AI in ATM: transparency, explainability, conformance, situation awareness and trust

A White Paper

INTRODUCTION

Increasing airspace demand requires an increase in the effectiveness and efficiency of the ATC system. Automation, specifically Machine Learning (ML), may present good prospects for increasing system performance and decreasing the workload of ATCOs. AI, however, is typically a “black box” making it hard to include in a socio-technical environment.

Also according to OPTICS and OPTICS2 projects’ analysis - which investigated European aviation research to assess if it is on the right track towards Flightpath 2050 to provide recommendations to steer EU Aviation Safety and Security research - Artificial Intelligence Explainability (XAI) is one of the hottest themes still to be explored in aviation.

As stated in the “Final Safety and Security integrated recommendations” report published in 2021, “Finding the right Human-AI partnerships will be key to future aviation safety. The Intelligent Assistant (IA) in the cockpit and on the ground will be the crucial stepping stone toward fuller Artificial Intelligence (AI) by 2050. Research is urgently needed to determine how humans and AIs can work together productively and safely, including human supervision and recovery in case of ‘aberrant behaviour’ by AI systems”.

This White Paper describes and highlights research progress on using Artificial Intelligence in Air Traffic Management, with a specific focus on Explainability. All the results and lessons learned come from five different SESAR Exploratory Research Projects:

AISA
ARTIMATION
MAHALO
SAFEOPS
TAPAS

Before reading the white paper, if you want to know more about the projects objectives and scope, please consider to have a look to the "AI in ATM SESAR projects" video.

Towards an Automated and exPlainable ATM System

Project duration: June 2020 – November 2022

6

days of Real Time Simulations

2

different operational use cases performed

+80GB

of data processed

WHAT IS TAPAS'S FOCUS?

TAPAS (Towards an Automated and exPlainable ATM System) addressed the effectiveness of introducing Artificial Intelligence (AI) / Machine Learning (ML) solutions to increase the levels of automation in Air Traffic Management (ATM), considering the need of the operator to trust the system (taken as the ability to understand and explain its behaviour and outcomes).

The main objective of the project was the exploration of highly automated eXplainable Artificial Intelligence (XAI) scenarios through Human-In-The-Loop validation activities and Visual Analytics, to identify needs and strategies to address transparency and explainability in the operational cases considered, paving the way for the application of these AI/ML technologies in ATM environments.

WHAT DID THE PROJECT DO?

The scope of TAPAS was the systematic exploration of AI/ML solutions towards increasing levels of automation in specific ATM scenarios, through analysis and experimental activities, with the objective to deliver principles of transparency, enabling the application of AI/ML supported automation in ATM. Specifically, TAPAS managed to:

- **Describe and analyse in detail two operational use cases:** conflict detection and resolution applied to Air Traffic Control (tactical phase), and Air Traffic Flow Management (pre-tactical phase).

- **Develop XAI methods**, addressing the requirements of both operational cases, which focus on the needs of operators (and other potential actors) concerning the quality and transparency of solutions generated by the XAI prototypes.

- **Apply Visual Analytics techniques to assess and enhance explainability** of AI/ML systems in ATM.

- **Run validation experiments** that assess the applicability of XAI methods in the various levels of automation considered, exploring different ways of interaction and information exchange to ensure operator's trust in XAI technologies.

