EXPLORATORY RESEARCH

D3.3 Human Factors Analysis of the Impact of Providing Probabilistic Risk Information in Real Time

Deliverable ID:	D3.3
Dissemination Level:	PU
Project Acronym:	SafeOPS
Grant:	892919
Call:	H2020-SESAR-2019-2
Topic:	SESAR-ER4-06-2019
Consortium Coordinator:	TUM
Edition date:	27 May 2022
Edition:	00.03.00
Template Edition:	02.00.03

Authoring & Approval

Authors of the document

Name / Beneficiary	Position / Title	Date
Elizabeth Humm / DBL	Project Team Member	26/04/2022
Carlo Abate / DBL	WP3 Leader	26/05/2022

Reviewers internal to the project

Name / Beneficiary	Position / Title	Date
Lukas Beller / TUM	Project Coordinator	27/05/2022
Philipp Kurth / DFS	Project Team Member	30/03/2022

Reviewers external to the project

Name / Beneficiary	Position / Title	Date

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

Name / Beneficiary	Position / Title	Date

Rejected By - Representatives of beneficiaries involved in the project

Name and/or Beneficiary	Position / Title	Date

Document History

Edition	Date	Status	Name / Beneficiary	Justification
0.0.1	03/12/2021	Draft	Elizabeth Humm	First draft created
0.0.3	29/03/2022	Draft	Elizabeth Humm Carlo Abate Lukas Beller Philipp Kurth	Draft ready for final revision
0.1.0	31/03/2022	Initial Submission	Elizabeth Humm	Draft ready for submission

			Carlo Abate	
0.2.0	26/04/2022	Revised Version	Elizabeth Humm Carlo Abate	Revised draft to include the reviewer's comments
0.3.0	26/05/2022	Revised Version	Carlo Abate	Newly-revised draft to include the second round of reviewer's comments

SafeOPS

SAFEOPS – FROM PREDICTION TO DECISION SUPPORT. STRENGTHENING SAFE AND SCALABLE ATM SERVICES THROUGH AUTOMATED RISK ANALYTICS BASED ON OPERATIONAL DATA FROM AVIATION STAKEHOLDERS.

This Deliverable is part of a project that has received funding from the SESAR Joint Undertaking under grant agreement No 892919 under European Union's Horizon 2020 research and innovation programme.

Abstract

The SafeOPS project investigates the possibility to develop an artificial-intelligence-based software to support air traffic controllers (ATCOs) in decision making, and the impact this tool has in routine air-traffic operations. The context selected for this investigation is the missed approach initiated by the flight crew of a landing aircraft and the subsequent go-around. In particular, the SafeOPS project is currently developing a system which is intended to be capable of predicting imminent, pilot-induced Go-Arounds and consequently warn the ATCOs to enhance their situational awareness and support their decisions on how to best manage the evolving traffic situation.

It is important to consider the human component in all aspects of the system, including the operational, organisational and management processes associated with a new technological solution and its operators. One part of the system which is particularly critical for effective integration of the human, and one which generally receives most attention, is the Human Machine Interface. One aspect of the SafeOPS project that calls for deeper analysis and contemplation in terms of its impact on Human

Factors, is that of probabilistic information presentation. In SafeOPS, the tool being proposed will ‘learn’ the likelihood of a missed approach through the data associated with previous missed approaches and go-arounds and create a probabilistic model of this data.

This study aims to consider the Human Factors of the integration of the SafeOPS tool, largely considering that which is associated with the Human Computer Interaction aspect. Part of the Human Factors analysis performed in this study is aimed at determining how the provision of probabilistic information changes the tasks the ATCOs perform and how the display of such information influences their decision making and subsequent actions. For this purpose, a series of eleven workshops were conducted between July 2021 and February 2022. The key part of the study was to conduct a design evaluation of the SafeOPS tool, display interface to identify design issues that had the potential to result in safety and usability problems. This was done by assessing the design against a series of design requirements and design heuristics, with the user at a number of workshops. This document provides a summary of the feedback from these workshops arranged in themes. Finally, according to the experience and feedback from the users in the SafeOPS workshops and in alignment with current Human Factors knowledge and research, guidelines on how best to present SafeOPS tool to the user were detailed.

Table of Contents

Abstract 3

1 Introduction *9*

1.1 Human Factors 9

1.2 Scope and limitations of the present study 12

2 Literature review of similar tools *16*

2.1 Conflict Alerting Tools 16

2.2 Area Proximity Warning (APW) 21

2.3 Minimum Safe Altitude Warning (MSAW) 21

2.4 Approach Path Monitor (APM) 22

2.5 AI-based tools investigated in other ongoing projects 22

3 The SafeOPS tool *24*

3.1 Brief description of the tool 24

3.2 Expert evaluation 29

4 Discussion and Recommendations *36*

4.1 Guidelines on how to present the information 36

5 Conclusion and Outlook *42*

6 References *43*

Appendix A Design Heuristics *44*

Appendix B SafeOPS Human Factors Design Requirements *51*

List of Tables

Table 1. Main properties of the three most common types of Conflict Alert systems. 17

Table 2. Precision and Recall of GA events from traffic data at Airport 1. 25

Table 3. Precision and Recall of GA events from traffic data at Airport 2. 26

Table 4. Example of the questions asked at the workshops with ATCO to investigate design heuristics. 33

Table 5. Commonly-used design heuristics. 44

Table 6 (following page). List of questions for the semi-structured interviews of the ATCOs to identify the best design properties to deliver the GA information within the SafeOPS tool. A summary of the answers collected from the interviewees during the workshops is reported below each question, marked with the letter 'A'. 45

Table 7. Human Factors Design Requirements specific to the SafeOPS Tool 51

List of Figures

Figure 1. Current notification alert design: a blinking red [4]. 18

Figure 2 Snapshot of a traffic simulation based on real flight data at Airport 2. 28

Figure 3. Snapshot of a traffic simulation based on real flight data at Airport 2 in which an imminent GA is predicted. 28

Figure 4. Snapshot of a traffic simulation based on real flight data at Airport 2. After the missed-approach procedure has occurred, the call sign of the GA flight turns back to the original colour (yellow). 29

List of Abbreviations

A

A/C	*Aircraft*
AI	*Artificial Intelligence*
AIM	*Accident Incident Model*
ALA	*Approach and Landing Accidents*
ANSP	*Air Navigation Service Provider*
AP	*Airport*
ASCOS	*Aviation Safety and Certification of New Operations and Systems*
ATC	*Air Traffic Control*
ATCO	*Air Traffic Controller*
ATM	*Air Traffic Management*

B

BBN	*Bayesian Belief Nets*

C

CFIT	*Controlled Flight Into Terrain*
CW	*Calendar Week*

D

DBL	*Deep Blue Srl*
DFS	*Deutsche Flugsicherung*

E

EFSS	*electronic flight strip system*

F

FC	*Flight Crew*
ft	*Feet (unit of altitude)*

G

GA	*Go-Around*

I

IMC	*Instrument Meteorological Condition*
INX	*Innaxis*

L

LDG	*Landing*
LOC-I	*Loss of Control in Flight*

M

MA	*Missed Approach*
MAC	*Mid-Air Collision*
MAP	*Missed Approach Procedure*
MARIA	*Model of ATM Reality in Action*

N

NM	*Nautical Miles (unit of distance)*
NOTAM	*Notice(s) to Airmen*

R

RWY	*runway*

S

SID	*Standard Instrument Departure*

T

T/O	*Take-Off*
TUM	*Technical University of Munich*

U

UA	*Unstabilized Approach*

W

WP	*Work Package*

1 Introduction

1.1 Human Factors

Although many industries focus on improvements through technological solutions e.g. equipment, hardware and software, these technologies must be designed, operated, maintained and supported by people. This is generally the case regardless of the nature, degree of complexity or technological sophistication of the solution being considered.

History has shown us that if the human component of the system is not correctly considered from the start, from the conception of a technological solution through its development, integration, use and final disposal after use, many adverse consequences can occur. These include, but are not restricted to, increased risk of accidents and incidents, higher training costs, reduced performance and effectiveness, breaches in duty of care, scarcity of appropriately skilled personnel, delays to the development, and procurement project schedule and substantial increases in design or redesign costs. When focussing on the human component in the design of a system, the elements that are considered are known as Human Factors (HF), furthermore the Systems Engineering process, of effectively encompassing these in the design of the system, is called Human Factors Integration (HFI).

