a conventional Executive/Planner couple or some other future combination of roles (e.g. in case of a multi-sector planner), they will be joined by AI. As usual, ATCOs work in their CWP, which they use to gather information, build their SA, make decisions and implement actions.<br><br>The vision for the future automation concept of en-route ATC operations includes a human-machine distributed TSA with a sector team consisting of executive ATCO, planning ATCO, and AI (actors). Actors will be able to continually monitor each other's states, AI being aware of the probable states of human actors through traffic situation analysis. Tasks will be dynamically allocated according to actor states, including graceful degradation of automation to ensure business continuity. The system should monitor traffic and help guide ATCO's attention to those tasks that are best suited for humans, such as decision making. Also, results from queries related to system state, i.e. detecting the performance degradation, can be forwarded to air traffic safety electronics personnel (ATSEP), and results from queries related to workload or demand-capacity balancing in general can be sent to shift supervisors (SUP) or flow management position (FMP).<br><br>In current ATC operations, each human team member, executive, or planner ATCO, is aware of the following.<br><br>- traffic situation (by looking at the radar screen),<br>- their state (e.g. feeling rested or tired),<br>- other team member state (by verbal/nonverbal communication), and<br>- system state (by inspecting the error messages, warning lights etc). |





Co-funded by the European Union

| | |
|---|---|
| | On the other hand, in current ATC operations, the system is unaware of the state of the ATCOs, it is unaware of the traffic situation, and it has very limited awareness of its state. |
| **What are the tasks implemented in AISA?** | 1. Conformance management<br>2. Detect incoming planned flight<br>3. Assume, identify and confirm aircraft<br>4. Assess if exit conditions are met<br>5. Conflict management<br>6. Execute Aircraft's plan<br>7. Transfer aircraft<br>8. Maximize quality of service<br>9. Workload monitoring<br>10. Identifying missing information<br>11. Monitor the status and performance of ATC sub-systems |
| **How does AISA know about itself and its performance?** | This knowledge is obtained by a subset of tasks. While AISA performs several routine questions about its state, it can know about its state. The tasks related to evaluate its own status are:<br>- Monitor status of ATC sub-system.<br>- Monitor ML modules performance<br>And, not directly but related to this item, AISA can merge different streams of data (AIXM, FIXM, etc) and detect if there is a discrepancy between the data or if the data is missing. |
| **Are differences between the tasks defined at the project and concept level?** | Yes. The PoC system performs only monitoring tasks that are automated. In the future, AISA is expected to respond to different actions such as "respond to safety net alerts" or "respond to received coordination from adjacent sectors". A comparison of the task perspectives depending on the AISA level considered can be found in D2.1 [23]. |

The next sections cover questions related with specific areas of AISA.





Co-funded by the European Union

## 3.2 Questions related to ML

This section shows questions related to ML modules and their performance in the AISA system:

| | |
|---|---|
| **What information does the ATC receive from ML modules?** | Currently AISA works with three ML modules:<br>- Trajectory prediction module: trajectory prediction and a level of confidence<br>- Conflict detection module: conflict detection metrics such as estimation of the minimum distance and distance to the closest point of approach.<br>- Complexity module: estimation of the current complexity level |
| **How is evaluated the validity of ML modules and how is integrated in the output?** | Queries monitor data inputs and solutions provided by ML modules. Inputs are checked to determine whether they are within operating parameters of the ML modules (e.g. check if the inputs are something that neural network was trained to use as inputs). If they are not, KG considers the results of the ML algorithm unreliable and, if necessary, alerts the ATCO (i.e., 'known unknowns'). Results will be queried to determine if they are reasonable compared to the facts stored in the knowledge base (e.g. unreasonably high or low predicted speeds, altitudes, etc.), and if they are not search for alternative solutions, or just discard the results. |
| **Can AISA identify that ML does not have enough historical data or it has not been trained for one particular situation?** | This is something that AISA evaluates based on the Metadata. AISA can identify that it has not been trained for one particular situation based on the deviation of the output compared with the statistics of the training data. At PoC level, there is no alarm because the system is neither integrated with the CWP nor working in real-time.<br><br>Two types of error can appear: 1) AISA says it is not covered but really it is (it implies that AISA does not work well with the Metadata), and 2) AISA says it is covered by the metadata but, in reality, it does not. In both cases, it is a wrong performance of the plausibility of the ML models. |
| **Can AISA identify a ML output is wrong?** | Similar to previous question. |
| **Is it necessary to provide the reasoning from the ML output to the ATC on live?** | The AISA system is not supposed to explain the prediction of the ML module. It can only assess the plausibility of the output and determine if the inputs were correct. It can be analysed in the post-processing of the situations identified as non-conformance by the ATCO and to evaluate the accuracy degradation of the ML modules. |
| **In the case a ML module does not** | If the ML Module or any other data source is unresponsive (not providing the data), AISA will detect such case and can trigger an alert. |





Co-funded by the European Union

| | |
|---|---|
| **work, what does AISA do?** | |
| **Reliability of the ML modules about a high number of wrong predictions** | Currently, the ATC trusts on the functionalities of the CWP because it has been tested and complies with some reliability, e.g. 90%. Some requirements about the reliability and trustworthiness of the ML modules should be included, similar to current predictions.<br><br>It should be interesting that the introduction of the ML modules should be performed in conjunction with the current modules. This will ease the ATCO to trust these new methods because they will see that they fail (at least on theory) fewer than the current ones. |

## 3.3 Questions related to KG

This section shows questions related with KG system:

| | |
|---|---|
| **How does the KG work?** | KG is populated by aeronautical data provided in AIXM and FIXM. These data are already available in structured formats complying to a strictly specified schema, which enables drawing the semantic relationships between data. However, ATC-specific rules, which are not encoded in aeronautical data, must be included before the system can provide reasoning capabilities. The reasoning engine would then be able to derive new knowledge and determine logical consequences from the available data. Also known as a rules engine, semantic reasoner, or just reasoner, it is a piece of software that is meant to infer logical consequences from a set of facts. Although the two terms are sometimes synonymous, reasoning engine is a more general term than inference engine. The Reasoning Engine in AISA is used to implement rules that cannot be described via the KG. This will expand the capability of the system to implement complex rules and produce queries based on them. |
| **How does the reasoning engine work?** | A reasoning engine usually works by applying either forward chaining or backward chaining to the data. The two methods present a mirror approach to data – forward chaining starts with the available data and 'reasons' its way to an answer, while backward chaining starts with the answer and tries to prove that it is correct by searching the available data. For these reasons, they are sometimes called 'data-driven' and 'goal-driven' methods, respectively. |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| | |
|---|---|
| **How is defined the limits of the AISA knowledge?** | The limits of AISA knowledge are defined in D4.4. It is described the different information that can be extracted from AISA and, in turn, the limits of the AISA knowledge. |
| **Does the ATCO know the limits of the AISA knowledge?** | ATCO should know which tasks are automated and which are not. They don't need to know exactly what knowledge is available to AISA. |
| **How is the data updating process in the system?** | In the PoC, some parts of the process are performed automatically, and parts are performed manually. In the future, obviously everything should be automatic: data entry is hybrid, schema mapping is automatic, and mapping between KG and Prolog is automatic. |

## 3.4 Questions related to AISA system

This section shows questions related to the AISA system:

| | |
|---|---|
| **Do ATC and AISA get access to the same information?** | AISA and ATCO have access to the same information; however, AISA can access to more information than the ATCO, e.g., historical traffic data. One of the main concepts of AISA is to populate the KG from the same data sources as the data displayed to the ATCO (AIXM, FIXM, WXXM). Other sources of knowledge must be used to be able to reason using the same facts as ATCO. This knowledge is encoded manually from subject matter experts, thus mimicking the knowledge that the ATCO obtained via training and experience. One limitation is that the system will never have the general level of knowledge as ATCO. |
| **Does AISA work in real time? What is the updating time span?** | AISA PoC does not work in real time because it is out of the scope of this project. However, in the future it should be developed to work on real time and with an updating time span that provides the information to the ATCO accordingly to their needs. |
| **In the case ADS-B information is lost, can AISA know it?** | If the data source is lost, AISA can detect that there are gaps in the KG. This can be done automatically. Ideally, the ATCO or ATSEP should be notified. Depending on which data are missing, this can be a small inconvenience or a big failure. |
| **Can AISA identify when ADS-B data is not complete or with errors?** | In the PoC, AISA shows a command line output that states that there are missing data. In the future concept of operations, it would inform ATCO or ATSEP personnel. |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| | |
|---|---|
| **Which is the primary source for AISA and ATC?** | In the PoC, AISA works with ADS-B data now, within the project, but this is one limitation of the PoC. In future ConOps, it is envisaged that the AISA will have the same data sources available as ATCO. |
| **How is the AISA HMI designed?** | The HMI is outside of the scope of the project. The results are displayed in the command line window although the results are not immediately readable by ATCO. Besides, it has been experimented with the introduction of some inputs via voice. In future ConOps, AISA should be able to communicate via voice or display, whatever is more accessible to ATCO and more appropriate for a given data type, depending on further research. |

## 3.5 Questions related to the AISA performance in conjunction with the ATCO

This section shows questions related to the AISA system in conjunction with the ATCO

| | |
|---|---|
| **Can AISA explain the reasoning followed?**<br><br>**Can ATCO understand this reasoning?** | Artificial SA enables the system to possess reasoning capabilities which are explainable, completely transparent, and generalizable. Unlike most AI today, the goal of the AISA project is to avoid the so-called “black box” issue by combining ML with the reasoning engine. Regarding the 'black box' issue, the user usually controls the input and receives an output from the machine that cannot be explained. The machine lacks the ability to explain and justify the processes that lead to the generation of the visible result. In AISA the situationally aware system is not simply obeying certain laws and rules and react accordingly, but is also capable of explaining why it made a certain decision.<br><br>Whereas ML systems effectively work as black boxes, reasoning engine can explain the results it provides. It can also be used to check results obtained from ML systems for logical inconsistency or implausibility in a way similar to that of the human in determining when those results are faulty. Furthermore, when plugging in several ML modules doing the same type of estimation or prediction in different ways (even produced by different vendors), artificial SA can be used to arbitrate which of the ML modules is correct. |
| **Does the ATCO know which information from AISA is reliable or is an estimation?** | One of the uses of the AISA system will be to help integrate information from various sources into a single semantic representation. Some of the sources, such as aeronautical information publications, provide information of the highest integrity; information that does not need to be checked by the user. When using such information ATCOs will not think twice whether it is correct or not. On the other hand, some source of information is not as trustworthy. These are mainly sources related to prediction |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| | |
|---|---|
| | or some sort of forecast (e.g. weather forecast, conflict detection) and those predictions might not, and often do not, come true.<br><br>Combining both types of information in one display may cause ATCOs to become annoyed or to disregard information deemed unreliable altogether. Checking the plausibility of a tool's result increases the mental workload of ATCOs, which is especially perilous in traffic situations with many complex interactions. With increased adoption of ML-based tools, it can only be expected that ATCOs will have to perform an increasing number of plausibility checks of results provided by different tools in the future. |
| **Does AISA inform the ATC based on a procedure (oral or in the screen)?** | In the PoC system, AISA cannot perform actions, but rather provide suggestions to the ATC. Depending on the tasks, AISA can suggest improvements to the quality of service, it can provide suggestions in terms of priority for conflict resolution, it can provide reminders for assuming or transferring the flights, it can provide alerts if an aircraft is non-conforming to plan or instructions, etc. It has been experimented the oral input by ATC but it demands further analysis in the future system. Besides, it should be evaluated jointly with ATC which is the best way to provide the information (oral or displayed). |
| **Can the ATCO understand the reasoning followed by AISA based on the KG to provide some information?** | Reasoning can be provided for the tasks that were done in Prolog. All tasks can be done in such way but not all of them will be done in Prolog during the project. Explanations provided by Prolog are not easily readable therefore additional effort should be done to make a system which can explain in simple, human-readable, terms what the reasoning was. |
| **Can AISA provide the reasoning explanation in real time?** | The reasoning in the PoC, and the explanation thereof, is not done in real time. AISA itself is working slower than real-time, and there are significant hurdles to be overcome before it can work in real-time. In the future development, it should be identified jointly with the ATC what reasoning can be provided in real time and which can wait until the post-processing. |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

# 4 Hazard identification and analysis

To identify the different hazards, the study has been divided into three main areas.