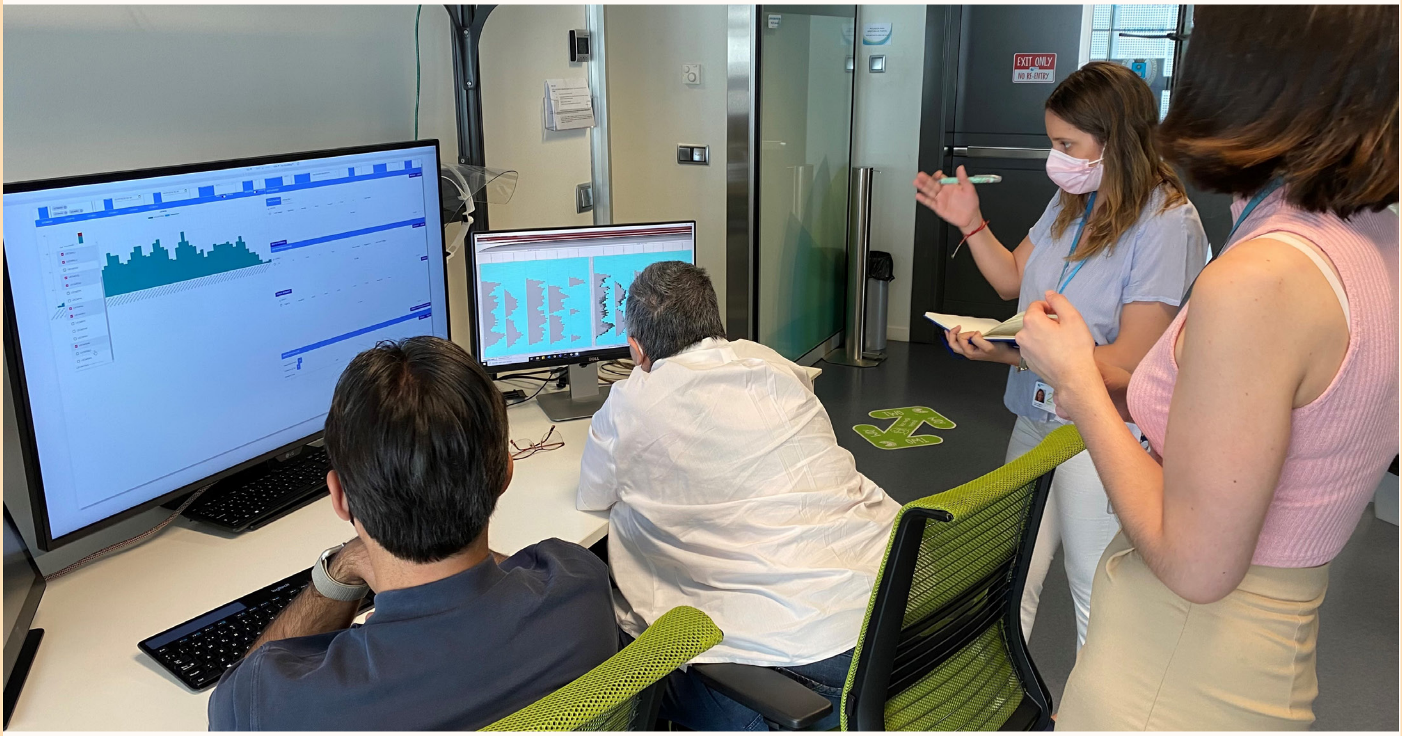
- **Extract conclusions, principles and recommendations** based on these experimental results and analysis, which will serve as an enabler for the implementation of XAI methods in higher levels of automation in ATM. evaluation of assumptions and requirements.

WHAT DID THE PROJECT ACHIEVE?

TAPAS results contributed to the achievement of:

- the development of a **general framework for the identification of principles and criteria for Artificial Intelligence / Machine Learning transparency / explainability** in the Air Traffic Management domain, and

- the **selection and implementation of suitable and explainable Artificial Intelligence / Machine Learning methods** for the testing and validation of the operational use cases addressed: Air Traffic Flow Management (non-safety critical) and Conflict Detection and Resolution (safety critical).



Real-Time test campaign at CRIDA's premises- first round of simulations, June 2021

WHAT'S IN IT FOR AVIATION?

Artificial Intelligence and Machine Learning are technologies intended to enable higher levels of automation, improving ATM capacity beyond current limitations. The research conducted by TAPAS has provided the initial steps in a field of technology that is a key enabler for the uptake of Artificial Intelligence / Machine Learning technologies in a domain as sensitive as Air Traffic Management, setting the technological and operational framework for future applications of the ATM Master Plan Automation Level 2 and Level 3.

WHAT'S NEXT?

The line of research focused on Transparency and Explainability in Artificial Intelligence / Machine Learning application should be continued by deeper exploring the ways in which the human and the machine would interact and, in an ideal case, team up. This potential follow-up, based on the foundations of explainability and trust, is necessary to overcome the second round of challenges, related to the synergies of the joint operations between the human and the machine, moving from the "Human-In-The-Loop" concept to the "Human-In-The-Team" one.