It is important to consider the human component in all aspects of the system, including the operational, organisational and management processes associated with the technological solution and its operators. One part of the system which is particularly critical for effective integration of the human, and one which generally receives most attention, is the man-machine/computer interface. Currently this is the most relevant part in the SafeOPS project.

1.1.1 A User Centric Approach

User-centred Design and *User-Driven Development* are approaches that seek to accommodate human needs within the design and development processes of technological solutions. They consider the user, e.g. operator, maintainer, as the main driver for design and development decisions by understanding the user, their demands, priorities and experiences. Often this process encourages an iterative design method in which prototypes are 'tested' on users and their feedback is integrated back into the design solution until an optimal design is arrived at. This approach ensures that the design of technological solutions is optimised for how the human can, wants, or needs to use the equipment, rather than forcing the human to change their behaviour and expectations to accommodate the equipment.

1.1.2 Human Factors in Aviation

One of the earliest examples of User-Centred design and the integration of Human Factors (or HF hereinafter, i.e. the characteristics attributed to the Human) in Aviation, is that of the flaps handle and landing gear lever in aircraft. In WW2, the United States lost a multitude of aircraft due to 'pilot error'. This was particularly pertinent to the Boeing B-17 Flying Fortress, in which the flaps and landing gear had identical switches that were co-located and were operated in sequence. Although the pilots were highly trained, they continually made the error, often in high workload scenarios like landing, of retracting the gear instead of the flaps. In 1942 the Army Air Force Psychologist, Alphonse Chapanis observed this and thus fixed a small rubber wheel to the landing gear lever and a small wedge-shape to the flap lever. Subsequently this kind of 'pilot error' almost completely disappeared. The landing gear lever and flaps handle are still designed this way today.

Although this illustrates the need for effective HF Integration in the physical design of equipment, it is equally as important, and one could argue more so, in modern systems that have increasingly higher degrees of complexity and autonomy, but reduced transparency. There is most likely always going to be a human component in all systems, no matter how much automation and 'responsibility' is afforded to the 'machine', and so the interface between the two should remain paramount in the design and development of such systems.

Aviation Human Factors therefore has a long history and is one of the key areas for the application of HF in all industries, alongside other forms of transport e.g. rail transport, medicine and the nuclear industry. The aviation community's interest in HF comes from the need and desire to improve safety and minimise accidents as a primary goal, and to maximise usability and commercial productivity as a secondary aim. Within the Air Traffic Management environment, close HF management is particularly critical due to the cognitive demands placed on Air Traffic Control Operators (ATCOs) resulting from an often high level of workload (peak amplitude but also sustained high level over a time period), a considerable urgency required in decision making, high safety criticality and a need for significant situational awareness, in a spatially and temporally complex environment.

Therefore, it is understandable that HF needs to be carefully considered when integrating any new technologies, systems or procedures into the ATM environment. Indeed, this consideration is pertinent when evaluating of the SafeOPS concept. The SafeOPS project is one which aims to explore the design and integration of a decision support tool for the time in advance prediction of missed approaches, and subsequent go-around (GA) of an aircraft when attempting to land. This decision tool classifies as Artificial Intelligence (AI) as it is based on big data and machine learning (M/L) algorithms, that provides probabilistic data to ATCOs and potentially effecting their decision-making process in the approach handling.

1.1.3 Probabilistic Information

One aspect of the SafeOPS project that calls for deeper analysis and contemplation in terms of its impact on HF and more significantly Human Machine Interface (HMI) design, is the fact the information being presented to the user is probabilistic. In SafeOPS, the tool being proposed will calculate the likelihood that an incoming aircraft approaching an airport performs a GA by comparing its flight data (e.g. aircraft type, speed, direction, descent rate, approach angle, weather conditions, etc.) with the historical data of previous traffic that landed or performed a GA at that airport.

There is a fundamental difference between the probabilistic information generated by the proposed SafeOPS tool and other sources of information presently available to ATCOs. For example, the radar screen displays data collected and transmitted by on-board equipment in real time, e.g. the call-sign, position, speed and direction of all aircraft in the airport area. Some of this information can potentially be incorrect in the unlikely event of a malfunctioning instrument, but the data itself refers to real, observed properties of the environment and of the agents moving in this environment. Other tools exist (e.g. the conflict-alert systems, cf. Sect. 2.1) that warn the ATCOs about imminent conflicts. This information is a *prediction* which results from a *deterministic* model. Such model calculates that, based on current data, if two objects continue moving along their trajectories and *nobody reacts to alter the course of events*, after a certain amount of time the two objects *will* collide in a certain point of space. By contrast, the SafeOPS tool is based on a *probabilistic* model. Consequently, the information generated by such a model has a certain probability to be correct and an inherent uncertainty associated with it. Hence the user of this information needs to decide whether they consider this uncertainty sufficiently small to rely on the prediction. The uncertainty inherently associated with the information generated by the proposed SafeOPS tool[1] is what makes it fundamentally different from other sources of information presently in use in ATM. The users, who are familiar with deterministic information (i.e. "something happened" or "something will happen if I don't react before") need not to be confused by this difference and need to learn how to process this information.

This is relevant to the user in the 'perception-decision-action-feedback' loop, which is a key principle linked to display and control design. Essentially, if the right information (regardless of its type) is provided to the user at the appropriate time, the user stands a reasonable chance of making a good decision. A decision which then needs translation into an action through the provision of easy to use controls. Finally the user needs to know if what they did was what they intended to do and that there has been an effective outcome from their action. Indeed, a lot of the role of an ATCO is in the tasks of monitoring and supervisory control of the air traffic 'system', however when non-standard and emergency situations occur, an ATCO is required to fulfil a role that incorporates more of the 'perception-decision-action-feedback' principle.

Consequently, when the SafeOPS concept is providing information to the ATCO as a decision support, the accuracy of this data needs to be high and the presentation of the information needs to encourage the correct understanding of its meaning by the ATCO, in order for it to be useful in ensuring a safe, timely and effective action is undertaken by them. Moreover, the safety criticality of ATM calls for reliable information, in order to make appropriate decisions. When any level of automation is introduced into a system, this generally reduces the transparency in a system; which can lead to problems with the user understanding what the system is doing or trusting it, and subsequent difficulties when the automation fails for any reason.

Nonetheless, whilst there is a comprehensive regulatory framework for ATM through the likes of the International Civil Aviation Organization (ICAO), European Aviation Safety Agency (EASA), the Federal Aviation Authority (FAA) in the U.S, country specific Civil Aviation Safety Authorities, etc., ground-based ATM systems are not required to be formally certified in the same way that aircraft avionics are. This is because ATM systems remain decision support tools with very limited automation [1]. This is also true for the SafeOPS concept in which it would be classified as 'Level 1 AI/ML: assistance to human' on the EASA classification of AI/ML applications scheme [2].

Probabilistic information and machine learning are constructs used in AI where the level of automation and autonomy of the machine can be considerable high. Indeed, the more autonomous an AI technology is, the stronger the HMI design needs to be e.g. highly visible and transparent, rapidly accessible, allows the user to abort the machine proposed automation etc. However, the ATM industry is inherently conservative and integration of AI/ML is slow in this area. It is believed that its integration will most likely be driven by specific use cases which exist at the interface between ATC and other more innovative aviation sectors, e.g. unmanned aerial systems traffic management.

There is an inherent problem in the reporting of probabilistic information in that it is potentially too complex for a user to understand its meaning and process it in a rapid and correct manner. As an example, what does 80% mean on an ATC display, compared to 85% or 99%. Usually, probabilistic

[1] And other similar predictive tools based on machine learning.

predictions would serve to inform some level of autonomy, in which the user did not need to interpret the predictions, but the AI would interpret the predictions and take the appropriate action.

With the introduction of increasing complexity or autonomy in a system, trust in the technology by the user is called into question. “Trust is… a social construct that becomes relevant to human–machine relationships when the complexity of the technology defies our ability to fully comprehend it” [1].