- Machine Learning,
- Knowledge Engineering,
- ATC Tools.

Once the various areas related with possible failures that may arise in the system have been studied, the final analysis corresponds to the AISA system as a whole, and it is where all hazards related to that system are collected. This fourth area is:

- AISA system.

Besides, each area is grouped into different families. In the area of ML, the aim is to look for hazards that may arise once the ML modules are developed. For example, there could be hazards while implementing modules due to bad training data, just because the ones used were inappropriate or the training algorithm was incorrect. Once the hazards have been found, it has been tried to group each of them into the following two families described in Table 5.

**Table 5. Hazard's family: Machine Learning.**

| Family | Description |
|---|---|
| **Data issues** | This family refers to problems with the data sets used to train the ML models. Errors in data sets could lead to unreliable ML models as they have learned from erroneous information. |
| **Model's performance** | This family refers to problems with the performance of ML models. It can include those hazards related with the ML model's output and also how they adapt to the scenario. |

Secondly, hazards related to 'Knowledge Engineering' have also been studied, trying to find different risks from this technology that could affect the system in terms of information, processing, storage, and management. As explained before, knowledge engineering creates rules that apply to data in order to imitate the process of human reasoning, and during this process several hazards can appear while managing all needed data. These hazards are grouped into five families described in Table 6.

The last area analysed covers ATC tools, which refers to threats that could be found when humans make use of various ATC tools that currently exist. Until now, these have also been grouped into various families, which can be seen in detail in Table 7.

These three areas constitute the pillars of the AISA system. The last step is to evaluate the AISA system to identify different hazards related to the previous areas mentioned and studied that may appear depending on the development of the system. Hazards have been linked to 5 families, as described in Table 8.



EUROPEAN PARTNERSHIP

Co-funded by the European Union

**Table 6. Hazard's family: Knowledge Engineering.**

| Family | Description |
|---|---|
| **Data Quality** | Data quality is concerned with the accuracy and completeness of the data, and it needs to be suitable for its intended uses. |
| **Design** | Aspects related with an incorrect design of the system: both ontologies and Knowledge Graph. |
| **Schema Modelling** | Problems related to the organization or structure of a database. A flawed schema impacts negatively on information quality. |
| **System** | Including hazards related to the system itself and its adaptation to changes. |

**Table 7. Hazard's family: ATC Tools.**

| Family | Description |
|---|---|
| **Design** | This family refers to risks derived from errors in the design of the ATC tools. The system has not been designed with the correct parameters or cannot achieve the correct performance. |
| **Functionality** | Process or task that has been specified to be performed by the ATC tool. |
| **Use of models and decision making** | This family refers to discrepancies between the use of the tools, the purpose for which they have been designed, and the way that ATC uses them. |

**Table 8. Hazard's family: AISA system.**

| Family | Description |
|---|---|
| **Information source** | The sources from which the data sets used by AISA are derived are numerous and varied. Knowledge graph and ML models make use of them and therefore it is important to avoid aspects such as data duplication or heterogeneity. |
| **KG design** | This family refers to risks derived from the design and performance of the AISA KG. It covers the way in which it has been designed to receive, process, and make use of the input and output information. |
| **ML models** | This family refers to hazards related to the responses that ML models provide to the AISA system and their integration. |
| **AISA interface** | Hazards related to software that allows interaction between AISA and the user. |
| **AISA-ATC reasoning** | This family refers to hazards related to AISA reasoning, when AISA reasoning differs from ATCO reasoning or is not expressed in a timely manner. |

Finally, it has been interesting for the analysis to separate the hazards of the AISA system into two levels: 1) Hazards that can be evaluated at the PoC level and 2) Hazards related to the Project level.

In the identification and analysis of hazards in the different areas, the whole AISA consortium has participated. One of the strong points of this identification and analysis is that experts from





Co-funded by the European Union

different fields participated in this analysis that covered all elements of the AISA system. When doing so, there is a parallel increase in knowledge of the AISA system. This is the reason why this step is meaningful for the project. In addition, a continuous analysis of the system is carried out repeatedly in order to find potential hazards that may affect the system in terms of risks. In this way, Table 9, Table 10, Table 11, Table 12 and Table 13 show the hazards identified for ML, Knowledge Engineering, ATC tools and AISA system, respectively.



EUROPEAN PARTNERSHIP

Co-funded by the European Union

**Table 9. Hazards related to the area of Machine Learning.**

| Identifier | Family | Hazard | Description |
|---|---|---|---|
| ML001 | Data | Chaotic data | Chaotic data means its behaviour is impossible to predict. Chaotic systems behaviour depends on multiple variables and are very sensitive to initial conditions. This means that it is very difficult to get a model that correctly predicts this behaviour. A small disturbance or a small change in the initial conditions generates an enormous effect in the future which makes them little observable and difficult to predict (sometimes it is confused with a random behaviour, and therefore it is said that the data that will be obtained can be considered as random). |
| ML002 | Data | Noisy data | The noise in the data can be defined as a random signal that overlap the original behaviour. Noise can hinder their behaviour for ML algorithms to learn it. If the noise cannot be removed, the machine learning algorithm will "think" that the data is random. |
| ML003 | Data | False data | False data are those that do not represent reality because they are not true and therefore can alter the rest of the database. False data disrupt the construction of the ML algorithm by assuming false data as real data, which pervert the predictions of the ML model. |
| ML004 | Data | Incomplete data | Once the ML models have been developed and the user want to apply it to some specific sample, it can occur that some of the data received from the information sources are incomplete. Incomplete data don't provide information on all the variables required by the ML model. |
| ML005 | Data | Unrepresentative data | The database with which the model is trained does not correspond to the reality or the context / field in which it will be applied. As a consequence, failures arise because the model does not recognize the space in which it is being applied. |
| ML006 | Data | Outliers (data out of range) | Normally, ML models create a definition of what they consider statistical behaviour to classify as anomalies all those data that do not fit that definition. This can cause several problems:<br>-These models increase their complexity with the dimension and size of the data.<br>- These models adulterate the sample by trying to consider outliers as normal data by the model. |
| ML007 | Data | Insufficient data | ML models can learn the relationships in a data set from a sample large enough. In the case that data is insufficient, the ML algorithms can establish relationships that will not be correct and will provide a large generalization error. |
| ML008 | Models | Generalization problem (Overfitting or Underfitting) | Both problems are denoted as fit problems and refer to a failure of the model to generalize the knowledge that they want to acquire. This can happen either because there is little training data available (the ML model will not be able to generalize, "underfitting"), or because it is over-adjusted or "overfitted" (the ML model has |

| Identifier | Family | Hazard | Description |
|---|---|---|---|
| | | | memorized the training data so well that it has learned guidelines that are too specific and irrelevant to the new data). |
| ML009 | Models | Lack of scalability and performance degradation | Models become obsolete as data grows. As the model moves towards production, it is typically exposed to larger data volumes and mode of data transport. Over time it will be necessary to monitor the equipment and solve the performance and scalability challenges that will appear. |
| ML010 | Models | Low model accuracy/reliability | The predictability of the ML model is reduced either because the classification capacity is low or because there is a high numerical error. |
| ML011 | Models | Lack of portability of the models | This problem happens when there is a lack of ability to easily migrate a ML model to another environment. This is a typical problem in which there are incompatibilities between the formats of the ML models due to the great disparity of applications that are used nowadays (Python, MatLab, C++, etc.) |
| ML012 | Models | Model interpretability | Interpretable ML models provide valuable information to understand the mathematical reasoning they have followed to make predictions and to understand the pattern of the input/output data |
| ML013 | Models | Real time requirement | ML models, after their implementation, may require a data pre-processing that takes some time, preventing the ML model to be used in real time |
| ML014 | Models | Failure to detect certain anomalies in the outputs / predictions. | ML is unable to identify anomalies in the output predictions, which are significantly different from most of the results. |

**Table 10. Hazards related to the area of Knowledge Engineering.**

| Identifier | Family | Hazard | Description |
|---|---|---|---|
| KE001 | Data Quality | Multiple data sources | Multiple data sources generate semantic heterogeneity which leads to data quality issues. In the case there is no single primary source, the information needed must be gathered through data accumulation. Heterogeneous data must be eliminated because when extracting data from different sources, the same data can be represented differently. |
| KE002 | Data Quality | Use of different representation formats in data sources | The information coming from certain sources need to be transformed from one format to another in order to be used in certain parts of the system. This can lead to incompatibilities in data quality. |
| KE003 | Data Quality | Changes in source systems | Unexpected changes in source systems cause Data Quality problems. The system requires certain information for its operation from some specific sources, if there have been changes, this may affect the KG. |
| KE004 | Data Quality | Data staging ETL | The data staging area is the place where all 'grooming' is done on data after it is culled from the source systems. Staging and ETL (Extraction, Transformation and Loading) phase is considered to be most crucial stage of data warehousing and different risk can be found there, such as:<br>- “lack of capturing only changes in source files”<br>- “Improper extraction of data to the required fields”<br>- “Loss of data during the ETL process (rejected records)” |
| KE005 | Design | Misunderstanding of the domain | The fact that air navigation experts do not participate in the design or construction of both the ontologies and the knowledge graph constitutes a problem since they do not have complete knowledge of the information that needs to be incorporated into the KG. For example, dependencies between classes (such as subclass relationships) are modelled incorrectly leading to wrong conclusions and query results. |
| KE006 | Design | Limitations the level of detail | When designing an KG, the level of detail at should be clear enough. It is important to know the level of detail for the processes of interrogation, information retrieval, knowledge discovery, etc. It is necessary to know up to what level of detail the system is able to understand and with what detail it will provide its response. In the case there are similar classes, it can be difficult to understand the difference between those classes. |
| KE007 | Design | Lack of descriptions | An ontology is designed by a teamwork different from the one that could use the ontology. Therefore, properties, relation and details that have been considering while designing this ontology must be clearly described. This is important to those who have not been present during the design. |
| KE008 | Schema modelling | Rules and queries are out-of-synchronisation with the KG schema | In case of updates to the KG schema, the knowledge engineer has to update the queries and rules that operate on top of that schema as well. If it doesn't happen this way, failure to do so leads to wrong results to conclusions. |
| KE009 | Schema modelling | Wrong assumptions about data quality | When formulating rules and queries the knowledge engineer may make wrong assumptions about data quality (especially completeness) in the KG leading to wrong results and conclusions. |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| | | | |
|---|---|---|---|
| KE010 | Schema modelling | Incorrect meaning of the domain | When formulating rules and queries the knowledge engineer may make wrong assumptions about the meaning of the domain model (KG schema) and hence about the meaning of the data. |
| KE011 | Schema modelling | Incomplete KG schema | Rules and queries may be incomplete or incorrect leading to missing conclusions or wrong conclusions. |
| KE012 | System | Semantic interoperability problems | Semantic interoperability problems may arise and this will make it impossible to make good use of the information available. In the case there is not only one common view of all the data, it is not possible to formulate queries in the different sources. |
| KE013 | System | Scalability | It is the adaptability and response of a system with respect to its performance as it changes its size or configuration to adapt to changing circumstances. This is a weak point because they are designed especially for single server architectures, growth is a (mathematical) challenge. The main problem of the knowledge graph is its construction, since as it gets bigger, more relationships and more elements are needed to be able to work. There are two problems:<br>- This requires time and work to prepare and adapt the knowledge graph.<br>- Once built, an associated problem will be the expansion of the knowledge graph as it progresses in its development. |
| KE014 | System | Lack of response-time with third-party SW | The lack of response-time guarantees of third-party SW used may lead to a miss of critical events. Lack of response time guarantees from the third-party software used may result in the loss of critical events. The system requires information from other software to finalize its predictions, and if this information arrives late, the results will not be issued in real time. This will lead to omitting important circumstances. |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