CONTACTS

Project Coordinator

José Manuel Cordero - *CRIDA*
Rubén Rodríguez - *CRIDA*

Communication Manager

Enrique Iglesias - *CRIDA*

tapas-atm.eu | Sesarju.eu/projects/tapas | TAPAS Project | TAPAS_SESAR_ER4

AI Situational Awareness Foundation for Advancing Automation

Project duration: June 2020 – November 2022

80% of monitoring tasks automated

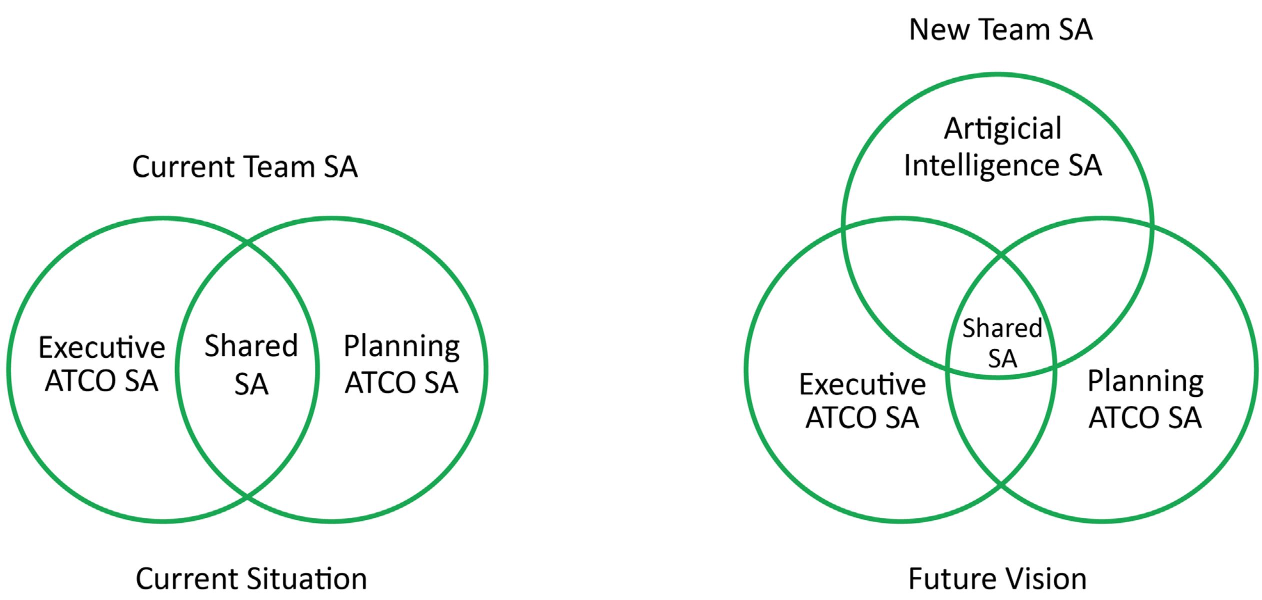
36 licensed ATCOs participated in the experiments

160 ATCO exercises performed

2 human-in-the-loop experiments conducted

WHAT IS AISA'S FOCUS?

The AISA project investigated how to increase automation in air traffic management. To achieve this, the project explored domain-specific application of transparent and generalizable artificial intelligence methods. This project focused on building a foundation for automation by developing an intelligent situationally-aware system. The developed artificial situational awareness allows the AI to take part in shared (or team) situational awareness alongside ATCO team members. The AI is also able to explain the reasoning behind its decision when confronted with the same problem as the ATCO.



Concept of distributed situational awareness for future automated systems

WHAT DID THE PROJECT DO?

The first part of the project was dedicated to creating a knowledge graph containing all relevant information regarding en-route operations. It was based on AIXM and FIXM models with the addition of project-specific knowledge missing from these models (e.g. aircraft performance, Letters of Agreement, ML predictions). Two experiments were conducted in Skyguide facilities, first of which had no AISA system input in order to gather the data from simulation exercises to feed the system with the necessary traffic data. This data was then used as input for the AISA system to create situational awareness and make predictions. The second experiment put to the test the human-machine team situational awareness.

Real-Time test campaign at DLR's premises - second round of simulations, April 2022

Data gathered for the purpose of situation awareness comparison between AI and ATCO was the following:

- Eye tracking recording
- Screen recording
- Frontal recording
- Biometrical data

Although improvements are possible and necessary for the proof-of-concept system, AISA is deemed a success. An analysis of the collected data was performed and it was determined that the concept of human-machine team situation awareness is feasible and the system can be made aware of its own state.

WHAT DID THE PROJECT ACHIEVE?

A unified UML diagram was created that combines both AIXM and FIXM data related to ATC en-route air traffic operations, as well as some additional information needed for this project. The developed methodology is important for other research which uses aeronautical data stored in AIXM and FIXM. A functioning proof-of-concept system was created and validated. AISA system's reasoning engine is based on ATM expert knowledge and uses a knowledge graph through which it interacts with the machine learning modules. The reasoning engine is capable of explaining obtained results and provides insight into inconsistencies and improbable results. This allows the system to be closer to the process of thinking that ATCOs have when making decisions and reaching conclusions.

WHAT'S IN IT FOR AVIATION?

The AISA project worked toward fulfilment of one of SESAR goals to build a foundation for successful cooperation between a human and a machine. AISA also tackled challenges of transparency and generalization, and the project presents a vision of automation in the ATM operational environment of en-route ATC. The AISA system contributes to increased safety by introducing an additional safety net performing tedious monitoring tasks with high reliability. It also contributes to the improvement of interoperability between different systems by enhancing data handling, which comes as a by-product of using knowledge graphs for data management.

WHAT'S NEXT?

Further development of the system is needed as it is currently only a proof-of concept system. The main missing ingredient is the awareness of controller's intent. Being able to analyze the controller's intent will enable the system to provide adaptable assistance to the controller. In a way, the system is out-of-the-loop when it comes to understanding the needs of the controller. Furthermore, the system is currently not operating in real-time but this has been deemed achievable considering that the processing time of a single situation graph is similar to the refresh rate of an ATCO's working position. Deeper integration of ML modules is needed, as well as exploration of ways to compare integrity levels of different ML modules.