This highlights that trust in an AI system is not just a result of trustworthiness or reliability of the system. Trust is more a response and/or acceptance of the system. This is different from trustworthiness of the system in the sense that the latter is easily ascertained and regulated through the Validation, Qualification and Certification (VQ&C) process. Therefore, it is not sufficient to simply deliver a system that is near to 100% reliable and expect a user to trust it. There is more of a need to engineer a solution where the machine human interaction is optimised for trust by using human factors to inform the design. Three of the basic components of an effective HMI, in relation to engendering trust are transparency (in more recent literature this is referred to as ‘explanatory ability or explainability’), minimal system failures, meaningful error reporting.

1.2 Scope and limitations of the present study

This document aims to consider the human component (Human Factors) of the integration of a machine learning tool, which produces probabilistic data in predicting a possibly missed approach and subsequent go around.

For the purpose of the SafeOPS project, the HF being considered is largely associated with the Human Computer Interaction (HCI) aspect. This is because the interface between the SafeOPS tool and the human will be at the point of the display systems currently being used for the role of air traffic control. According to this, the scope in this document encompasses the display design from a software perspective (not hardware) including any auditory components of the software.

When considering a method to effectively design and evaluate the HCI of such a system, one can consider the following as high-level Human Factors Systems Requirements for the use of the HCI:

REQ ID	Title	Requirement	Rationale
REQ-SafeOPS 22-TS-HFSR-001-1	Cognitive Capabilities	The system shall accommodate the cognitive capabilities of the specified user population	This requirement ensures that the system is designed to support human information processing, addressing complex human operations such as decision-making, the limits of human attention and time-sharing, the human component of situation awareness and mental workload.
REQ-SafeOPS 22-TS-HFSR-001-2	System Inputs	The system shall provide appropriate means for the human to make inputs to the system	This requirement ensures that the system is designed to provide the most appropriate means of making control inputs and the configuration of the control device, whether mediated through a physical device or through a graphical user interface. The focus of this requirement is on the transmission of user intent to the system and the feedback that the system provides.
REQ-SafeOPS 22-TS-HFSR-001-3	Information Presentation	The system shall provide appropriate means of information presentation to the human user	This requirement ensures that the design of information displayed to the user and the physical display technology or material on which the information is conveyed is appropriate to the physical and cognitive abilities of the human, the task requirements and the context in which the information is conveyed. It therefore includes aspects such as visibility, discriminability and meaning. It includes all type of information, including both visual and auditory.

REQ-SafeOPS 22-TS-HFSR-001-4	Sensory Characteristics	The system shall accommodate the sensory characteristics of the specified user population	This requirement ensures that the design of the system accommodates the characteristics of human vision, hearing, as well as other sensory characteristics such as smell and touch.
REQ-SafeOPS 22-TS-HFSR-001-5	Communications Needs	The system shall accommodate the communication needs of the specified user population	This requirement ensures that the system is designed to support human means of communicating (including verbal and non-verbal means).
REQ-SafeOPS 22-TS-HFSR-001-6	Human Error	The system shall integrate humans in ways which minimise the opportunity for human error	This requirement ensures that the system is designed so that it does not induce either human errors or behavioural violations and allows the system to recover from those errors where they are committed.
REQ-SafeOPS 22-TS-HFSR-001-7	Physical and Cognitive Stress	The system shall protect the human from physical and cognitive stressors resulting from system use	This requirement ensures that the human is protected from both physical stressors arising from a number of sources (e.g. external environments, workplace environments and physical workload) and cognitive stressors such as mental workload and time pressure.

Under these high level system requirements there are many lower-level requirements that guide the designer in achieving the high level aims. These are described later in the document (section 4.1) when the SafeOPS design is evaluated and discussed.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

1.2.1 Assumptions and Limitations

Several assumptions and limitations exist. Because the SafeOPS projects is considering the integration of the novel technology into current systems, it is assumed that the physical, and environmental integration of the human in the system has already been effectively managed. This includes workstation design, placement of physical controls and placement of displays in the workstation, data entry and control devices, communications equipment, seating position, lighting and illumination, temperature and humidity and noise and vibration. Therefore, none of these will be considered in the integration of the SafeOPS tool unless there is something very specific and relevant to the SafeOPS tool.

Limitations on the integration of the SafeOPS tool are the boundaries set by current procedures for conducting any tasks in the context of the go-around e.g. go around procedures, and any conventions and regulations for ATM system design characteristics. In ATC there are very specific procedures for undertaking air traffic operations and it is important that the integration of a new technology adheres to them or clearly defines where they would need to be adapted. This would thus have an impact on training. Although important, the operating procedures and training protocols however are not going to be considered in this review, unless they are specifically relevant. Nor are personnel or organisational aspects being covered.

1.2.2 Structure of the document

The structure of the document takes the following format:

- An introduction to HF and the user centric approach to design describes human factors specifically in aviation and ATM, an overview to probabilistic data and how this relates to the human component in the system, the scope of the report and the assumptions and limitations
- Subsequently, a review of existing ATM tools presents the design and use of the current tools, any details on the provision of risk information, identified benefits in their use and any known issues and limitations associated with them which may also apply to the design and implementation of the SafeOPS tool.
- A description of the SafeOPS tool follows with details of its evaluation by ATCO and HF experts, covering the workshops conducted and the outcome of a Hierarchical Task Analysis and results from the display design evaluation.
- Discussion and Recommendations; Guidelines on how to present the information of the SafeOPS tool to the user from a HF perspective and lessons learned from more literature about HF with AI-based tools.

2 Literature review of similar tools

This chapter introduces the reader to other tools in the ATM system which are implemented for the purpose of increasing safety. They have been reviewed here in order to ascertain any aspects of their design and functioning which are relevant for the design of the SafeOPS tool.

Air traffic controllers are assisted by several tools and services that support the detection and identification of emerging safety issues. ‘Safety Nets’ is a term that describes a group of tools that are either airborne or ground based which, “help prevent imminent or actual hazardous situations from developing into major incidents or even accidents”[3]. Those which are ground based are an integral part of the ATM system and serve to bring the ATCOs attention, usually through a series of alerts, to a potentially hazardous situation which he/she is expected to act on, issuing clearances and instructions to prevent or resolve conflicts. All ground-based Safety Nets operate with an alerting system which has a ‘look-ahead’ time (relative to the potential conflict) of 2 minutes. In the current configuration of ATM systems, the following exist under the category of Safety Nets:

- Short Term Conflict Alert (STCA)
- Area Proximity Warning (APW)
- Minimum Safe Altitude Warning (MSAW)
- Approach Path Monitor (APM)

There are also other conflict alerting tools akin to the STCA, which operate at a longer ‘look ahead’ time than 2 minutes and so aren’t considered ‘Safety Nets’, however they are important ATM tools to introduce in this chapter. They are specifically the ‘Tactical Controller Tool’ and the ‘Medium Term Conflict Detection’ tool. The review commences with the introduction to the reader, of the conflict alerting tools and concludes with the introduction of the other Safety Net tools.

2.1 Conflict Alerting Tools

Air Traffic Control systems are largely human centric in which a highly trained individual, the ATCO, is required to cognitively process a continuous stream of information to manage air traffic through the issuance of clearances and instruction to prevent conflicts between aircraft. However, cognitively processing large amounts of complex information is susceptible to vulnerabilities including errors, biases and the channelling of attention. As a result, deviations by an aircraft from clearances or instructions can go unnoticed and can result in reduced aircraft proximity. There are tools in the ATM system which warn controllers of conflicts and cover increasing ‘look ahead’ times. The three most-commonly-implemented are the following: the Short-Term Conflict Alert (STCA), the Tactical controller tool (TCT), and the Medium-Term Conflict Detection (MTCD). Their main characteristics and look-ahead times are summarised in Table 1 and briefly discussed as follows.

Table 1. Main properties of the three most common types of Conflict Alert systems.