**Table 11. Hazards related to the area of ATC Tools.**

| Identifier | Family | Hazard | Description |
|---|---|---|---|
| ATC001 | Design | Capacity / demand balance during design for operation | The system has been designed to manage a specific number of aircraft and this capacity must be in accordance with the expected future demand. In case that the future demand is not considered, the system obtained won't be able to manage the number of aircraft that could exist in the future. Namely, the system has been designed for a demand that does not correspond to the expected demand in the future. |
| ATC002 | Design | Design performance error | The ATCO works without knowing the optimal error with which the system has been designed. Without this information, the system may create doubts once modifications to the system are required, since it is not known with which error the system is working. |
| ATC003 | Design | Performance degradation | The system degrades its performance throughout its life cycle. Therefore, when making use continuously of the model, it can be noticed that it loses a significant discriminatory power over time. It means, its speed of detection or action can become slower. |
| ATC004 | Design | Insufficient learning feedback loop | Some errors may be detected in one implementation but not corrected for the next one, giving rise to errors that had already been detected but not corrected. |
| ATC005 | Functionality | Conflict alert | The system must be able to warn about conflicts. The problem is when this alert does not work correctly just because they are false-alarms (it warns of a conflict that then does not take place) or due to missed-alarms (the system does not warn of a conflict that finally occurs). |
| ATC006 | Functionality | Compliance monitoring | The system does not detect correctly deviations from the planned route through the airspace, either due to not being able to analyse the conformity of the actual flight level with the last level assigned by the controller or the conformity of the actual flight heading in front of the last heading assigned by the controller. |
| ATC007 | Functionality | Restricted Area Warning | The system does not detect correctly potential incursions into restricted areas in the medium term. |
| ATC008 | Use of models and decision making | Insufficient user training | ATCO has not been trained enough in order to work with such type of tools based on AI and it could lead to frictions between ATCO and AI system as it is not clear enough the prevalence between ATCO's reasoning and AI's system recommendations.<br>This hazard tackles two situations: 1) ATCOs may follow the recommendation provided by a traffic management tool without question these instructions, or 2) ATCOs don't follow the recommendation given. |
| ATC009 | Use of models and decision making | Failure to consider human factors in decision making | Typically, the error is considered to be associated with the machine, but human reasoning can also be wrong by disregarding the results generated by a system. |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

**Table 12. Hazards related to the area of AISA system PoC.**

| Identifier | Family | Hazard | Description | Operational Context |
|---|---|---|---|---|
| ASPoC001 | Information sources | Loss of ADS-B information | This hazard represents the situation in which one aircraft does not emit or present problems with ADS-B information, which implies that AISA (KG and ML modules) cannot use ADS-B information as inputs. | ADS-B information is used by KG to update the information of the aircraft and by ML modules to perform the predictions.<br>It does not affect to monitoring issues because ATCOs employ the radar information. |
| ASPoC002 | Information sources | Metadata management is out of date | The metadata with which AISA (KG and ML) is going to work does not represent the current scenario (air traffic flow patterns) because is not updated, i.e., the data that has been entered and builds AISA knowledge does not represent the current situation. | Metadata is used by AISA to build the knowledge available for the system (KG) or to analyse the feasibility of the inputs/outputs of ML modules. In the case the metadata is not update, the information/predictions provided could not match with the real scenario. This hazard covers a problem of continuity and integrity. |
| ASPoC003 | KG Design | Information integration | The information coming from the AIXM / FIXM has to be transformed from UML to RDFS / SHACL and this can lead to incompatibilities. | The IT engineers have developed a process to transform the information coming from AIXM/FIXM to RDFS/SHACL. This process can fail which will inhibit the transformation and the data feed to the KG. |
| ASPoC004 | ML models | Invalid input data for the ML model | This hazard addresses a problem when the inputs for predictions out of the scope of the training set. It means that this input data has no resemblance with the one used to train the model. | The problem once AISA model receives invalid input data for the ML modules is twofold: 1) ML models do not identify them as invalid inputs and provide an invalid prediction without knowing it, and 2) ML identifies them and do not provide predictions but the safety barrier is working properly. |
| ASPoC005 | ML models | Lack of trustworthiness in ML modules | This hazard is related when ML works correctly (fulfilling the accuracy expected) but some of them are not accurate or does not provide the expected result. | When any of the ML modules give a faulty/non-accurate prediction, this led to serious errors and mistrust from the ATCO once he/she notifies that it is not working correctly. For instance:<br>- CD predicts a minimum separation of 10 NM when it should be 5 NM. This isolated prediction is wrong but the set of ML predictions work under expected performance.<br>- It is expected that ML work 95% with some performance and what happens with predictions from the other 5%? |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| Identifier | Family | Hazard | Description | Operational Context |
|---|---|---|---|---|
| ASPoC006 | ML models | AISA system as barrier for ML erroneous predictions | The KG analysis statistically the feasibility of the outputs from the ML modules, based on the monitoring-rules previously implemented.<br>When a ML module generates a false or erroneous prediction, it could happen that the KG does not identify it as erroneous or false. | When any of the ML modules give erroneous/faulty predictions to the KG, it analyses it and does not identify it as erroneous/faulty. By the end, the ATCOs receive an erroneous prediction and may lead to a critical situation. |
| ASPoC007 | AISA-ATC reasoning | AISA-ATCO misunderstandings or distractions | The AISA knowledge must be well defined in advance to clearly know the extension where AISA can provide assistance, in what areas or situations. AISA provide information that is not timely correct and can distract the ATCO's attention. | The problem arises when ATCOs expects AISA some information that is not considered in the AISA knowledge or AISA provides information that is not expected. This relates to a problem of misunderstanding: ATCOs expect some information, AISA provides another type of info. |
| ASPoC008 | AISA-ATC reasoning | Divergence in the AISA-ATCO reasoning | AISA provides information/prediction or act in a way that is different from something expected from the ATCO perspective. | Situations will arise in which AISA and ATCO do not agree on the reasoning, giving rise to situations of uncertainty due to not knowing which is the correct or most appropriate reasoning in a given situation. This increase the ATCO's workload since they have to analyse what AISA is doing. |
| ASPoC009 | AISA-ATC reasoning | AISA is unable to explain the reasoning | Although it is intended that AISA avoids the "black box" effect, it is possible that at some point it will not be able to explain the reason for the answer it is giving. | This is a twofold problem: on the one hand, it will increase the ATCO workload in the case AISA cannot explain its reasoning and, on the other hand, it could lead to the risk of mistrust from the ATCO's perspective |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