CONTACTS

Project Coordinator **Tomislav Radišić - *FTTS*** - tradisic@fpz.unizg.hr

Dissemination Manager **Roland Guraly - *SLOT* -** rolandguraly@slotconsulting.hu

aisa-project.eu | Sesarju.eu/projects/aisa | AISA Exploratory Project

SafeOPS

From Prediction To Decision Support - Strengthening Safe And Scalable ATM Services Through Automated Risk Analytics Based On Operational Data From Aviation Stakeholders

Project duration: January 2020 - December 2022

+600.000

real world approaches

+1.850

real world go-arounds used to train a go-around prediction

WHAT IS SAFEOPS'S FOCUS?

SafeOPS investigates an AI-based decision support tool for Air Traffic Control in the context of go-arounds. The project researches how Air Traffic Controllers handling go-arounds could use the predictive information to adapt their strategies, avoiding knock-on effects that can accompany go-arounds; and thereby increasing safety and resilience.

WHAT DID THE PROJECT DO?

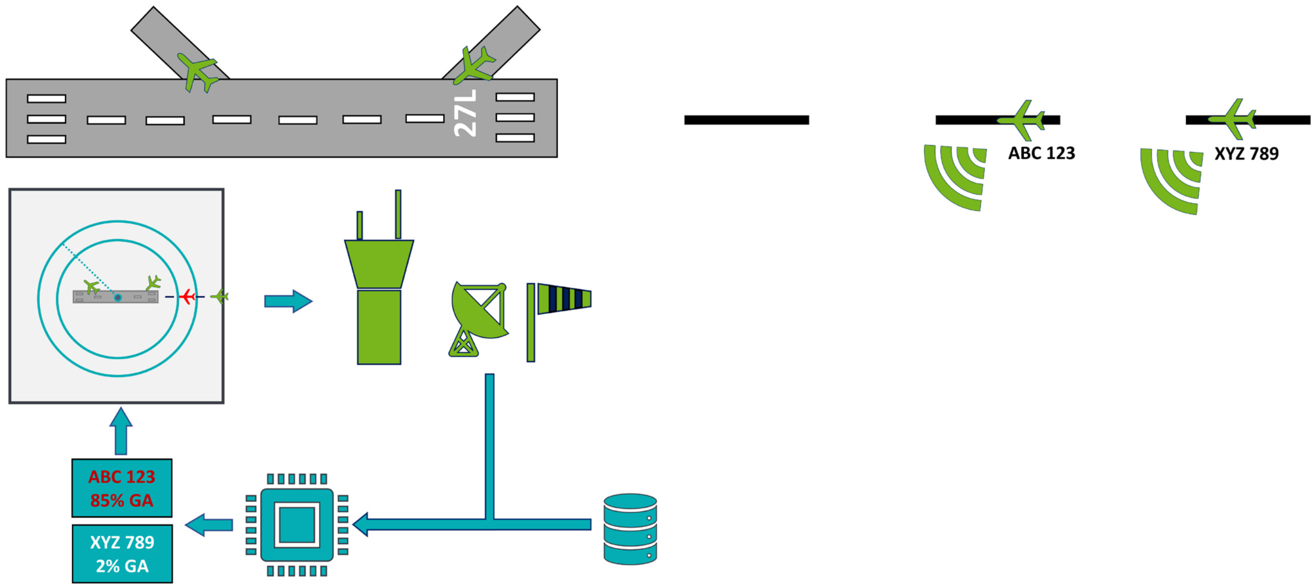
The project employs a three-layered approach that includes an operational layer, a predictive layer, and a risk framework. In workshops with Air Traffic Controllers, the *operational layer* developed scenarios in which a go-around prediction could provide benefits for Air Traffic Management. This led to an initial concept of operations and the definition of use cases and requirements in the initial phase of SafeOPS.

Based on this, the *predictive layer* developed an initial machine learning model for go-around predictions, yielding insights in achievable accuracies and the transparency of the predictions. In parallel, the *risk framework* assessed the risks and benefits of the foreseen decision support concept in terms of operational safety, including also human factors considerations.

In the final phase of the project, the operational layer designed a simulation exercise based on the *risk framework's* and the *predictive layer's* findings, to investigate the impact of the SafeOPS concept on safety and resilience. Thus, SafeOPS focused on separation challenges, workload, and capacity of the Tower operations in the initially defined scenarios.

WHAT DID THE PROJECT ACHIEVE?