Tool	Typical Look ahead time	Input	Use
Short Term Conflict Alert (STCA)	1-2 minutes	Surveillance data (actual flight level derived from SSR mode C and assumption that the aircraft will maintain its present track). There are some implementations that also include pilot input received via SSR mode S or controller input of CFL.	Safety net - to warn the controller that a separation breach is imminent.
Tactical controller tool (TCT)	Up to 5-8 minutes depending on local implementation	Surveillance and/or FPL data, depending on local implementation.	Conflict resolution and clearance verification tool - to inform the controller if the current clearance results in a separation breach, so that it can be corrected if necessary.
Medium Term Conflict Detection (MTCD)	Up to 20-30 minutes depending on local implementation	Flight plan data and controller input (clearances).	Planning tool (to inform the controller of potential conflicts so that a proper plan is made in advance).

2.1.1 Short Term Conflict Alert (STCA)

Short-Term Conflict Alert (STCA) capability can provide an independent technology-based solution through the insertion of an alerting logic in the control loop. This alerting logic would draw the attention of the ATCO to existing or pending situations in which user defined minimum separation distances, between any pair of surveillance tracks is, or is predicted to be violated in the very near future. As such the STCA capability is intended to function in the short term, providing warning times up to 2 minutes.

The STCA uses deterministic data to present actual or predicted conflicts. It uses surveillance information derived from radars, ADS-B or multilateration (MLat) as well as environmental data and flight plan information to predict the movement of aircraft

A STCA is in fact a concept rather than a specific system and so there are differing implementations of this concept in use in the ATM environment across the world. As a concept, it is defined as a ground-based safety net, with the sole purpose of monitoring the environment of operations, in order to provide timely alerts of an increased risk to flight safety, which may include resolution advice. Most STCA's comprise of a visual alert on the radar display and sometimes, the addition of an accompanying audible alert. Usually, an alert is triggered when an aircraft does not meet the separation criteria (5 nautical mile longitudinal, 1000 ft altitude), leading to a flashing red 'ALRT' notification on the aircraft track label (Figure 1).

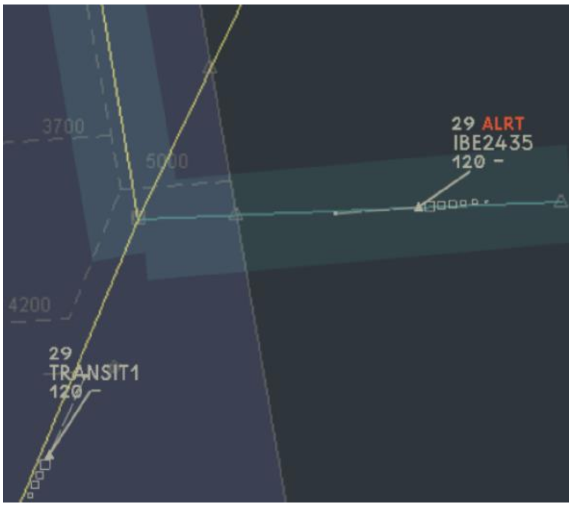

Figure 1. Current notification alert design: a blinking red [4].

An aspect of human cognition that the STCA aims to provide a safety net for, is that of confirmation bias. Human beings tend to look for information that 'confirms' and supports already made assumptions and decisions, and assume it to be true information. At the same time as avoiding information that does not conform to their beliefs or assumptions, thus rejecting it as false or irrelevant. The STCA presents alerts about the real-life situation when it is clear that the ATCO may have overlooked or formed a cognitive bias about the current air-picture.

The STCA is essentially an alerting capability and such systems usually have distinct HF considerations associated with them. This is in order to optimise system usability and to ensure acceptance of the tool amongst the ATCO user group. These considerations are encompassed as recommendations in STCA implementation guidance material [5]. Stated as a minimum, the STCA implementation should be optimised for the following:

Appropriateness and Timeliness

The rules set for generating alerts should be appropriate to the role of the ATCO and other systems in place. The addition of alerts should take into consideration the current alerting hierarchy. Its design should avoid any disagreement with or interruption of the current control practices and decision support tools currently in place. Timeliness of the alerts are critical for provision of sufficient warning time to resolve hazardous situations. Moreover, the number of nuisance alerts needs to remain below an acceptable threshold.

Effectiveness

The design should ensure that the ATCO becomes aware of, through the alert, the reasons for a potential conflict, as the controller may have already failed to detect the potential conflict earlier, by not perceiving a risk in that particular scenario.

Comprehensibility and Performance Monitoring

The integration and continued use of an STCA system should be supported by initial and continuation training, to address increasing complexities of STCAs and the environment. How the ATCOs ultimately use the STCA, and their perception of its effectiveness, should be periodically evaluated and also following any changes to the design. Lessons learned in accidents and incidents in which the STCA was involved should be shared amongst relevant communities, e.g. design, regulatory, training etc.

Known issues and limitations of STCA

It has been shown that the alerting mechanism employed by the STCA, consisting of a small blinking red" ALRT" tag, could remain undetected, especially under high workload settings [4], [6]. Also the algorithms underlying STCA vary in their sophistication. Some, fail to consider the planned or cleared route changes an aircraft is to make. Instead, they perform simple straight-line extrapolation of flight path. This can obviously result in a false alarm if an aircraft intends to level off directly under another [7].

2.1.2 Tactical Controller Tool (TCT)

Tactical Controller Tool (TCT), operating in both enroute and terminal manoeuvring area (TMA)[2], warns the executive (tactical) controller of potential conflicts within the sector. It combines current aircraft tracks with a tactical trajectory based on the aircraft's current behaviour, although some implementations are based only on surveillance data.

The TCT acts as a separation assurance aid and aims to reduce workload for the tactical controller by providing accurate monitoring and conflict detection. It helps not only in detecting problems but also in showing that no problems exist. Warnings from the system are usually provided in both the vertical and horizontal planes and when no TCT warnings are indicated, the controller can be assured that no potential conflicts exist at that time. TCT also indicates when a "critical missed manoeuvre" could occur, these are situations in which loss of separation would occur if an aircraft failed to make a planned manoeuvre.

TCT is designed to fit with the MTCD sub-system and has a look ahead time of five to eight minutes. However the look ahead time may be extended to match the MTCD's. TCT may be activated manually (by selecting the aircraft concerned) or automatically (when the required criteria are met).

The main *benefits* of the TCT are as follows:

Time saving: the system calculates the conflicts, which allows the controller to focus on other tasks (e.g. hear-back, planning, scanning, etc.).

Reduces unnecessary controller interventions: TCT is usually more precise than the combination of standard range and bearing measurement tools and manual calculation. As a

[2] Designated area of controlled airspace surrounding a major airport. It is also known as Terminal Control Area (TCA) in the U.S.

result, in many cases the aircraft are subject to simpler (e.g. locking of headings) or no intervention.

Mitigates "Blind Spots": TCT may discover most short-term conflicts.

Provides continuous feedback on clearance effectiveness: in case the controller chooses to apply vectoring, the TCT immediately evaluates the new heading(s).

Enables the controller to perform "softer" interventions: if the controller calculates the conflict mentally, they usually apply techniques to accommodate errors (e.g. round values to worst case scenarios, "safety buffers" for measuring tools to compensate the uncertainty of the closest point of approach, etc.). The ATS system calculates the closest point of approach with great accuracy which allows the controller to issue clearances that result is smaller aircraft deviations.

Resolution advisories: the TCT may be equipped with an advisory feature that suggests action for conflict solving (e.g. "ACFT A should turn 10 degrees to the right").

2.1.3 Medium-Term Conflict Detection (MTCD)

Medium-Term Conflict Detection (MTCD) is a flight data processing system for providing a warning of a potential conflict between flights in the controller's area of responsibility, up to 20 minutes ahead of the conflict.

MTCD is a suite of predictive tools that predict and notify on the probable loss of the required separation between two aircraft and aircraft penetrating segregated or otherwise restricted airspace, but also aircraft-to-aircraft encounters where each aircraft is blocking airspace that might have been used by the other. As such, akin to the STCA, the term MTCD isn't a specific piece or collection of equipment, but it is any system designed to achieve the goals mentioned here.

The motivation behind the development of MTCD included the following:

I. To move from a largely reactive form of air traffic control to more pro-active control;

II. Balancing more evenly the workload associated with tactical and planning tasks;

III. Enhancing sector team efficiency.

The key difference between MTCD and the STCA concept is that STCA is a safety net function for the improvement of safety of ATC service provision, however MTCD is a controller tool.