**Table 13. Hazards related to the area of AISA system PL.**

| Identifier | Family | Hazard | Description | Operational Context |
|---|---|---|---|---|
| ASPL001 | Information sources | Loss of ADS-B information | This hazard represents the situation in which one aircraft does not emit or present problems with ADS-B information, which implies that AISA (KG and ML modules) cannot use ADS-B information as inputs. | ADS-B information is used by KG to update the information of the aircraft and by ML modules to perform the predictions.<br>It does not affect to monitoring issues because ATCOs employ the radar information. |
| ASPL002 | Information sources | Different surveillance sources of information | This hazard refers to the problem of having different sources of information about the same element (e.g., ADS-B and radar on aircraft positioning). | The problem arises when the ATCO does not know what source of surveillance information is the primary and which is the support one. The information obtained from both of them can improve the surveillance capacity but can generate compatibility issues in the system in the case it is not clear how AISA uses them. |
| ASPL003 | Information sources | The AIP information is out of date or wrong | AISA system requires certain information from the AIP for its operation. This hazard relates with the last update or modifications of the information from the AIP. | If it is not update, and there have been changes, this may affect the KG and ML modules and their respective predictions.<br>In the upper airspace, the interesting AIP information are the boundaries of the airspace sector and airways. If ML modules are making the predictions and giving the information based on obsolete AIP information, this could lead to erroneous predictions. For example, if airspace boundaries have changed but this information is not updated, CD module cannot perform the prediction correctly because it is trained in another scenario. |
| ASPL004 | Information sources | Metadata management is out of date | The metadata with which AISA (KG and ML) is going to work does not represent the current scenario (air traffic flow patterns) because is not updated, i.e., the data that has been entered and builds AISA knowledge does not represent the current situation. | Metadata is used by AISA to build the knowledge available for the system (KG) or to analyse the feasibility of the inputs/outputs of ML modules. In the case the metadata is not update, the information/predictions provided could not match with the real scenario. This hazard covers a problem of continuity and integrity. |
| ASPL005 | Information sources | Heterogeneity of data producers | As there is no single authoritative flight source that can be found in any of the central data sources, flight information must be gathered through data accumulation. | What generates duplication and heterogeneity in the data and, if a good data filtering is not carried out, it could affect the development of the ontology by introducing errors from the ontology design. Heterogeneous data must be eliminated because, when extracting data from different sources, the same data can be represented differently. |
| ASPL006 | Information sources | Lack of additional information in the data | AISA system does not receive all the flight information required. Flight information is constructed by accumulating data from multiple sources to synthesise the properties of a flight instance and link it appropriately. | The problem arises when these sources of information do not provide all the necessary data, as may be the case for ADS-B. It is the main source of information used to initiate the flight data fusion process but the ADS-B files are missing additional information: aircraft registration, aircraft manufacturer and airline, among other elements. |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| Identifier | Family | Hazard | Description | Operational Context |
|---|---|---|---|---|
| ASPL007 | Information sources | Problems in the representation of a flight path | A trajectory in an ontology is represented as a sequence of explicit instances of track points. Each tracking point corresponds to a specific reporting time when the speed of an aircraft (its latitude, longitude and altitude) are captured and transmitted to ground systems. | This representation, while adequate for many needs, is detailed and leads to a proliferation of trace points that weakens the efficiency of SPARQL query responses. |
| ASPL008 | KG Design | Little or excess information | When the ATCO requests some type of information or the system identifies that he needs it, it may happen that they receive not necessary or little information when generating this information. | This situation will produce an extra workload to the controller due to the fact of having to do extra reasoning when receiving information that is not required (as well as losing time while having to pay attention to all the information given). |
| ASPL009 | KG Design | AISA knowledge limit | The knowledge limits refer to what information can be extracted from AISA system. The AISA knowledge must be well defined in advance to clearly know the extension where AISA can provide assistance, in what areas or situations. | The problem arises when ATCOs ask AISA some information that is not considered in the AISA knowledge, or actions could be requested to AISA but it cannot provide any type of help. This problem can generate an increase on ATC workload and mistrust on AISA system. |
| ASPL010 | KG Design | Problems with information processing capacity | As the knowledge Graph increases its domain, its information store and its ability to process information is slower and less useful in real time. | In particular, reasoning in the KG (i.e., executions of rules and queries) occurs in AISA in fixed time intervals.<br>In cases of high workload, the system (KG of AISA) may become unable to complete the execution of rules and queries in time. |
| ASPL011 | KG Design | Timeframe too large | The refreshing or updating time of the AISA system is too high, e.g., since the moment AISA provides a result, until it recalculates with the updated information, the updating time does not match the ATCOs need. | This aspect can lead to omitting situations that can affect the ATCOs decision-making process or to provide the information delayed. |
| ASPL012 | KG Design | Scalability issues | The capacity of the system may not be sufficient to keep up with the increases in data and knowledge that it will experience, and its performance may decrease significantly. | If the system does not performance well with scalable systems, it is a problem of design. In a design phases, one of the requirements must be that the system must to be able to handle some specific number of queries at specific time intervals, or number of aircraft, etc. |
| ASPL013 | KG Design | Information integration | The information coming from the AIXM / FIXM has to be transformed from UML to RDFS / SHACL and this can lead to incompatibilities. | The IT engineers have developed a process to transform the information coming from AIXM/FIXM to RDFS/SHACL. This process can fail which will inhibit the transformation and the data feed to the KG. |
| ASPL014 | KG Design | Information identification problems | This is a problem of repetition/duplication of information in different timeframes. When different information with the same identification is | In case that past information is not eliminated and is repeated, can generate incompatibilities with the one that should be used in the present. |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| Identifier | Family | Hazard | Description | Operational Context |
|---|---|---|---|---|
| | | | generated, such as for example when the same aircraft with the same identification flies regularly. Each time it flies, different information with the same identification is generated. | |
| ASPL015 | KG Design | Incomplete domain model (and KG schema) | AISA Situational awareness is the knowledge engine to make it act as a human. If domain model is not complete, it will fail to make deductions due to missing information. | During the development of AISA system, the AISA members have developed some reasoning engineering based on ATCOs expertise to achieve situational awareness. There is a gap between the reasoning engineering put into the KG and the overall ATCO knowledge. Two problems can arise: 1) ATCOs ask AISA some information that is not considered in the AISA knowledge, and 2) ATCOs request actions to AISA but it cannot provide any type of help. |
| ASPL016 | ML models | Invalid input data for the ML model | This hazard addresses a problem when the inputs for predictions out of the scope of the training set. It means that this input data has no resemblance with the one used to train the model. | The problem once AISA model receives invalid input data for the ML modules is twofold: 1) ML models do not identify them as invalid inputs and provide an invalid prediction without knowing it, and 2) ML identifies them and do not provide predictions but the safety barrier is working properly. |
| ASPL017 | ML models | Unavailability | When any of the ML modules is not working, that means it is not available at that moment and it is not possible to take advantage of its predictions. | The ATCO must know the ML model is out of service and AISA cannot provide predictions. This is a problem of availability. |
| ASPL018 | ML models | Lack of trustworthiness in ML modules | This hazard is related when ML works correctly (fulfilling the accuracy expected) but some of them are not accurate or does not provide the expected result. | When any of the ML modules give a faulty/non-accurate prediction, this led to serious errors and mistrust from the ATCO once he/she notifies that it is not working correctly. For instance:<br>- CD predicts a minimum separation of 10 NM when it should be 5 NM. This isolated prediction is wrong but the set of ML predictions work under expected performance.<br>- It is expected that ML work 95% with some performance and what happens with predictions from the other 5%? |
| ASPL019 | ML models | AISA system as barrier for ML erroneous predictions | The KG analysis statistically the feasibility of the outputs from the ML modules, based on the monitoring-rules previously implemented.<br>When a ML module generates a false or erroneous prediction, it could happen that the KG does not identify it as erroneous or false. | When any of the ML modules give erroneous/faulty predictions to the KG, it analyses it and does not identify it as erroneous/faulty. By the end, the ATCOs receive an erroneous prediction and may lead to a critical situation. |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| Identifier | Family | Hazard | Description | Operational Context |
|---|---|---|---|---|
| ASPL020 | AISA interface | Visualization of confusing information | The way in which the information is presented to the ATCO could be confusing and not clear, which may cause problems to the ATCO's reasoning. | This is a problem of ergonomics that affect the labour of ATCOs |
| ASPL021 | AISA interface | Human-machine non-clear situations | Problems with the distribution of the SA between human-machine. | In this hazard, there are some situations (not identified during the design phase) that it is not clear the distribution of tasks. This increases the ATCO workload because it has to analyse why AISA does/does not some specific task as the ATCO expected. |
| ASPL022 | AISA interface | Conflict alert timing | AISA provides information too early to the ATCOs about conflict detection. | Conflict inputs can be provided too early (e.g. due to far range), ATCOs need to look for the aircraft AISA was informing which implies an additional workload for ATCOs. Moreover, to provide information by voice/audible is a problem for ATCO because it means an urgent action and they have to react and stop doing what they are doing. |
| ASPL023 | AISA-ATC reasoning | AISA-ATCO misunderstandings or distractions | The AISA knowledge must be well defined in advance to clearly know the extension where AISA can provide assistance, in what areas or situations. AISA provide information that is not timely correct and can distract the ATCO's attention. | The problem arises when ATCOs expects AISA some information that is not considered in the AISA knowledge or AISA provides information that is not expected. This relates to a problem of misunderstanding: ATCOs expect some information, AISA provide another type of info. |
| ASPL024 | AISA-ATC reasoning | Divergence in the AISA-ATCO reasoning | AISA provides information/prediction or act in a way that is different from something expected from the ATCO perspective. | Situations will arise in which AISA and ATCO do not agree on the reasoning, giving rise to situations of uncertainty due to not knowing which is the correct or most appropriate reasoning in a given situation. This increase the ATCO's workload since they have to analyse what AISA is doing. |
| ASPL025 | AISA-ATC reasoning | Time spent explaining the reasoning | AISA should be able to explain each of the decisions it makes in the case ATCOs demands it. However, this explanation can be time-consuming for the ATCO due to the time required for AISA to explain the reasoning. | ATCO, which must pay attention to the reasoning said to understand the decision that has been made.<br>When little information is available it is easy to work with this and perform calculations or inference. However, when there is a lot of information, it is or it can become difficult to work in real time with this type of system. |
| ASPL026 | AISA-ATC reasoning | AISA is unable to explain the reasoning | Although it is intended that AISA avoids the "black box" effect, it is possible that, at some point, it will not be able to explain the reason for the answer it is giving. | This is a twofold problem: on the one hand, it will increase the ATCO workload in the case AISA cannot explain its reasoning and on the other hand, it could lead to the risk of mistrust from the ATCO's perspective. |
| ASPL027 | AISA-ATC reasoning | Insufficient user training | ATCO has not been trained enough in order to work with such type of tools based on AI and it could lead to frictions between ATCO and AI system as it is not | This hazard tackles two situations: 1) ATCOs may follow the recommendation provided by a traffic management tool without question these instructions, or 2) ATCOs don't follow the recommendation given. |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| Identifier | Family | Hazard | Description | Operational Context |
|---|---|---|---|---|
| | | | clear enough the prevalence between ATCO's reasoning and AI's system recommendations. | |
| ASPL028 | AISA-ATC reasoning | Not considering human factors in AISA decision making | This problem is related with the reasoning that the ATCO must do when one information is received from AISA. | Once some information is provided to the ATCO from AISA, it is considered that they have to work together as a Team Situational Awareness. This implies that the system must be developed considering human factors in the AISA decision-making process. |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

# 5 Risk assessment

## 5.1 Initial risk quantification

Once hazards have been identified and evaluated by the AISA consortium, risks have been analysed based on the quantification of likelihood and severity. As explained in Section 1.2, the risk assessment session follows the ICAO methodology and, in particular, the scales proposed in Table 1 and Table 2. Then, every hazard receives a number (1-5) based on the level of likelihood, a letter (A-E) indicating the degree of severity of its consequences, and, finally, the combination of them constitutes the risk quantification. In order to facilitate the interpretation of the results, the results are shown in different tables following the colour scale proposed by the ICAO tolerability matrix.

Table 14 shows the values of the risk assessment for ML hazards. It can be shown that one hazard is denoted as acceptable risk, 13 are tolerable risks, and there are no inacceptable risks.

**Table 14. Likelihood and severity of Hazards related to the area of Machine Learning.**

| Identifier | Family | Hazard | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|
| ML001 | Data | Chaotic data | 3 | D | 3D |
| ML002 | Data | Noisy data | 3 | D | 3D |
| ML003 | Data | False data | 1 | C | 1C |
| ML004 | Data | Incomplete data | 4 | E | 4E |
| ML005 | Data | Unrepresentative data | 2 | C | 2C |
| ML006 | Data | Outliers (data out of range) | 4 | E | 4E |
| ML007 | Data | Insufficient data | 4 | D | 4D |
| ML008 | Models | Generalization problem (Overfitting or Underfitting) | 3 | D | 3D |
| ML009 | Models | Lack of scalability and performance degradation | 4 | D | 4D |
| ML010 | Models | Low model accuracy/reliability | 2 | B | 2B |
| ML011 | Models | Lack of portability of the models | 2 | B | 2B |
| ML012 | Models | Model interpretability | 3 | D | 3D |
| ML013 | Models | Real time requirement | 2 | C | 2C |
| ML014 | Models | Failure to detect certain anomalies in the outputs / predictions. | 3 | C | 3C |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

Table 15 shows the values of the risk assessment for Knowledge Engineering. It can be shown that 5 hazards are denoted as acceptable risks, 9 are tolerable risks, and there are no unacceptable risks.

**Table 15. Likelihood and severity of Hazards related to the area of Knowledge Engineering.**

| Identifier | Family | Hazard | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|
| KE001 | Data Quality | Multiple data sources | 2 | D | 2D |
| KE002 | Data Quality | Use of different representation formats in data sources | 3 | D | 3D |
| KE003 | Data Quality | Changes in source systems | 3 | D | 3D |
| KE004 | Data Quality | Data staging ETL | 2 | D | 2D |
| KE005 | Design | Misunderstanding of the domain | 2 | C | 2C |
| KE006 | Design | Limitations the level of detail | 2 | D | 2D |
| KE007 | Design | Lack of descriptions | 2 | D | 2D |
| KE008 | Schema modelling | Rules and queries are out-of-synchronisation with the KG schema | 2 | B | 2B |
| KE009 | Schema modelling | Wrong assumptions about data quality | 2 | C | 2C |
| KE010 | Schema modelling | Incorrect meaning of the domain | 2 | C | 2C |
| KE011 | Schema modelling | Incomplete KG schema | 2 | C | 2C |
| KE012 | System | Semantic interoperability problems | 2 | D | 2D |
| KE013 | System | Scalability | 2 | B | 2B |
| KE014 | System | Lack of response-time with third-party SW | 2 | C | 2C |

Table 16 shows the values of the risk assessment for the ATC tools. It can be shown that no hazards are denoted as acceptable risks, 9 are tolerable risks, and there are no unacceptable risks.