SafeOPS developed an AI-model, which can predict go-arounds, based on publicly available data sources. Based on this, it investigated risks and benefits of a decision support concept for Air Traffic Management in a complementary evaluation, consisting of a risk framework and simulation exercises. With this impact evaluation, SafeOPS could demonstrate that the envisioned concept can provide benefits in terms of safety and resilience. Even false predictions do not have negative effects on safety, which is an important finding; however, false predictions can negatively affect the capacity of airports.



SafeOPS Concept visualization on a high-level basis. Aircraft transmit performance data, which is received by ANSPs' radar. Together with weather information, the SafeOPS concept predicts the likelihood of a go-around for each approach and provides this information to the Controller via the radar screen.

WHAT'S IN IT FOR AVIATION?

The next generation Air Traffic Management systems are driven by two goals that are hard to combine: the increasing demand for capacity and cost-efficiency, as well as for safety and resilience. SafeOPS proposes a proactive approach for safety management, capable of predicting safety hazards in real-time. Also, the probabilistic nature of the concept will contribute to the introduction of new liability models, regulations, processes, and risk assessment methodologies in the aviation community.

WHAT'S NEXT?

Based on its results, SafeOPS identified some topics for future research. First, the AI-based prediction tool must be enhanced to become an online capable tool, based on ANSP radar data, and to meet the EASA guidelines for AI in aviation[2]. Second, the safety and resilience vs. capacity trade-off must be investigated. With the potential effects on safety, resilience, and capacity now explored, a cost analysis including capacity losses is necessary to find the minimum acceptable accuracies for the machine learning component. Lastly, the concept must be investigated for wider operational spectrum, using Monte Carlo methods complementary to the simulation exercises.

CONTACTS

Project Manager

Lukas Beller - *TUM* - lukas.beller@tum.de

Project Partners

Carlo Abate - *Deep Blue* - carlo.abate@dblue.it
Elizabeth Humm - *Deep Blue* - elizabeth.humm@dblue.it

Pablo Hernandez - *Innaxis* - ph@innaxis.aero

Philipp Kurth - *DFS* - philipp.kurth@dfs.de

safeops.eu | Sesarju.eu/projects/safeops

[2]EASA Concept Paper: First usable guidance for Level 1 machine learning applications (www.easa.europa.eu/en/downloads/134357/en).

Modern ATM via Human/Automation Learning Optimisation

Project duration: July 2020 - December 2022

34

real en-route ATCOs involved

2

main simulations performed, in Italy and Sweden

18

different scenarios per each controller played

WHAT IS MAHALO'S FOCUS?

*Artificial Intelligence (AI) provides both **opportunities and considerable challenges** to the continued growth of Air Traffic Control (ATC) services. The MAHALO project focused on two constructs thought to underlie human-AI interaction. The first is **conformance**, which the project has defined as the apparent strategy match between human and AI systems. The second construct, **transparency**, refers to the degree to which the system makes its internal processes apparent to the operator. MAHALO set out to experimentally manipulate these two constructs, and to explore their main and interactive effects on a broad number of **human performance measurements**, including conflict detection performance, automation acceptance, and rated workload.*

WHAT DID THE PROJECT DO?

MAHALO conducted two field simulations at two sites, with 34 participants, **to evaluate the impact of conformance and transparency manipulations on controller** acceptance, agreement, workload, and general subjective feedback, among other measures. Each simulation consisted of two phases. First was a training pre-test in which controllers interacted with scripted traffic scenarios that presented two-aircraft closing conflicts, and which recorded controllers' resolution strategies. Second was a main experiment phase, in which the same controllers interacted with Machine Learning (ML) solved analogues of the pre-test scenarios.

Conformance of advisories had an impact on controllers' response, but not in a uniform direction. Although personalized advisories received more favourable responses in many cases, there were also cases when the optimal or group advisories were favoured. There was no strong effect of transparency on controllers' responses. An in-depth analysis was made dividing participants in two groups depending on how close their separation distance preference was to the target separation distance aimed for by the optimal model's advisory. The analysis revealed a reoccurring pattern emerging where the group of participants, whose average separation distance measured in the training pre-test was closer to the separation distance aimed for by the optimal advisory, showed unchanged or more positive responses to the advisory with increasing transparency. That is, **their acceptance of advisories and ratings of agreement, conformance, and understanding was higher** compared with the other group.

WHAT DID THE PROJECT ACHIEVE?

Over the project runtime the MAHALO project has achieved several results:

- Conducted a state-of-the-art review of Machine Learning (ML) advances.
- Developed and demonstrated a ML capability.
- Designed an experimental user interface and simulation capability.
- Conducted human-in-the-loop validation trials of the user interface.
- Integrated ML capabilities with the simulator and experimental interface.
- Conducted a first full simulation to demonstrate the entire test platform.
- Specified experimental design for the final simulation sessions.
- Defined some guidelines on how to incorporate conformance and transparent mechanisms of AI solutions to conflict detection and resolution.