The main functions of the MTCD are as follows:

Trajectory prediction: The creation of future trajectories for each aircraft;

Conflict detection: The identification of potentially conflicting trajectories, where future position of 2 or more aircraft might fall below specified minima (not necessary the separation minima);

Trajectory update: The updating of predicted trajectories, this functions is performed following an external input (i.e. human intervention) or due to a change in aircraft trajectory and/or automatic integration of detected aircraft position;

Trajectory edition: Allowing the human interaction with the predicted trajectory of one or more aircraft.

2.2 Area Proximity Warning (APW)

It has been documented that in the past some actual or potential hazardous situations related to aircraft position have remained unnoticed by ATCOs. Therefore, the purpose of the Area Proximity Warning (APW) system is to independently warn the controller about unauthorised penetration of an airspace volume or a future situation, based on the position and speed of an aircraft relative to that airspace volume. The APW system can also provide further detail on the aircraft status e.g. controlled flights moving into restricted airspace or uncontrolled flights moving into controlled airspace. APW is intended to function in the short term, if applicable providing warning times of up to 2 minutes.

Known issues and limitations of APW

With past implementations of APW [8], the system alerts would annunciate for all aircraft tracks regardless of flight status (assumed by the controller or not) and SSR code, including Operational Air Traffic and General Air Traffic. Also, the system would alert on predicted entry into an APW area as well as on exit. This would result in too many alerts being displayed to controllers who were not working the aircraft.

When the TSA map was displayed it overrode the corresponding APW area (which would automatically become an eggshell white when it became active) hiding the APW boundary area. This means that if a controller is already displaying the TSA map for an APW area on his screen, there is no visual cue when the APW area activates or deactivates. Furthermore, the APW label was always drawn at the most easterly corner of the defined shape. Because some of the polygons share the same co-ordinates, this would lead to the overlapping of APW labels, which were then unreadable.

The APW areas changed between the active and non-active state with no visual pre-warning for the controller. The APW area simply appeared on the screen when it became active, and automatically disappeared on deactivation.

There were a small number of situations, where an alert was generated even though the aircraft was not going to proceed into the area. Some of these nuisance alert situations would be predicted to clip the corner or edge of an APW area.

It is not known if these issues exist in more recent implementations of APW systems, as these are taken from documentation from 2009, however they do highlight some of the problems associated with attempting to integrate such systems and the unpredictable nature of integration issues.

2.3 Minimum Safe Altitude Warning (MSAW)

The Minimum Safe Altitude Warning (MSAW) is the most common ground warning system in use. MSAW is a software programme linked to the radar data processing system that continuously and automatically compares the Mode C readout of all aircraft that meet the MSAW criteria (e.g. only IFR flights) operating in a defined zone (MSAW zone). The system alerts controllers, either acoustically or optically, or both, if a Mode C level indication is detected below a pre-determined altitude. This altitude is normally 400 or 500 feet below the MFA or MRA applicable, to avoid false and frequent alerts. An

MSAW alert is transmitted immediately to an aircrew for altitude correction, with the advantage that controllers can offer enhanced information with regard to an escape route or suggest a heading toward lower terrain.

Guidance on implementing a MSAW system suggest that the following should be explored and designed with the users input [9]:

- The presentation of alert messages to the display of the controller: visual and/or aural;
- In order to avoid a large number of false MSAW alerts, ATCOs of MSAW equipped units must be able to deactivate certain SSR tracks and suppress MSAW alarms for aircraft with incorrect mode C read out and aircraft that are performing procedures such as a leaving or joining of an IFR flight plan, or executing visual approaches or visual climbs with full reference to terrain;
- Test function, which is very important in order to ascertain that at all times, the MSAW feature is operational. Each software change should be tested;
- Special attention must be given to zones around airports in the final approach and initial climb out sectors where fine-tuning is necessary in order to determine an acceptable solution between real alerts and a high number of false alerts. MSAW should be customised to each airport;
- Important to involve controller representatives during developing, testing, tuning, implementing, training and introducing MSAW systems in ATC units;
- Controllers must have guidelines and standards (eg standard R/T phraseology) on how and when to use MSAW systems

2.4 Approach Path Monitor (APM).

An Approach Path Monitor (APM) is a ground-based Safety Nets intended to warn the controller about increased risk of controlled flight into terrain accidents by generating, in a timely manner, an alert of aircraft proximity to terrain or obstacles during final approach.

APM should be used for both precision and non-precision instrument approaches. However, APM is not expected to be used for circling approaches.

This Safety Net uses surveillance data, flight plan data (specifically the destination airport) and user-defined limits of the acceptable runway approach paths to achieve its purpose.

2.5 AI-based tools investigated in other ongoing projects

Three other projects, which received funding from the SESAR Joint Undertaking under the Horizon 2020 programme, are currently investigating the benefits and shortcomings of introducing AI-based technology in the routine of aviation actors. Although the maturity level of these projects is relatively low, we summarise their main outcomes at present, as they can potentially produce exploitable knowledge and understanding relevant to SafeOPS.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

HARVIS (2019 – 2021)[3] investigated the use of AI based digital assistants able to support pilots especially in decision making processes in complex situations. To do so, the project evaluated two use cases: a tool able to suggest the best alternative route and airport in case of on-board emergencies (e.g. technical failures, medical problems) and a tool suggesting the go around manoeuvre. Test have been carried out in pilots in the loop simulations, highlighting on one side experts' interest in that type of support and the feasibility of such tools, on the other side the low trust in AI, to be enhanced by training and by the post-flight analysis of AI data and the reasons behind the proposed solutions

MAHALO (2020 – ongoing)[4] investigates the impact on acceptability and performance of two different approaches: AI explainability and conformity of solutions proposed by AI. To do so the project is carrying out simulations with ATCOs testing a conflict prediction and resolution tool based on AI trained to maximise specific parameters (e.g. mileage, fuel consumption) and on AI trained to mimic real ATCOs behaviour. The first AI can propose solutions that are distant from the way ATCOs normally deal with conflicts but more "optimal"; different levels of explainability provide the ATCOs with the motivation of the proposed solution. The second one proposes solutions that may be less optimal but are easier to be accepted. Results will be available by the end of the year.

Artimation (2021 – ongoing)[5] investigates the impact of combinations of different types of explainable AI algorithms and visualisation techniques, and aims at evaluating the impact on ATCOs acceptance and performance. In the project two different tools are developed and used as use cases to test experts reaction: a conflict prediction and resolution tool and a delay propagation prevision tool. Simulations with ATCOs will investigate different levels of explainability (e.g. black box vs. domain based explanation) and different visualisation types (e.g screen based vs. 3D vysor based). Results will be available by the end of the year.

[3] http://www.harvis-project.eu/

[4] http://mahaloproject.eu/

[5] https://www.artimation.eu/

3 The SafeOPS tool

Chapter 2 described ATM safety tools that currently exist (or are being studied), largely to represent a protective mechanism for any cognitive biases or attention narrowing of the ATCO and to assist them in being fully aware, in a timely manner, of any forthcoming conflicts, excursions or potentially hazardous situations. The knowledge of how the current ATM systems have been implemented and any issues associated with them, can inform and influence the design of new ATM systems and, in this current study, to the design of the SafeOPS tool. In this chapter, the SafeOPS tool and design evaluation is described in more detail, followed in chapter 4 by guidelines for the design of the tool based on the feedback from ATCOs, literature and HF best practice.

3.1 Brief description of the tool

This section summarizes the SafeOPS concept to the point available at the moment of writing this deliverable. Therefore, this section is based on the information of section 4 of D2.1 [10] available on https://safeops.eu/ and deliverable D4.1, which is submitted in parallel to this document.

The SafeOPS concept, following the *from prediction to decision* paradigm, envisions better informed decisions by ATCOs in go-around scenarios by providing probabilistic and time in advance information on how an approach is likely to evolve – as landing or go-around. The basic concept therefore is based on two distinct parts:

- The big data based predictive tool and
- The visualization of the predictions in the radar screen of the ATCOs.