Co-funded by the European Union

Table 16. Likelihood and severity of Hazards related to the area of ATC Tools.

| Identifier | Family | Hazard | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|
| ATC001 | Design | Capacity / demand balance during design for operation | 2 | C | 2C |
| ATC002 | Design | Design performance error | 2 | B | 2B |
| ATC003 | Design | Performance degradation | 3 | C | 3C |
| ATC004 | Design | Insufficient learning feedback loop | 2 | B | 2B |
| ATC005 | Functionality | Conflict alert | 2 | B | 2B |
| ATC006 | Functionality | Compliance monitoring | 2 | B | 2B |
| ATC007 | Functionality | Restricted Area Warning | 2 | C | 2C |
| ATC008 | Use of models and decision making | Insufficient user training | 2 | B | 2B |
| ATC009 | Use of models and decision making | Failure to consider human factors in decision making | 2 | B | 2B |

Table 17 shows the values of the risk assessment for the AISA PoC level. It can be shown that 2 hazards are denoted as acceptable risks, 6 are tolerable risks, and there is 1 unacceptable risk.

Table 17. Likelihood and severity of Hazards related to the area of AISA PoC level.

| Identifier | Family | Hazard | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|
| ASPoC001 | Information sources | Loss of ADS-B information | 3 | D | 3D |
| ASPoC002 | Information sources | Metadata management is out of date | 3 | B | 3B |
| ASPoC003 | KG Design | Information integration | 2 | D | 2D |
| ASPoC004 | ML models | Invalid input data for the ML model | 3 | D | 3D |
| ASPoC005 | ML models | Lack of trustworthiness in ML modules | 4 | B | 4B |
| ASPoC006 | ML models | AISA system as barrier for ML erroneous predictions | 4 | C | 4C |
| ASPoC007 | AISA-ATC reasoning | AISA-ATCO misunderstandings or distractions | 4 | C | 4C |





Co-funded by the European Union

| | | | | | |
|---|---|---|---|---|---|
| ASPoC008 | AISA-ATC reasoning | Divergence in the AISA-ATCO reasoning | 4 | D | 4D |
| ASPoC009 | AISA-ATC reasoning | AISA is unable to explain the reasoning | 2 | D | 2D |

Table 18 shows the values of the risk assessment for the AISA project level. It can be shown that 9 hazards are denoted as acceptable risks, 18 are tolerable risks, and there is 1 unacceptable risk.

**Table 18. Likelihood and severity of Hazards related to the area of AISA Project level.**

| Identifier | Family | Hazard | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|
| ASPL001 | Information sources | Loss of ADS-B information | 3 | D | 3D |
| ASPL002 | Information sources | Different surveillance sources of information | 2 | C | 2C |
| ASPL003 | Information sources | The AIP information is out of date or wrong | 2 | B | 2B |
| ASPL004 | Information sources | Metadata management is out of date | 3 | B | 3B |
| ASPL005 | Information sources | Heterogeneity of data producers | 2 | D | 2D |
| ASPL006 | Information sources | Lack of additional information in the data | 3 | E | 3E |
| ASPL007 | Information sources | Problems in the representation of a flight path | 2 | E | 2E |
| ASPL008 | KG Design | Little or excess information | 2 | E | 2E |
| ASPL009 | KG Design | AISA knowledge limit | 3 | C | 3C |
| ASPL010 | KG Design | Problems with information processing capacity | 2 | B | 2B |
| ASPL011 | KG Design | Timeframe too large | 3 | E | 3E |
| ASPL012 | KG Design | Scalability issues | 3 | B | 3B |
| ASPL013 | KG Design | Information integration | 2 | D | 2D |
| ASPL014 | KG Design | Information identification problems | 2 | D | 2D |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| | | | | | |
|---|---|---|---|---|---|
| ASPL015 | KG Design | Incomplete domain model (and KG schema) | 3 | D | 3D |
| ASPL016 | ML models | Invalid input data for the ML model | 3 | D | 3D |
| ASPL017 | ML models | Unavailability | 2 | B | 2B |
| ASPL018 | ML models | Lack of trustworthiness in ML modules | 4 | B | 4B |
| ASPL019 | ML models | AISA system as barrier for ML erroneous predictions | 4 | C | 4C |
| ASPL020 | AISA interface | Visualization of confusing information | 2 | D | 2D |
| ASPL021 | AISA interface | Human-machine non-clear situations | 3 | C | 3C |
| ASPL022 | AISA interface | Conflict alert timing | 3 | C | 3C |
| ASPL023 | AISA-ATC reasoning | AISA-ATCO misunderstandings or distractions | 4 | C | 4C |
| ASPL024 | AISA-ATC reasoning | Divergence in the AISA-ATCO reasoning | 4 | D | 4D |
| ASPL025 | AISA-ATC reasoning | Time spent explaining the reasoning | 2 | C | 2C |
| ASPL026 | AISA-ATC reasoning | AISA is unable to explain the reasoning | 2 | D | 2D |
| ASPL027 | AISA-ATC reasoning | Insufficient user training | 2 | B | 2B |
| ASPL028 | AISA-ATC reasoning | Not considering human factors in AISA decision making | 2 | B | 2B |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

## 5.2 Risk mitigation and re-quantification

The purpose of the last step of the risk assessment is to propose strategies that will lead to a reduction in the risk of the previous hazards identified. The process continues by proposing mitigation measures and finally a re-quantification of the risk quantification considering the mitigation measures proposed. The tables in the previous section show, according to the colours, which hazards are riskier. Those shaded in red are at an unacceptable level of tolerability and so, searching and finding mitigating measures for them is essential in order to reduce the risk. However, those shaded yellow are in a tolerable region. This does not mean that they have not been looked away from them. Instead, they should be reviewed from time to time to ensure that the risk does not increase and become part of the unacceptable region. Only those in green can be overlooked at the moment, as they are at an acceptable level in terms of safety.

In the case of the AISA project, several risks have been identified and measures have been proposed to mitigate them. As can be seen in the following tables, a proposal of mitigation measures has been carried out for all hazards, regardless of the level of tolerability, and the re-quantification of the risk considering how the mitigation measures could reduce the initial risks.

Some of these proposals are only described, and further study is needed to establish more precise measures. On the other hand, there are measures that should be studied to analyse if they are cost effective because, even if the action could reduce the risk, it should be studied how expensive and affordable it is to introduce the mitigation measures. For all these reasons, it is emphasised that, although a wide variety of results have been obtained from the evaluation carried out, not all of them are directly applicable. It will need to be studied in more detail at a later stage to ensure that they are reliable, efficient, and cost-effective proposals.

The main problems during the completion of the risk assessment have been identifying the correct mitigation measures (due to the novelty of the technologies and the lack of knowledge in some areas). The boundaries between the AISA PoC and the AISA project are problematic because, in several areas, it has been difficult to understand the differences between them. In addition, the re-quantification of the likelihood and severity was challenging to evaluate the impact of the mitigation measures on the system in this conceptual stage.

Lastly, the overall results of this mitigation process allowed one to reduce from two unacceptable risks to two tolerable risks. The number of tolerable risks has decreased to 16 and the number of acceptable risks increased to 58. These results confirmed the goal of this work of reducing the risks for the AISA system.



EUROPEAN PARTNERSHIP

Co-funded by the European Union

Table 19. Mitigation measures of Hazards related to the area of Machine Learning.

| Identifier | Family | Hazard | Mitigation | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|---|
| ML001 | Data | Chaotic data | • Evaluate data control, data integration and data exploration until clear data is obtained<br>• Analyse the generalization error to know if the model has "understood" the data patterns or is being random. | 2 | D | 2D |
| ML002 | Data | Noisy data | • Reduce data noise by using algorithms for this purpose.<br>• Analyse the generalization error to know if the model has "understood" the data patterns or is being random. | 2 | D | 2D |
| ML003 | Data | False data | • Define a procedure for obtaining data that ensures the generation of real data<br>• Develop a procedure to notify the identification of false data for a post-analysis | 1 | D | 1D |
| ML004 | Data | Incomplete data | • Evaluate data control, data integration and data exploration until clear data is obtained<br>• Develop an impute method based on algorithms, e.g., control algorithms or machine learning algorithms that predicts what value we are missing by learning from the cases in which we have data.<br>• Develop a procedure to notify the identification of false data for a post-analysis | 2 | E | 2E |
| ML005 | Data | Unrepresentative data | • Check that the data represent reality through statistical analysis before discarding characteristics, and make use of tools that are not biased<br>• Build and develop different databases for different conditions covering the majority of situations | 1 | D | 1D |
| ML006 | Data | Outliers (data out of range) | • Remove outliers from graphic descriptions, and perform statistical tests<br>• Define a procedure for obtaining data that ensures the generation of real data<br>• Develop a procedure to notify the identification of false data for a post-analysis | 2 | E | 2E |
| ML007 | Data | Insufficient data | • Include minimum requirements in the amount of data during the system design<br>• Define a minimum generalization error based on the database size | 2 | D | 2D |
| ML008 | Models | Generalization problem (Overfitting or Underfitting) | • Include minimum requirements in the amount of data during the system design<br>• Define a minimum generalization error based on the database size based on regularization techniques | 1 | D | 1D |

| Identifier | Family | Hazard | Mitigation | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|---|
| ML009 | Models | Lack of scalability and performance degradation | • Include minimum requirements in the amount of data able to handle during the design phase<br>• Monitor and periodically evaluate the model's performance in order to find degradation or bias<br>• Monitor and solve performance and scalability challenges<br>• Develop a consistent approach to production analysis | 2 | E | 2E |
| ML010 | Models | Low model accuracy/reliability | • Include minimum requirements in the threshold to consider the model's reliability during the design phase | 1 | B | 1B |
| ML011 | Models | Lack of portability of the models | • Include this requirement in the system design phase<br>• Analyse the ability to be used based on the technical capabilities of the software or its compatibility with other environments to ensure proper operation | 1 | B | 1B |
| ML012 | Models | Model interpretability | • Include the interpretability of the model in the system design phase<br>• Develop requirements considering final users (ATCOs) | 2 | E | 2E |
| ML013 | Models | Real time requirement | • Include the real-time operation of the model in the system design phase<br>• Define what means real-time operation and to include them as requirements to provide predictions | 1 | C | 1C |
| ML014 | Models | Failure to detect certain anomalies in the outputs / predictions. | • Implement algorithms to detect anomalies based on statistical analysis<br>• Develop a procedure to notify the identification for a post-analysis | 2 | D | 2D |