PUBLICATIONS:

Westin, C., Hilburn, B., Borst, C., Kampen, E. Van, Bang, M., (2020), Building Transparent and Personalized AI Support in Air Traffic Control, Digital Avionics Systems Conference, 2020

Westin, C., Nunes Monteiro, T., Borst, C., Bonelli, S., (2020), Balancing transparency and conformance of an AI conflict detection and resolution support system, Proceedings of the Tenth SESAR Innovation Days, 2020.

Nunes Monteiro, T., Borst, C., Kampen, E. Van, Hilburn, B., & Westin, C. (2021), Human-interpretable Input for Machine Learning in Tactical Air Traffic Control. Proceedings of the Eleventh SESAR Innovation Days, 92, 1--6. + presentation (WP3 & WP4).

Bonelli, S., Borst, C., Brambati, F., Cocchioni, M., Hilburn, B. & Monteiro-Nunes, T., (2022), Transparent and Personalised AI Support in Air Traffic Control: First Empirical Results, The International Conference on Cognitive Aircraft Systems, 2022.

A participant interacting with the Simulator. The monitor shows a frozen scenario with the highest level of transparency

WHAT'S IN IT FOR AVIATION?

An implicit assumption going into research was that **transparency fosters understanding, acceptance** and **agreement**. As a thought experiment, however, consider the case where poorly functional automation is outputting advisories. In this case, transparency might have the opposite effect and lower controllers' agreement and acceptance of the system. The notion here is that if transparency involves making clear to the operator the inner workings of the algorithm, it does not necessarily increase agreement and acceptance, but should optimize them. Transparency and explainability should **increase acceptance and agreement for an optimal algorithm, which should also decrease acceptance and agreement for a sub optimal algorithm.**

WHAT'S NEXT?

Results of the MAHALO project, while valuable in themselves, suggest clear avenues of exploration for future research. One avenue is the potential utility of **personalisation**, or tuneable parameters that might allow for a hybrid of the optimal and personal model view, such that controllers could tune certain parameters within the confines of an optimal advisory system. Second is the potential benefit of an **adaptive systems approach**, to automate that tuning such that parameters can be adjusted in step with learning performance.

Third is the **potential hybridisation of machine learning together with adaptive systems**, both for task performance but also as a trigger for task reallocation. Finally, future research is required to explore **potential benefits of advisory transparency** on advisory acceptance and system trust. In contrast to expectations, measures of acceptance, agreement and workload did not benefit overall from increased transparency. This suggests that transparency alone may not be suitable as a measure for increasing operators' acceptance.

CONTACTS

Project Coordinator

Stefano Bonelli - *Deep Blue* - stefano.bonelli@dblue.it

Project Partners

Matteo Cocchioni - *Deep Blue* - matteo.cocchioni@dblue.it

Clark Borst - *TUD* - C.Borst@tudelft.nl
Erik-Jan van Kampeen - *TUD* - E.vanKampen@tudelft.nl

Carl Westin - *LiU* - carl.westin@liu.se

Brian Hilburn - *CHPR* - brian@chpr.nl

Supathida Boonsong - *LFV* - supathida.boonsong@lfv.se

Mahalo.eu | Sesarju.eu/projects/mahalo | Mahalo project | H2020_MAHALO

ARTIMATION

Transparent Artificial Intelligence and Automation to Air Traffic Management Systems

Project duration: July 2020 - December 2022

2

use cases developed

3

dimensions for immersive analytics

0

cables used by the EEG helmet to gather neurophysiological data

WHAT IS ARTIMATION'S FOCUS?

The project assessed the impact of both different visualisation techniques for AI algorithms to provide explainability to the ATC task of Conflict Detection and Resolution, and the impact of an Explainable AI (XAI) tool on Air Traffic Controllers' trust to optimise the use of runways, supporting the task of Delay Prediction.

WHAT DID THE PROJECT DO?

ARTIMATION gathered information about three different kind of visualisation techniques to represent an AI outcome for the task of Conflict Detection and Resolution. It then tested them with two different experimental groups, i.e., ATC experts and ATC students. Differences in human factors indexes (e.g., stress, workload, approach-withdrawal, acceptance, trust) were assessed both with neurophysiological measures, and traditional self-report questionnaires and semi-structured debriefing interviews. Moreover, three different algorithms have been trained to support the task of Delay Prediction to optimise the runway use. The three algorithms select different parameters to show to Tower Air Traffic Controllers to understand what selection provides the higher trust level and the higher agency towards the delay.

WHAT DID THE PROJECT ACHIEVE?

As far as the Conflict Detection and Resolution visualisation use case, ARTIMATION assessed the human performance differences between ATC experts and students. The project explored the feasibility of introducing Explainable AI tools in training sessions for ATC students, to always provide a different point of view and a tool to compare the students' own decision-making process with an optimal solution designed considering several parameters.