3.1.1 Big Data based Predictive Models

The big data related requirements defined for SafeOPS were detailed in D2.1 section 3. Based thereon, a technical problem statement for a data driven predictive tool is formulated. This problem statement tackles the following tasks which were foreseen to be essential by the writing of D2.1:

- Data acquisition
- Building a data pipeline that automates:
    - Data cleaning
    - Data fusion
    - Feature analysis
    - Data labelling
- Training and evaluation and benchmarking of predictive algorithms for the SafeOPS use cases.

Based on the documents stated above, D4.1 describes the technical implementations of the first version of the predictive tool. To avoid long duplications, this section will summarize the important results of D4.1 for D3.3, being the benchmarking and evaluation of the predictive algorithms.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

In D4.1, a data set containing 32 features is created at locations from 4NM – 10NM from the runway threshold is created for both airports under investigation in SafeOPS. The features include per flight information on the flight itself, weather, approach performance, airport information and a label indicating a go-around or landing. It is described in detail in D4.1 section 5.2.3. Based on this data set, several machine learning models are trained for predicting a go-around to happen based on the feature data set as input. This means, that at the point of 4NM from runway threshold, the model will indicate whether an approach under investigation will perform a go-around or not. Out of initially nine different models, D4.1 selected the one with the best results regarding accuracy and reliability. Section 5.4.2 in D4.1 describes the benchmarking process, showing that the Gradient Boosting (Light GBM) method performed the best for both airports.

For Airport 1, the results summarized in Table 2 were obtained in D4.1. The two important metrics are the precision and recall for both the go-around (true) and landing (false) case.

The recall value of 0.4281 for the true go-around describes that the chosen model labels on average 4281 out of 10000 go-arounds as such. This on the contrary means that 5719 out of 10000 go-arounds are not detected by the predictive algorithm.

The precision for the true go-around case of 0.8974 describes that on average, from 10000 cases the model predicts a go-around to occur, around 8974 approaches will perform a go-around, whereas in 1026 cases the aircraft will perform a landing.

Similarly, for the false go-around case or landing case, the recall of 0.9998 indicates that for 10000 landings, the predictive model will label 2 go-arounds and 9998 landings.

The precision metric of 0.9980 for the landing case describes that in case the tool predicts a landing, out of 10000 flights, on average 20 will perform a go-around whereas 9980 will perform a landing.

Table 2. Precision and Recall of GA events from traffic data at Airport 1.

	Airport 1	
Go-Around	**Precision**	**Recall**
True	0.8974	0.4281
False	0.9980	0.9998

For Airport 2, equivalent results to Table 2 for Airport 1 are summarized in Table 3.

The recall value of 0.2135 for the true go-around describes that the chosen model labels on average 2135 out of 10000 go-arounds as such. This on the contrary means that 7865 out of 10000 go-arounds are not detected by the predictive algorithm.

The precision for the true go-around case of 0.9111 describes that on average, from 10000 cases the model predicts a go-around to occur, around 9111 approaches will perform a go-around, whereas in 889 cases the aircraft will perform a landing.

Similarly, for the false go-around case or landing case, the recall of 0.9999 indicates that for 10000 landings, the predictive model will label 1 go-around and 9999 landings.

Table 3. Precision and Recall of GA events from traffic data at Airport 2.

Airport 2		
Go-Around	**Precision**	**Recall**
True	0.9111	0.2135
False	0.9972	0.9999

The precision metric of 0.9972 for the landing case describes that in case the tool predicts a landing, out of 10000 flights, on average 28 will perform a go-around whereas 9972 will perform a landing.

For both models, we can observe that they are very selective, meaning that a lot of go-arounds are not detected as such. While for Airport 2 only 21% of go-arounds are detected as such, for Airport 1 42% are captured. This can have several reasons. First of all, it is the first model developed in SafeOPS, so the feature engineering process is far from perfect still. Furthermore, the data sources used are ADS-B and METAR data, both being rather sparse in the temporal domain. ADS-B data provides performance information roughly with a 10s resolution, METAR is released only every 20 minutes. Since however, the predicative model is supposed to be deployed in a Tower and QAR data are not available in real time, no other data source is available to SafeOPS to demonstrate a real time prediction, although MLAT would be a possibility to refine the temporal resolution. Also, ADS-B also in combination with METAR solely captures the aircrafts performance on the approach and can thus only possibly cover pilot induced go-arounds and therefore does not provide possible features or contributing factors for controller induced go-arounds.

If the models however predict a go-around to occur, the chances that the approach under investigation is actually performing a go-around are around 90% for both models. For the landing case Recall as well as Precision values are above 99%. This on the first glance looks very good but has to be brought into perspective regarding the imbalanced prior probability of go-arounds and landings per approach. In the data for both airports we found that from 10000 approaches 35 perform a go-around. Therefore, for the Airport 1 case out of 10000 approaches and thus 35 go-arounds, the predictive model will label 14-15 correctly as go-arounds and 20-21 falsely as landings. From the remaining 9965 landings, 9963 will be correctly labelled as landings and 2 will falsely be labelled as go-arounds. Similarly, for 10000 approaches at Airport 2, 7-8 out of 35 go-arounds are labelled as such, with remaining 27-28 not detected and labelled as landing. For the 9965 landings, 1 will be wrongly labelled as go-around and 9964 will correctly be labelled as landing. Therefore, for both, Airport 2 and Airport 1, for 8 correctly labelled go-arounds 1 approach, labelled as landing, will perform a go-around.

A detailed discussion about how to improve the performance of the predictive model is beyond the scope of this deliverable and is presented in D4.1, in which the algorithms developed for the first ML case study and their limitations are thoroughly described. In the following, we summarise the main aspects and we refer the reader to D4.1 for a complete discussion. Different approaches are possible to improve the performance of the predictive model. The first and most obvious one, although not necessarily the simplest, is to get more data. Although go-arounds are a rare event and therefore the imbalance problem in the dataset would remain, increasing the data level could help improve the overall performance of the model. Treat missing and outlier values. The unwanted presence of missing and outlier values in the training data often reduces the performance of a model. Modifying the current pre-processing stage could achieve a performance improvement. The feature engineering phase can also be further explored by developing new features to extract more information from the data or to obtain information that better reflects the operational scenario. In addition, a feature selection process can also be carried out, whereby features that do not add value to the prediction are eliminated. Finally, the use of more powerful models can also be explored, although usually at the cost of losing explainability. For a complete discussion on the results of the predictive models and in particular for a detailed description of possible strategies to improve the performance, the reader is kindly referred to Sect. 5.4 of the deliverable D4.1 of SafeOPS.

.

3.1.2 Visualization of Predictions in the Radar Screen

For the SafeOPS workshops, a python tool which mimics the radar of Tower Controllers is developed. Information from the AIP Germany, the HMI requirements defined in SafeOPS' deliverable D2.1 and input form ATCOs in the SafeOPS workshop guide the development. The tool allows to visualize the Scenarios and Use Cases, defined in D2.1 - Section 3, adding the additional information provided by the data-driven go-around prediction tool.

The objective of the visualization is to help ATCOs understand how the SafeOPS solution could be indicated on the Radar screen, and help the development team understand the potential benefits and threats of such a tool. The visualization tool can be fed with either real world recorded data such as FDM or ADS-B data, but also with data generated from aircraft simulators or aircraft models. Therefore, the desired trajectories needed to simulate the scenarios can be obtained via several methods.

In the following figures, airport1 is shown, with the CTR boundary visualized by a dotted polygon. In the center of the screen the 4 runways are illustrated in thick, white lines, while the thinner dashed lines represent the runway center line extensions. Each dash is 1 NM long, and the separation between the dashes are also 1 NM each.

The figure below shows two aircraft, one departing aircraft on RWY 25C, indicated by a white diamond, and an arriving aircraft indicated by a yellow diamond. The colour yellow is used to highlight aircraft, which are established on the ILS. To make the graphics more realistic, a call-sign is shown next to the marker, with corresponding flight level and airspeed.