**Table 20. Mitigation measures of Hazards related to the area of Knowledge Engineering.**

| Identifier | Family | Hazard | Mitigation | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|---|
| KE001 | Data Quality | Multiple data sources | • Filter the data and eliminate duplicate or heterogeneous data<br>• Develop a process to ensure no problems arises from data-sources heterogeneity | 2 | E | 2E |
| KE002 | Data Quality | Use of different representation formats in data sources | • Include it as system requirement during the design phase<br>• Develop a procedure to notify these divergences for a post-analysis | 2 | D | 2D |
| KE003 | Data Quality | Changes in source systems | • Include it as system requirement during the design phase<br>• Develop a procedure to notify these divergences for a post-analysis | 2 | D | 2D |
| KE004 | Data Quality | Data staging ETL | • Include it as system requirement during the design phase<br>• Monitor the consistency of the data quality<br>• Develop a procedure to notify these divergences for a post-analysis | 2 | E | 2E |
| KE005 | Design | Misunderstanding of the domain | • Train the system following ATCO expertise and reasoning and make periodical evaluations<br>• Develop a procedure to notify these divergences for a post-analysis | 1 | C | 1C |
| KE006 | Design | Limitations the level of detail | • The level of detail should be limited to the minimum necessary as system requirement<br>• Document clearly the level of detail and knowledge limits | 2 | E | 2E |
| KE007 | Design | Lack of descriptions | • Clearly describe during the system, design the labels of the classes, properties and relationships to understand the ontology<br>• Document clearly the level of detail and knowledge limits | 1 | D | 1D |
| KE008 | Schema modelling | Rules and queries are out-of-synchronisation with the KG schema | • Develop a procedure to notify these divergences for a post-analysis | 2 | C | 2C |
| KE009 | Schema modelling | Wrong assumptions about data quality | • Train the system following ATCO expertise and reasoning and make periodical evaluations<br>• Develop a procedure to notify these divergences for a post-analysis | 1 | C | 1C |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| Identifier | Family | Hazard | Mitigation | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|---|
| KE010 | Schema modelling | Incorrect meaning of the domain | • Train the system following ATCO expertise and reasoning and make periodical evaluations<br>• Develop a procedure to notify these divergences for a post-analysis | 1 | C | 1C |
| KE011 | Schema modelling | Incomplete KG schema | • Include data control, integration and exploration in the schema model as system requirement during the design phase<br>• Monitor the consistency of the data quality<br>• Develop a procedure to notify these divergences for a post-analysis | 2 | C | 2C |
| KE012 | System | Semantic interoperability problems | • Establish a common description of the data<br>• A consensus must be reached in which the same ATM data is expressed regardless of who they are generated by<br>• Develop a procedure to notify these divergences for a post-analysis | 1 | D | 1D |
| KE013 | System | Scalability | • Include it as system requirement during the design phase, e.g., using hardware or software approaches<br>• Develop an updating process as integrity/monitoring, similar to any aeronautical system | 1 | B | 1B |
| KE014 | System | Lack of response-time with third-party SW | • Include it as a requirement of the system during the design phase, e.g. using hardware or software approaches • To develop an update process as integrity/monitoring, similar to any aeronautical system | 1 | C | 1C |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

**Table 21. Mitigation measures of Hazards related to the area of ATC Tools.**

| Identifier | Family | Hazard | Mitigation | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|---|
| ATC001 | Design | Capacity / demand balance during design for operation | • Carry out a specific study to ensure that the capacity with which the system is going to work corresponds to the one which has been considered<br>• Include it as a system requirement during the design phase | 1 | C | 1C |
| ATC002 | Design | Design performance error | • Include the maximum admissible error as a system requirement<br>• Inform/train ATCO about the expected error from the system | 2 | C | 2C |
| ATC003 | Design | Performance degradation | • Analyse the potential degradation and to develop a process throughout its lifecycle to deal with<br>• Monitor the model performance and inform when it loses its performance | 2 | D | 2D |
| ATC004 | Design | Insufficient learning feedback loop | • Integrate feedback loops and requirements into the development of the system lifecycle<br>• Systematic monitoring and reporting of errors from the system | 1 | B | 1B |
| ATC005 | Functionality | Conflict alert | • Implement specific training must be provided to ATC personnel for the correct use of this tool<br>• Include a visual/audio alert on the CWP if minimum separation are or are predicted to be infringed<br>• Analyse and adjust the time range of the conflict alert threshold<br>• Perform statistical analysis of alerts to identify possible deficiencies in airspace design and ATC procedures<br>• Include a system requirement during the design phase to recalculate the conflict prediction every x seconds to avoid "missed" or false alarms<br>• Action mechanisms and procedures must be designed, in case the failure occurs. | 1 | B | 1B |
| ATC006 | Functionality | Compliance monitoring | • Include the probability of failure as a system requirement during the design phase<br>• Develop a procedure to notify these divergences for a post-analysis. | 2 | C | 2C |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| Identifier | Family | Hazard | Mitigation | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|---|
| ATC007 | Functionality | Restricted Area Warning | • Include the probability of failure as a system requirement during the design phase<br>• Develop a procedure to notify these divergences for a post-analysis.<br>• Database with all PRD areas must be referenced to a period of validity similar to AIRAC cycle | 1 | C | 1C |
| ATC008 | Use of models and decision making | Insufficient user training | • Include it as a system requirement during the implementation phase<br>• Develop a specific user-training process including training about: how the artificial intelligence model works in the general system, how to use the knowledge it generates and how and when to cancel its results. | 1 | B | 1B |
| ATC009 | Use of models and decision making | Failure to consider human factors in decision making | • Include it as a system requirement during the design phase<br>• Develop a whole human factor analysis. | 1 | C | 1C |





Co-funded by the European Union

**Table 22. Mitigation measures of Hazards related to the area of AISA system PoC level.**

| Identifier | Family | Hazard | Mitigation | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|---|
| ASPoC001 | Information sources | Loss of ADS-B information | • Implement an alert that inform ATCO when ADS-B information of one aircraft is lost and remind that ML or KG are not updating the information<br>• Include the radar information in conjunction with ADS-B for KG or ML in order not to be a loss of the ADS-B a loss of the ability of AISA to predict | 2 | D | 2D |
| ASPoC002 | Information sources | Metadata management is out of date | • The characteristics of the metadata (e.g. training sets) considered should be specified to ensure that the airspace and air traffic flows considered are valid to those used in airspace at present.<br>• Develop a covering function about what is considered in the Metadata and what it is out. The goal is to inform the ATCO that the predictions can get worse/not so accurate.<br>• Develop a process to ensure quality of the metadata throughout its life cycle | 2 | D | 2D |
| ASPoC003 | KG Design | Information integration | • Have a list of all the information that is necessary to provide to AISA: conformance tests in real-time.<br>• Generate error reports about the incompatibilities that have been generated for a post-review. | 2 | E | 2E |
| ASPoC004 | ML models | Invalid input data for the ML model | • Establish initial tests to ensure that the input provided to the ML has been trained to evaluate it | 2 | E | 2E |
| ASPoC005 | ML models | Lack of trustworthiness in ML modules | • The validation transition phase maintains the current system in order to compare their results<br>• Implement a certification process for ML modules<br>• Cross checking the results of the machine learning modules in real time with the knowledge graph<br>• Retrain the machine learning model | 3 | D | 3D |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| Identifier | Family | Hazard | Mitigation | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|---|
| ASPoC006 | ML models | AISA system as barrier for ML erroneous predictions | • Implement an algorithm based on the previous information to validate the predictions<br>• Develop a procedure to notify the errors identified and a post-analysis<br>• Keep tracking to the past performance or the plausibility<br>• Cross-checking of the safety barriers considered in AISA system | 3 | D | 3D |
| ASPoC007 | AISA-ATC reasoning | AISA-ATCO misunderstandings or distractions | • Train AISA following ATCO expertise and reasoning and make periodical evaluations<br>• Develop a procedure to notify these divergences for a post-analysis.<br>• Develop an updating process as integrity/monitoring, similar to any aeronautical system | 3 | D | 3D |
| ASPoC008 | AISA-ATC reasoning | Divergence in the AISA-ATCO reasoning | • Train AISA following ATCO expertise and reasoning and make periodical evaluations<br>• Develop a procedure to notify these divergences for a post-analysis.<br>• Develop an updating process as integrity/monitoring, similar to any aeronautical system | 3 | D | 3D |
| ASPoC009 | AISA-ATC reasoning | AISA is unable to explain the reasoning | • Establish a procedure to notify these cases<br>• Review the knowledge of AISA to try to avoid these black-box situations it in the future | 2 | D | 2D |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

**Table 23. Mitigation measures of Hazards related to the area of AISA system Project level.**

| Identifier | Family | Hazard | Mitigation | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|---|
| ASPL001 | Information sources | Loss of ADS-B information | • Implement an alert that inform ATCO when ADS-B information of one aircraft is lost and remind that ML or KG are not updating the information<br>• Include the radar information in conjunction with ADS-B for KG or ML in order not to be a loss of the ADS-B a loss of the ability of AISA to predict | 2 | D | 2D |
| ASPL002 | Information sources | Different surveillance sources of information | • Determine which surveillance means is going to be the primary source. The sources of information should be the same for a better complicity between the reasoning of both (ATCO and AISA)<br>• Inform ATCO when AISA changes the primary means of surveillance system | 2 | B | 2B |
| ASPL003 | Information sources | The AIP information is out of date or wrong | • The AIP information included must be referenced to a period of validity similar to AIRAC cycle | 1 | B | 1B |
| ASPL004 | Information sources | Metadata management is out of date | • The characteristics of the metadata (e.g. training sets) considered should be specified to ensure that the airspace and air traffic flows considered are valid to those used in airspace at present.<br>• Develop a covering function about what is considered in the Metadata and what it is out. The goal is to inform the ATCO that the predictions can get worse/not so accurate.<br>• Develop a process to ensure quality of the metadata throughout its life cycle | 2 | D | 2D |
| ASPL005 | Information sources | Heterogeneity of data producers | • Develop a process to ensure no problems arises from data-sources heterogeneity | 2 | E | 2E |
| ASPL006 | Information sources | Lack of additional information in the data | • Develop a process to ensure the quality of the information and, in the case a data lacks some information, to extract from previous information | 2 | E | 2E |

| Identifier | Family | Hazard | Mitigation | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|---|
| ASPL007 | Information sources | Problems in the representation of a flight path | • Carry out a cost-efficiency analysis, balancing the number of points that would need to be considered at the same time and the computational cost that this would entail with SPARQL queries. | 1 | E | 1E |
| ASPL008 | KG Design | Little or excess information | • Train AISA following ATCO expertise and reasoning and make periodical evaluations<br>• Develop a procedure to notify these divergences for a post-analysis | 1 | E | 1E |
| ASPL009 | KG Design | AISA knowledge limit | • Training courses for ATCOs and other personnel in order to be aware of AISA knowledge limits<br>• Inform ATCOs that the action or information requested is out of AISA knowledge limits | 2 | A | 2A |
| ASPL010 | KG Design | Problems with information processing capacity | • Include it as system requirement during the design phase<br>• Develop a roadmap to increase the efficiency during the lifecycle of the KG | 2 | C | 2C |
| ASPL011 | KG Design | Timeframe too large | • Include it as a system requirement during the design phase<br>• Carry out a study to find the appropriate time frame for updating the system and design it accordingly. | 2 | D | 2D |
| ASPL012 | KG Design | Scalability issues | • Include it as a system requirement during the design phase<br>• Develop a process to increase the scalability during the lifecycle of the KG | 2 | D | 2D |
| ASPL013 | KG Design | Information integration | • Have a list of all the information that is necessary to provide to AISA: conformance tests in real-time.<br>• Generate error reports about the incompatibilities that have been generated for a post-review. | 2 | E | 2E |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| Identifier | Family | Hazard | Mitigation | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|---|
| ASPL014 | KG Design | Information identification problems | • Include it as a system requirement during the design phase<br>• Bulk of information generated by the KG every X time out of the system. | 1 | D | 1D |
| ASPL015 | KG Design | Incomplete domain model (and KG schema) | • Experts should review the domain model in order to find the lack of interaction between aspects or missing data.<br>• Train or inform ATCOs about what reasoning engineering have been introduced in the KG | 2 | C | 2C |
| ASPL016 | ML models | Invalid input data for the ML model | • Establish initial tests to ensure that the input provided to the ML has been trained to evaluate it. | 2 | E | 2E |
| ASPL017 | ML models | Unavailability | • Include the probability of being out of service as a system requirement during the design phase<br>• Develop a procedure to notify the errors identified and a post-analysis | 2 | D | 2D |
| ASPL018 | ML models | Lack of trustworthiness in ML modules | • The validation transition phase maintains the current system in order to compare their results<br>• Implement a certification process for ML modules<br>• Cross checking the results of the machine learning modules in real time with the knowledge graph<br>• Retrain the machine learning model | 3 | D | 3D |
| ASPL019 | ML models | AISA system as barrier for ML erroneous predictions | • Implement an algorithm based on the previous information to validate the predictions<br>• Develop a procedure to notify the errors identified and a post-analysis<br>• Keep tracking to the past performance or the plausibility<br>• Cross-checking of the safety barriers considered in AISA system | 3 | D | 3D |