The Delay Prediction use case aims at optimising the runway use introducing explainability through the visualisation of parameters influencing an aircraft delay. Preliminary results will be illustrated during the conference.

ARTIMATION Conflict Detection and Resolution visualisation (picture taken during validation activities)

WHAT'S IN IT FOR AVIATION?

Aviation will benefit from ARTIMATION from both use cases. The Conflict Detection and Resolution use case will introduce the concept of Explainable AI in training, assessing the feasibility of different kinds of visualisations to have a better mental picture of a possible optimal resolution. The Delay Prediction use case will realise a proof-of-concept of the introduction of Explainable AI in Control Towers to optimise the use of runways.

WHAT'S NEXT?

Nowadays, explainable AI is not made for high-complexity tasks, such as Conflict Detection and Resolution. An environment that could benefit from the introduction of XAI can be the remote digital tower: this new environment will be conditioned more and more from digitalisation and the introduction of Artificial Intelligence. ARTIMATION's next steps involve understanding and assessing how to provide trust to Air Traffic Controllers dealing everyday with AIs.

CONTACTS

Project Coordinator

Mobyen Uddin Ahmed - *Malardalens University*

Project Partners

Shaibal Barua - *Malardalens University*
Mir Riyanul Islam - *Malardalens University*

Stefano Bonelli - *Deep Blue*
Ana Ferreira - *Deep Blue*
Nicola Cavagnetto - *Deep Blue*

Fabio Babiloni - *La Sapienza University*
Pietro Aricò - *La Sapienza University*
Giulia Cartocci - *La Sapienza University*

Christophe Hurter - *Ecole Nationale de l'Aviatione Civile*
Augustin Degas - *Ecole Nationale de l'Aviatione Civile*

artimation.eu | Sesarju.eu/projects/artimation | Artimation Project | ArtimationAI

LESSONS LEARNED
ER4 PROJECTS ON TRUST

To unpack AI, especially in high reliability organisations/ environments like ATM, trust in AI-based concepts/ solutions is paramount. Trust has to be built, not only with the end-users but on several levels, amongst others with stakeholders, regulatory bodies but also developers of the solutions/concepts. The presented projects therefore pursued various approaches to build trust in solutions/concepts based on AI components, for which we summarise the lessons learned in the following.

Explainability/ Transparency

One common approach to build trust on the **User Level**, is Explainable AI (XAI). **TAPAS** finds that during operations, rather than having explanations, the user needs to trust the system. This trust does not necessarily need to be built during operations, but can also be acquired in training and (de-)briefing , which is found more valuable by the users than online explanations, especially in time or safety critical situations.

Artimation, which supports the TAPAS findings, also concludes that Air Traffic Controllers(ATCOs), when under pressure, do not find the time to review explanations from the AI. Therefore, ATCOs prefer a simple design, presenting the AI output. ARTIMATION also finds significant differences in ATCOs' preferred visualization of AI results depending on their seniority. Experienced ATCOs prefer to receive only visualizations of proposed solutions. Furthermore, they have the confidence to reject AI solutions, if they do not match their experience. Students seem to be more flexible in receiving additional information and accept AI solutions, even if these solutions are contrary to their own approach.
The project therefore advocates for additional on-demand explanations that, in less stressful situations or in simulations, help the user develop trust through a better understanding of system behavior.

MAHALO concluded that increased transparency can benefit understanding of e.g. the system and/or situation, but does not necessarily benefit acceptance of a system and agreement with its advisories. The effect might be opposite, where increased transparency decreases acceptance and agreement simply because the system is offering an explanation that reveals that its reasoning is different from that of the operator. Transparency should ideally be individually tailored to facilitate a dialogue between humans and AI. Transparency should be provided in relation to what the user needs to understand, which requires the AI to be able to develop an understanding of the human it is interacting with.

Personalization/ Conformance

Understanding, and based thereon acceptance of AI solutions can be achieved through personalization of AI. This approach is investigated in **MAHALO**, which therein finds a way to overcome some of the challenges associated with integrating ML/AI techniques in ATC from the perspective of human-machine collaboration. A consequence from personalizing AI solutions is the increasing amount of necessary data for training purposes.

Building initial trust is one thing. Regaining it once it is lost, due to e.g. a system failure, is the other - and its hard. Therefore, focus on the safety and reliability of solutions must be paid. Both has to be achieved and demonstrated throughout the development process of a system with an AI constituent.

Stakeholder and Field Experts Involvement

Stakeholders are a part of the consortium for all projects. This does not only ensure relevant feedback during the experiments and testing (of the proposed concepts/solutions) but most importantly already during development stages. Following the *System Thinking* approach, involvement of Field Experts in a user-centric development approach, is a key aspect in designing and developing resilient and safe **systems**. That is all the more crucial, when developing solutions based on technology which is unprecedented. Assessing the impact of AI solution on the operation, already during development phases, requires close synchronisation of stakeholders, field experts and developers.