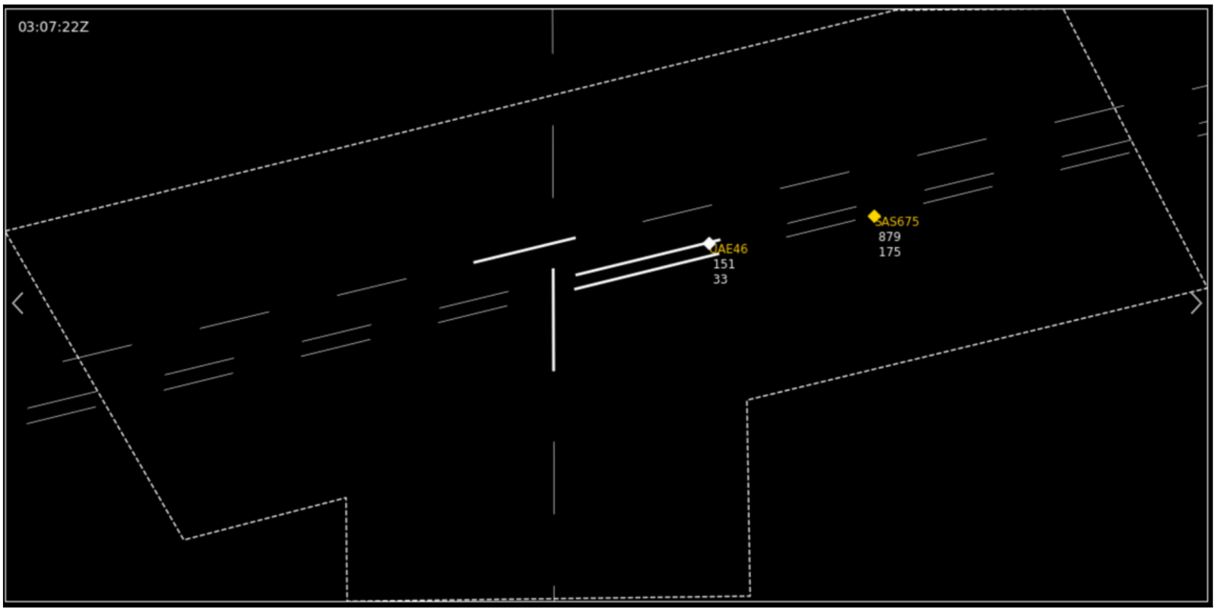

Figure 2 Snapshot of a traffic simulation based on real flight data at Airport 2.

The clock in the top left corner shows some imaginary time. Its whole purpose is to give the ATCOs a sense of time since the animation speed is variable.

Once the prediction tool detects a go-around, the yellow-coloured marker changes its color. In this example, the color is set to red however the final color to indicate a predicted go-around is not yet finally decided.

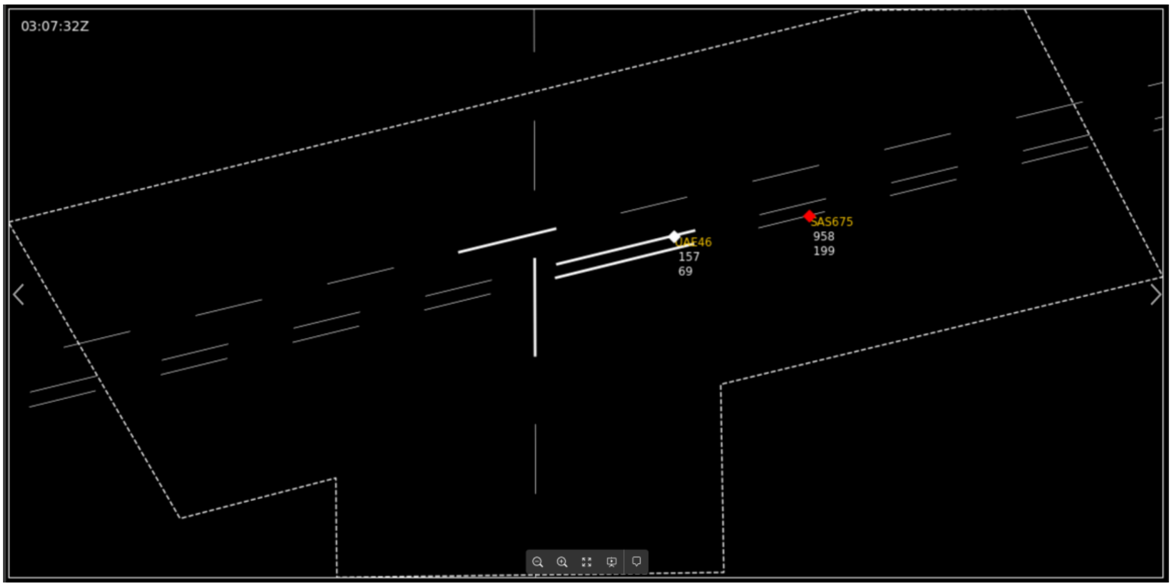

Figure 3. Snapshot of a traffic simulation based on real flight data at Airport 2 in which an imminent GA is predicted.

Once the aircraft reached the missed approach altitude, the marker is indicated in white again.

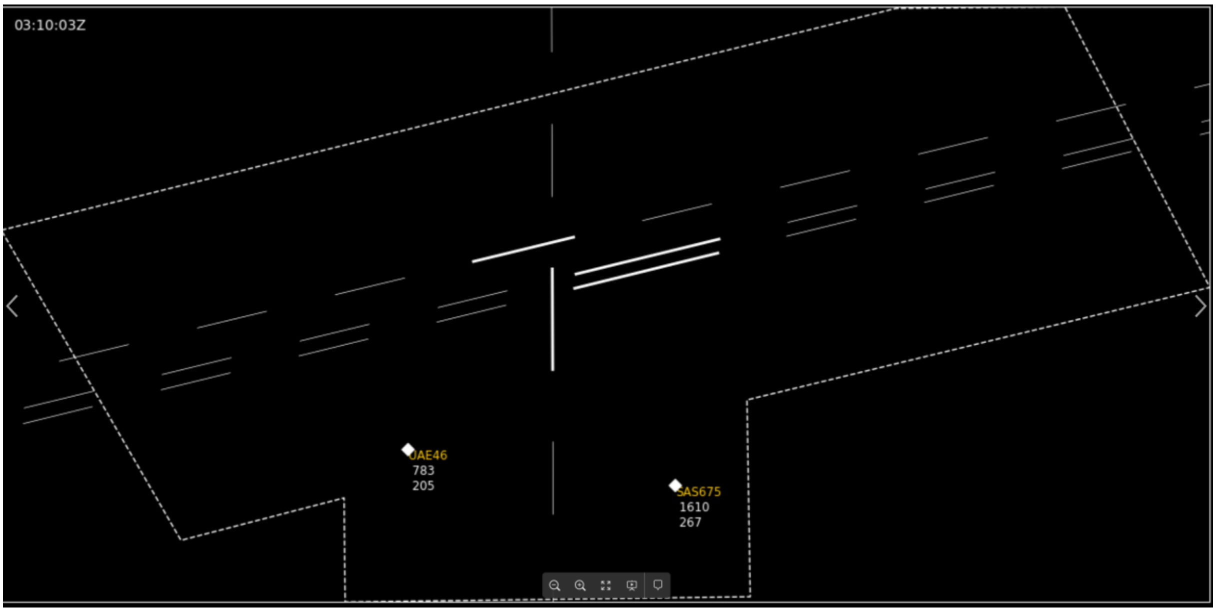

Figure 4. Snapshot of a traffic simulation based on real flight data at Airport 2. After the missed-approach procedure has occurred, the call sign of the GA flight turns back to the original colour (yellow).

Features of the Visualization Tool

According to the HMI requirements of D2.1 various features for the visualization are implemented in the visualization tool.

These features are summarized below:

- Variable coloring of the aircraft marker, depending on the confidence of the go-around prediction
- Variable indication of additional information like precursors
- Surrounding Traffic Indication, to provide a feeling on overall traffic situation
- Confidence in prediction of a go-around

3.2 Expert evaluation

3.2.1 User feedback

One of the objectives of the Human Factors analysis performed in Task 3.3 is to shed light on how the provision of probabilistic information changes the tasks the ATCOs perform and how the display of such information influences their decision making and subsequent actions. For this purpose, a series of eleven workshops were conducted between July 2021 and February 2022 with the aim of performing the following activities:

- *Hierarchical Task Analysis (HTA):* Identifying all actions the ATCOs take when handling traffic. These actions include the regular monitoring of the RWY visually and on the ground radar, the communications and instructions issued to the pilots, the coordination with other ATCOs, the insertion and update of flight status information in the local electronic flight strip system

(EFSS), the constant monitoring of the traffic situation of departing and landing aircraft visually and on the airspace radar.