| Identifier | Family | Hazard | Mitigation | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|---|
| ASPL020 | AISA interface | Visualization of confusing information | • Involve ATCOs during the design phase<br>• Train AISA following ATCO expertise and reasoning and make periodical evaluations | 1 | D | 1D |
| ASPL021 | AISA interface | Human-machine non-clear situations | • Train AISA following ATCO expertise and reasoning and make periodical evaluations<br>• Develop a procedure to notify these divergences for a post-analysis.<br>• Develop an updating process as integrity/monitoring, similar to any aeronautical system | 2 | D | 2D |
| ASPL022 | AISA interface | Conflict alert timing | • Involve ATCOs during the design phase<br>• Provide visualised information about conflict detection that is non-urgent<br>• Include audible alerts only for urgent separation infringements | 2 | D | 2D |
| ASPL023 | AISA-ATC reasoning | AISA-ATCO misunderstandings or distractions | • Train AISA following ATCO expertise and reasoning and make periodical evaluations<br>• Develop a procedure to notify these divergences for a post-analysis.<br>• Develop an updating process as integrity/monitoring, similar to any aeronautical system | 3 | D | 3D |
| ASPL024 | AISA-ATC reasoning | Divergence in the AISA-ATCO reasoning | • Train AISA following ATCO expertise and reasoning and make periodical evaluations<br>• Develop a procedure to notify these divergences for a post-analysis.<br>• Develop an updating process as integrity/monitoring, similar to any aeronautical system | 3 | D | 3D |
| ASPL025 | AISA-ATC reasoning | Time spent explaining the reasoning | • Include a maximum-response time as a system requirement<br>• Establish a procedure which notifies of those events in which the reasoning has not been explained in a suitable time. | 2 | D | 2D |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

| Identifier | Family | Hazard | Mitigation | Likelihood (1-5) | Severity (A-E) | Risk |
|---|---|---|---|---|---|---|
| ASPL026 | AISA-ATC reasoning | AISA is unable to explain the reasoning | • Establish a procedure to notify these cases<br>• Review the knowledge of AISA to try to avoid these black-box situations it in the future | 2 | D | 2D |
| ASPL027 | AISA-ATC reasoning | Insufficient user training | • Include it as a system requirement during the implementation phase<br>• Develop a specific user-training process including training about: how the artificial intelligence model works in the general system, how to use the knowledge it generates and how and when to cancel its results. | 1 | B | 1B |
| ASPL028 | AISA-ATC reasoning | Not considering human factors in AISA decision making | • Include it as a system requirement during the design phase<br>• Develop a whole human factor analysis. | 1 | C | 1C |



EUROPEAN PARTNERSHIP

Co-funded by the European Union

In particular, some recommendations for the further development of AISA emerged during the risk assessment session for particular risks:

- In the case that the severity or likelihood of one hazard acquires equality between two consecutive values, the solution adopted was to characterise it as the most restrictive.
- Audio alerts must be considered only for safety-critical situations (e.g. short-term separation infringement).
- Information and the timing at which the information is provided should be analysed in the future to avoid an increase of workload for the ATCOs. It is an issue about when AISA should provide some type of information because the ATCOs priorities can differ depending on the situation.
- For the risk 'lack of trustworthiness in ML modules', there was a balance between severity responses between B, C, and D due to the expertise of the different subject matter experts involved, i.e., there is no single perception of severity and must be analysed in depth in further safety evaluations.
- The introduction of AISA must follow two pillars: the definition of a transition phase covering all the potential hazards that could arise, and the training of ATCOs in this type of systems, not only about the way to work but about what AI implies.



EUROPEAN PARTNERSHIP

Co-funded by the European Union

# 6 Conclusions

This document presents the results of the AISA risk assessment. Risk assessment focuses on performing a safety analysis by identifying hazards, analysing them and their risk (based on probability and severity) and providing mitigation measures.

The first output of the AISA risk assessment has been the analysis of the concept, system, and requirements to correctly identify the performance and functionalities of the AISA system. This analysis allows for deepening on the AISA knowledge and to identify gaps in areas that have not been fully defined, considering the information provided by the ConOps and requirements.

Based on the identification and analysis of hazards, four hazard areas have been identified to separate the hazards according to the different technologies that constitute the AISA system. Machine Learning, Knowledge engineering, ATC tools and AISA system. In addition, the AISA system has been split into PoC and Project level. This division was necessary because the solution of the PoC system means some limitations that should not apply after the project-phase of AISA development. All these hazards and mitigation measures associated with them constitute a new AI library. This library is meaningful because it allows other researchers to not start from scratch, making it a cornerstone for further safety projects related to the integration of AI-based tools in ATC or even in the ATM.

The main result of the risk identification can be summarised in the constitution of a set of safety requirements for the further development of AISA and other AI-based systems. Risk assessment has identified areas, systems, or functions that can be critical or limit the development of AI systems. This implies that measures must be imposed to avoid the appearance of these risks or to mitigate the consequences. Many mitigation measures proposed in this work are related to the implementation of risks as safety requirements during the design phase of the system.

Considering the numbers, 74 hazards distributed in the different families (14 for ML, 14 for knowledge engineering, 9 for ATC tools, 9 for AISA PoC and 28 for the AISA Project) and more than 150 mitigation measures have been proposed. These mitigation measures allowed to diminish the system risk, by reducing the number of non-acceptable risks from 2 to 0, decreasing the number of tolerable risks from 55 to 16 and increasing acceptable risks from 17 to 58. Therefore, the results of the risk assessment conclude that the system could be considered safe with current conditions after the implementation of mitigation measures.

Finally, the main limitations of this work have been the lack of statistical data and the theoretical scope of risk assessment. This is the reason why a qualitative risk assessment was developed due to the fact that most of the intelligence technologies considered in this work are currently in development and no statistical data were available. However, the implementation of mitigation measures was not feasible to implement due to the scope of the AISA project. Throughout the evolution of AISA project and the development of different technological solutions, this risk assessment should be updated with the inclusion of new risks and the re-analysis of the current risks.



EUROPEAN PARTNERSHIP

Co-funded by the European Union

# 7 References


[1] EASA, “Artificial Intelligence Roadmap. A human-centric approach,” 2020.

[2] ICAO, Doc 9859, Safety Management Manual (SMM), 2010.

[3] H. B., G. B., P. B., J. D., M. E., NLR, “Accident Risk Assessment for Advanced Air Traffic Management,” 2001.

[4] SESAR, “Agent-Based Modelling of Hazards in ATM,” 2012.

[5] H. B., G. B., S.H. Stroeve, “SYSTEMIC ACCIDENT RISK ASSESSMENT IN AIR TRAFFIC BY MONTE CARLO SIMULATION,” 2009.

[6] HINDAWI, “Change-Oriented Risk Management in Civil Aviation Operation: A Case Study in China Air Navigation Service Provider,” 2020.

[7] J. C., Á. R. C., A. O., Carlos Capitán, “Risk Assessment based on SORA Methodology for a UAS Media Production Application,” 2019.

[8] Mathesia, “Aprendizaje automático en el control del tráfico aéreo,” 2020. [Online]. Available: https://mathesia.com/machine-learning-in-air-traffic-control/.

[9] ARC, “Predicción del retraso debido al control del tráfico aéreo mediante machine learning,” 2017. [Online]. Available: https://arc.aiaa.org/doi/abs/10.2514/6.2017-1323.

[10] EUROCONTROL, “Predicting flight routes with a Deep Neural Network in the operational Air Traffic Flow and Capacity Management system,” 2018.

[11] EUROCONTROL, “REVIEW OF TECHNIQUES TO SUPPORT THE EATMP SAFETY ASSESSMENT METHODOLOGY,” 2004.

[12] J. K. Toms Noskievic, “Air Traffic Control Tools Assessment,” *Magazine of Aviation Development,* 2017.

[13] W. H. G. B. H. Barry Kirwan, “Human error data collection as a precursor to the development of a human reliability assessment capability in air traffic management,” 2008.

[14] K. R. N., Ronald L. Stroup, “APPLICATION OF ARTIFICIAL INTELLIGENCE IN THE NATIONAL AIRSPACE SYSTEM – A PRIMER,” 2019.

[15] A. S. M., R. P., J. P. M., Christopher D. Wickens, “The future of Air Traffic Control: Human operator and automation,” 1998.

[16] R. W., E. Hollnagel, “Functional modeling for risk assessment of automation in a changing air traffic management environment,” 2008.

[17] D. K. S., Ranjit Singh, “A Descriptive Classification of Causes of Data Quality Problems in Data Warehousing,” 2010.

[18] L. T., M. Y., S. T., XiangyuWang, “Knowledge Graph Quality Control: A Survey,” 2021.

[19] M. A. J., C. Q., P. V., Matthias Jarke, “Architecture and Quality in Data Warehouses: An Extended Repository,” 1999.

[20] FAA, “Safety Risk Management: the 5 step process,” 2018.

[21] EUROCONTROL, ESSAR 4, 2001.

[22] FAA/EUROCONTROL, ATM Safety Techniques and Toolbox, 2007.

[23] SESAR, “Concept of Operations for AI Situational Awareness,” 2021.

[24] SESAR, “AI Situational Awareness,” 2021.

[25] R. M. Keller, Ontologies for Aviation Data Management, 2016.

[26] SESAR Exploratory Research, “AISA Project,” 2020. [Online]. Available: https://aisa-project.eu/.

# Appendix A Definitions

| | |
|---|---|
| **Air Traffic Control** | A service operated by appropriate authority to promote the safe, orderly and expeditious flow of air traffic. **[24]** |
| **Air Traffic Management** | The dynamic, integrated management of air traffic and airspace (including air traffic services, airspace management and air traffic flow management) — safely, economically and efficiently — through the provision of facilities and seamless services in collaboration with all parties and involving airborne and ground-based functions. [24] |
| **Consequence** | A consequence is defined as the potential outcome (or outcomes) of a hazard. The damaging potential of a hazard materializes through one or many consequences. [2] |
| **Control/ Mitigation** | Generally speaking, control and mitigation are terms that can be used interchangeably. Both are meant to designate measures to address the hazard and bring under organizational control the safety risk probability and severity of the consequences of the hazard. [2] |
| **Artificial Intelligence** | The ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings. The term is frequently applied to the project of developing systems endowed with the intellectual processes characteristic of humans, such as the ability to reason, discover meaning, generalize, or learn from past experience. [24] |
| **Knowledge Graph** | A knowledge graph is a programmatic way to model a knowledge domain with the help of subject-matter experts, data interlinking, and machine learning algorithms. [24] |
| **Machine Learning** | Machine learning is the science of getting computers to learn and act in the same way humans do, with improving their learning over time autonomously by being fed volumes of big data in the form of observations and real-world interaction. [24] |
| **Ontology** | An ontology is a type of data model that has emerged in recent years from a convergence of research in the artificial intelligence (AI), semantic web, and information science communities. [25] |
| **Hazard** | Condition or an object with the potential to cause injuries to personnel, damage to equipment or structures, loss of material, or reduction of ability to perform a prescribed function. [2] |
| **AISA** | AISA (AI Situational Awareness Foundation for Advancing Automation) is a SESAR Exploratory Research project investigating how to increase automation in |