Safety Assessment

To ensure safe operations, EASA suggests including AI/ML failure modes in the system safety assessment.For a binary classification, admittedly one of the simpler AI solutions, SafeOPS includes the false prediction scenarios, when assessing the impact of the concept on the operation. AI performance metrics like precision and recall allows weighting the different possible scenarios, including false predictions, according to their relative likelihood. The result is an impact evaluation of the complete and probabilistic output of the AI solution on the operational domain. How to generalize this approach to other, more complex AI solutions, is ongoing research and building trust on a **system level** thus still an open question.

Learning Assurance

The same is true on the **Development and Software Design level**. While there exist Software Assurance Levels (SWAL) to establish trust/confidence in a software being free from failures or bugs, for a trained AI solution (not the software but the learned hyperparameters of the algorithm) and the data involved, these do not generally apply.

To build trust in data sets and trained AI solutions, the AISA risk assessment emphasises the importance of Data Quality of the training, test and validation data sets. It concludes that deriving and validating *data quality requirements*, to monitor and document the quality of the data, building the baseplate of the AI. Additionally, the *generalization error* (the quality of predictions for inputs not in the training data set) of the trained AI solution should be analysed, to ensure the AI has “understood” the data patterns, detecting over- and underfitting in the training process.

A common observation of all projects is the cost- and time-consuming aspect of data acquisition - and especially - the data quality assurance for AI purposes. Therefore, an open access data lake/repository of the ATM/AI community with raw data, but also data that is pre-processed and cleaned, according to data quality standards might be an interesting approach to save time and costs, especially for ER projects which usually develop AI solutions on a proof-of-concept level, which is a topic, also highlighted in *The Fly AI Report*.

CONCLUSIONS

"Higher levels of automation supporting air traffic controllers' workload and reducing their stress are key for a future-proofed ATM system". This statement, made by SESAR 2020 Scientific Committee in 2019 as part of a long-term vision and research roadmap, remains as valid now as then. It clearly indicated the way forward for automation in ATM, in line with other key documents such as the European ATM Master Plan, Flightpath 2050, the International Civil Aviation Organization (ICAO) Global Air Navigation Plan, the NextGen implementation plan, or the SESAR Airspace Architecture Study: all of them, among others, identify higher levels of automation as key enablers for future ATM concepts.

Since then, some work has been done in this direction. This paper has presented the main results and conclusions of five SESAR 2020 ER4 projects (AISA, ARTIMATION, MAHALO, SafeOPS and TAPAS), converging in similar ideas from different and independent approaches and perspectives. All of them have applied Artificial Intelligence systems to use cases testing higher levels of automation than in current operations, pushing the boundaries and the state-of-the-art in line with the levels of automation defined in the European ATM Master Plan, which illustrates how ATM and aviation will evolve into an integrated digital ecosystem characterized by distributed data services as envisioned in Phase D of the ATM Master Plan.

As described along this paper, the projects provided evidence of technical and operational feasibility for AI systems in the Controller Working Position, in particular in Conflict Detection and Resolution applications, including higher levels of automation where the system would be able to take resolution decisions and implement them directly without intervention of the human operator, who would keep a monitoring role and full situational awareness, being able to intervene at any time without a negative impact in safety.

This research opens the door to new challenges, in particular those related to the human factors and potentially linked operational challenges. How to manage potential de-skilling of air traffic controllers when adopting a monitoring role, or keep situational awareness in a

system able to operate beyond regular human workload capacity as set in nowadays ATM operations are aspects which will need to be tackled carefully in future research.
On the other hand, these projects also focused on the explainability and transparency of the AI systems, fundamental for their effective uptake. The key conclusion is that operators identify as more important to trust the system than getting explanations on the AI decisions in the tactical phase. This trust, of course, has its foundations on the certification and training phases, where respective actors should dig into the system to understand its behavior, validate it and, eventually, build this fundamental trust. This may be accompanied by regulatory changes in terms of liability. The projects developed principles and recommendations which may serve as a starting point to address explainability in future AI applications in ATM.

Despite the progress made by these projects, human-centric full automation is still far away. The future ATM system will have needs in terms of scalability and resilience that will join to the usual operational needs in safety, capacity, environment or cost-efficiency, but this will need an evolution in the way that human and machine work together as a team, more than just break down activities between them. This teaming of both sides is considered a fundamental enabler to progress beyond ATM Master Plan automation level 3, in which research activities should be planned.

SUPPORTED BY
sesar
JOINT UNDERTAKING

EUROPEAN UNION

The projects presenting their findings in this white paper have received funding from the SESAR Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 892970, No. 894238, No. 892919, No. 892358, No. 892618.

Produce by Deep Blue with the support of projects consortia.

PRINTED IN DECEMBER 2022