Co-funded by the European Union

| | |
|---|---|
| | air traffic management. The project will explore domain-specific application of transparent and generalizable artificial intelligence methods. [26] |
| **Safety risk** | Safety risk is defined as the assessment, expressed in terms of predicted probability and severity, of the consequences of a hazard, taking as reference the worst foreseeable situation. Safety risk is a product of the human mind intended to measure the seriousness of, or "put a number" on, the consequences of hazards. [2] |
| **Safety** | The state in which the possibility of harm to persons or of property damage is reduced to, and maintained at or below, an acceptable level through a continuing process of hazard identification and safety risk management. [2] |
| **Severity** | The possible consequences of an unsafe event or condition, taking as reference the worst foreseeable situation. [2] |
| **Situational Awareness (SA)** | Is the perception of environmental elements and events with respect to time or space, the comprehension of their meaning, and the projection of their future status. [23] |

# Appendix B Risk assessment session

The risk assessment session was held online and the assistants were:

| Institution | Assistant |
|---|---|
| **Universidad Politécnica de Madrid** | Javier Alberto Pérez Castán<br>Luis Pérez Sanz<br>José María López Pellicer |
| **Faculty of Transport and Traffic Sciences at University of Zagreb** | Tomislav Radišić<br>Ivan Tukarić<br>Dorea Antolović<br>Kristina Samardžić<br>Mia Bazina |
| **Zurich University of Applied Sciences (ZHAW), School of Engineering** | Ruth Esther Häusler Hermann |
| **Johannes Kepler University Linz / Institute of Business Informatics** | Bernd Neumayr |
| **Slot Consulting Ltd** | Roland Gurály |
| **Skyguide Swiss Air Navigation Services Ltd** | Keiko Moebus<br>Christoph Herberth<br>Jennifer Burkhalter |
| **Technische Universität Braunschweig, Institute of Flight Guidance** | Lars Schmidt |
| **EUROCONTROL** | Rocío Barragán Montes |
| **Delft University of Technology** | Erik-Jan van Kampen |
| **Innaxis** | Pablo Hernández |

For the risk assessment session, it was structured the session as follows and this information was provided to the assistants in advance.

### Part 1: Risk assessment presentation

The first part of the session is focused on the presentation of the methodology developed to perform the risk assessment of the AISA system. The main goals, limitations and description of the process will be presented.

### Part 2: Risk assessment process

To facilitate the risk assessment session, 'risk assessment sheets' have been created that contain the most important information and that will help perform the risk assessment for each hazard. The risk assessment sheets are constituted as follows.

#### 1. Step 1: Hazard description and context

Step 1 consists in describing the hazard and providing some context about its implications on the AISA system. In the sheet, there are two areas that represent the hazard description (on the left)



EUROPEAN PARTNERSHIP

Co-funded by the European Union

and some operational context of AISA (on the right). The objective of these description is to provide to every expert a common understanding of the hazard.

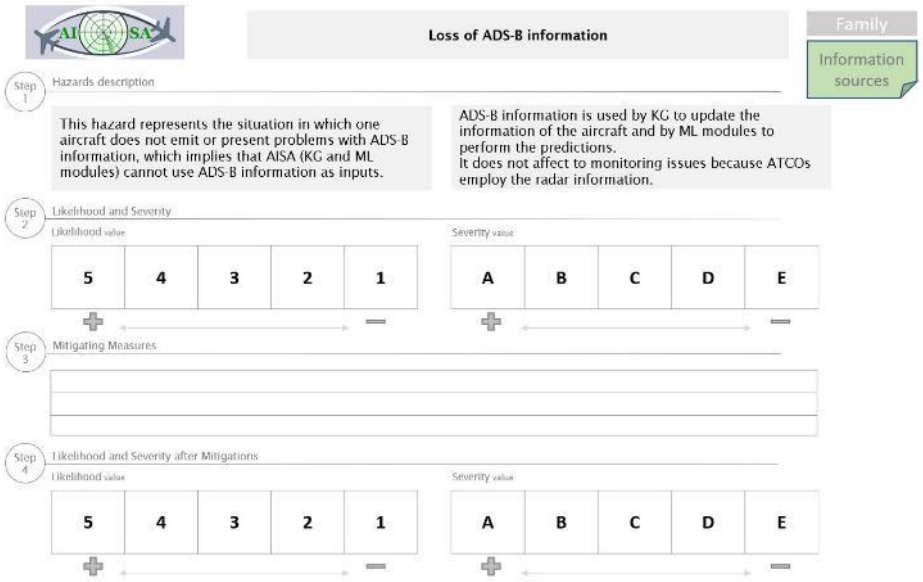


**Figure 9. Example of a risk assessment sheet, initial structure.**

## 2. Step 2: Likelihood and Severity quantification

Step 2 focuses on the risk assessment of the hazard before mitigation measures are considered. The risk is constituted by the likelihood and severity of the hazard proposed by the ICAO in the Table 1 and Table 2. Aimed to get a maximum agreement among the participants, it is very important that all of them keep in mind the descriptions and scales figures/letters provided in previous tables

To reach an agreement on probability and severity factors, it has been developed polling on the Teams meeting. Each assistant provides his/her expertise, and it will be quantified the risk by the majority votes.



EUROPEAN PARTNERSHIP

Co-funded by the European Union

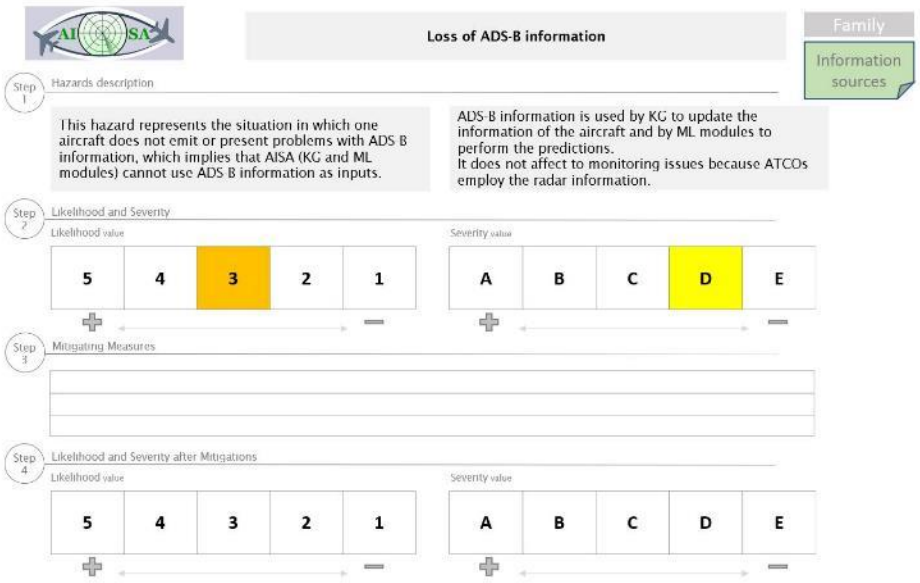


**Figure 10. Example of Risk Assessment Sheet, after Step 2.**

## 3. Step 3: Mitigation measures

This step consists of providing at least one mitigation measure that includes one barrier to avoid or mitigate the risk likelihood and/or severity.

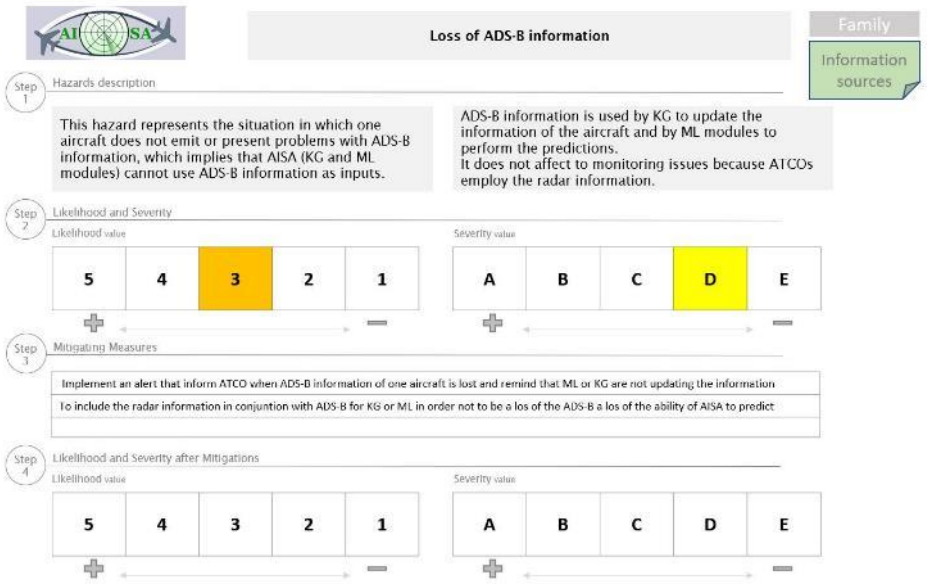


**Figure 11. Example of a risk assessment sheet, after step 3.**

## 4. Step 4: Likelihood and severity re-evaluation



EUROPEAN PARTNERSHIP
Co-funded by the European Union

The last step is to re-evaluate the likelihood and severity of the hazard considering the impact of the mitigation measures as a whole. The process will be the same as the previous quantification based on a new poll. Each assistant provides his/her expertise, and it will be quantified the risk by the majority votes.

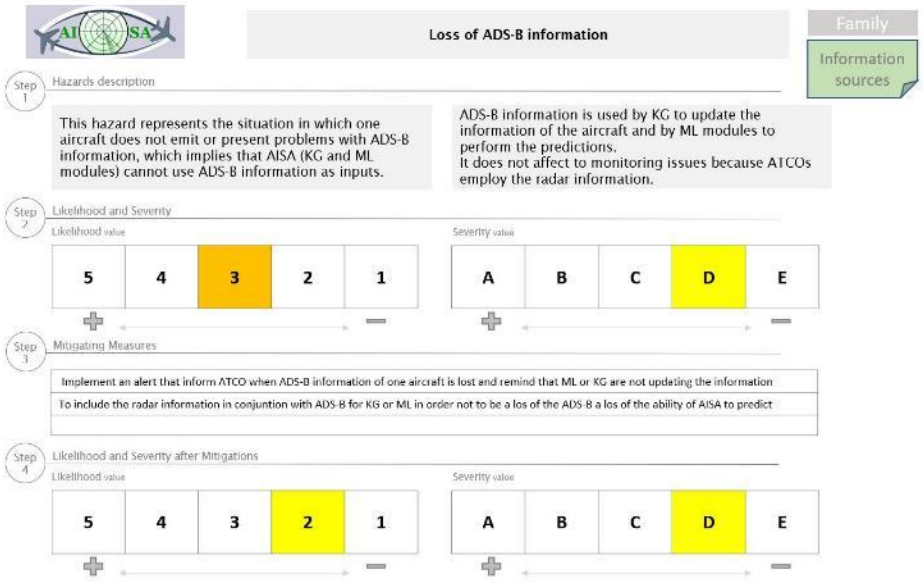


**Figure 12. Example of a risk assessment sheet, after step 3.**



EUROPEAN PARTNERSHIP
Co-funded by the European Union


AI SA
sesar
JOINT UNDERTAKING


University of Zagreb


Technische
Universität
Braunschweig


POLITÉCNICA
UNIVERSIDAD
POLITÉCNICA
DE MADRID


skyguide


zh
aw
School of
Engineering
ZAV Center for Aviation


JKU
JOHANNES KEPLER
UNIVERSITÄT LINZ


onsulting
lot






Co-funded by
the European Union