- *HTA when the predictive tool is in use*: Reviewing the results of the above-mentioned task breakdown to investigate how, with the addition of the GA prediction tool, the actions deviate from the current course and to understand how the ATCOs envisage their use of the information generated by the tool, for example as a function of the GA probability associated with the prediction.

- *SafeOPS Display Design Evaluation*: Laying down the preliminary display requirement associated with the predictive technology and subsequently evaluating and revising the display design to identify the design choices that help improve the usability and clarity of the predictions for the ATCOs.

The activities of the workshops were divided into three main exercises. Firstly, the ATCOs were guided through a HTA of their actions in the specific conditions of the scenarios described in Deliverable D2.1 [10]. Subsequently, because the predictive tool was not at a maturity level that made possible its adoption in a real-time simulation, the visualisation tool described in Sect. 3.1 was used to present two animations based on real traffic data corresponding to the two "Scenario 1" of each airport. These animations were used to convey the "look and feel" of the predictive tool and the way information could realistically be showing on the radar screen currently in use. In this experimental setting, the second exercises consisted in repeating the task analysis assuming the tool were in use and displaying GA predictions. The third exercise consisted in a semi-structured interview to evaluate the display of the new technology, in the context of the systems display currently used. The results of these exercises are described in some detail in the following.

3.2.2 Results of the Hierarchical Task Analysis

This section summarises the main results of the HTA performed with the ATCOs. We abstract from the details of the reference use cases considered in this analysis (namely Airport 1.UC.01 and Airport 2.UC.01 as described in deliverable D2.1 [10]) to present general results that are relevant to the context of GA operations and independent of specific airports[6].

In the current setup, ATCOs perform the following flow of actions:

- *Initial situation monitoring.* It consists in continuously performing visual checks on the RWY and on the ground radar to ensure it is not occupied, monitoring the EFSS screen to ensure that the departing aircraft (hereinafter, DEP) is marked as "ready", eventually communicating with the FC to ensure they can act rapidly, monitoring the distance and speed of the approaching aircraft (hereinafter, APP) to mentally establish whether there is enough time to clear the departure while maintaining the separation.

- *Communication and coordination.* It consists in the continuous radar monitoring and visual checks of the situation, in particular the status of the RWY, the correct execution of the line-

[6] The results of the detailed HTA performed with the ATCOs can be made available and attached to a revised version of the present document upon request, pending the official approval of information disclosure from DFS.

up and take-off operations, the distance and speed of the APP, the wind speed, and the wake category of the two aircraft to ensure that there is no risk of wake vortex encounter.

- *Monitoring of the evolving situation.* It consists in the continuous radar and visual monitoring of the situation, in particular the status of the RWY, the correct execution of the line-up and take-off operations, the distance and speed of the APP, the wind speed, and the wake category of the two aircraft to ensure that there is no risk of wake vortex encounter.

- *Managing and monitoring the landing phase.* This includes monitoring the execution of the approach operation on the radar, the communication of relevant information (wind speed, weather conditions, etc.) to the flight crew of the APP, issuing the landing clearance and update the status on the EFSS, continuously monitoring that the take-off operations of DEP are being executed as planned.

- *Managing the pilot-induced GA.* The ATCO is informed by the APP that the approach has been discontinued. Because the missed-approach procedure is a high-workload phase for the flight crew, this information might reach the ATCO several tens of seconds after the GA started. Although the ATCO might have noticed indications of the onset of a GA (e.g. the change in speed and descent rate from the radar screen or the change in bank angle if visual contact is established), the response actions effectively start after the MAP call from the APP has been acknowledged. The ATCO checks the distance, speed, intended trajectory and wake category of the surrounding traffic, particularly the DEP and any VFR flights in the area. During this procedure, the ATCO will guarantee separation for example by instructing the pilot to perform a non-standard MAP or, alternatively, by telling the DEP to maintain the ascent below a certain altitude or to climb straight ahead thus delaying a planned turn towards a direction that intersects the trajectory of the published MAP. In addition, the ATCO has to coordinate actions with the other controllers and adjacent sectors, to update the flight status information in the EFSS, to monitor that the GA is executed as instructed, and to transfer the aircraft to the radio frequency of the responsible to re-integrate the aircraft into the arrival sequence for another approach.

3.2.3 Expected task composition with the SafeOPS tool

The ATCOs identified the situations in which their current task composition would differ with the SafeOPS tool in use. Because at the time of the workshops the SafeOPS tool was not sufficiently mature to produce GA predictions based on real flight data, the ATCOs were asked to consider two alternative circumstances, one where the tool forecasts a GA probability close to 100%, one where the likelihood of GA is low (50% or less). *As such, the following analysis refers to **a hypothetical situation** in which the SafeOPS tool were deployed and operational, and the foreseeable response of the ATCOs to the evolving events.*

The Main Actions and tasks mentioned in the following refer to the HTA performed with the ATCOs[7].

Main Action 2 – "ATCO clears departure for TO and marks it in the EFSS". In present-day operations, in case of existing indications that a missed approach procedure was initiated, ATCOs must ensure as

[7] Cf. footnote 6 at page 21.

soon as possible the minimum radar separation of 3 NM. The ATCOs are aware of the fact that a GA can always happen at any point from the final approach fix until the wheels are on the runway and any deceleration device gets activated [11]. However, in a typical traffic situation (as e.g. Scenario 1.1 and 2.1 as described in D2.1 [10]) and without indications of an imminent GA, they can safely clear the DEP for take-off. In an event of a GA, the 3 NM minimum separation must be re-established as soon as possible (e.g. by turning the APP away from standard missed approach or the DEP from its planned trajectory). The ATCO may apply visual separation during good weather conditions in the area around the airport until radar separation has been re-established.

With the SafeOPS tool in use, the situation would be different. In this case, the indication that a GA is imminent would be available and thus the ATCO receives a trigger about this possibility. As a consequence, unless the prediction is displayed long in advance and the approaching aircraft is far from the airport (e.g. at 8-10 NM), the ATCO could decide to restrain from issuing the departure clearance to avoid the risk of inducing a possible separation infringement in case the GA prediction is correct and the missed approach indeed occurs. If the clearance was already issued, the ATCO would have the option to cancel it and, depending on the rolling speed, the take-off can be aborted. The situation would not differ for a low GA likelihood predicted by the tool. The ATCO must privilege safety over capacity and wait until the GA actually occurs or the aircraft lands.

The task flow of Main Action 2 would probably remain unaltered in case the departing and approaching aircraft belong to the same wake-turbulence category and with good weather and visibility conditions. In this situation, there would not be an immediate threat of wake turbulence and visual separation rules apply. Consequently, and more so in the case of a low chance of GA, the ATCO could consider issuing the take-off clearance to the DEP and closely monitoring the situation to rapidly react in case the missed approach effectively takes place.

Main Actions 3 – "Departure A/C overflown runway end or reduced RWY separation exists" – and 4 – "ATCO clears Arriving A/C to land". If the ATCO would decide to delay or cancel the take-off clearance in response to a GA prediction displayed by the SafeOPS tool, Main Action 3 ("Departure A/C overflown runway end or reduced RWY separation exists") does not take place. By contrast, Main Action 4 ("ATCO clears Arriving A/C to land") remains unaltered with or without the SafeOPS tool in operation. Indeed, if the APP has not initiated a missed-approach procedure, then the flight crew is expecting to receive the landing clearance. If the GA prediction is correct and the flight crew is initiating a missed-approach procedure, then they will either inform the ATCO directly, or will not respond because they are currently under pressure in a high-workload situation. The absence of a read-back would be an indirect confirmation of the forecast.

Main Actions 7a – "PL vectors FC Arrival on non-standard MAP" – and 7b – "ATCO vectors FC DEP on not briefed departure route (climb straight ahead)". The current management of the MAP phase would be modified with the SafeOPS tool in operation. In the absence of a departure ahead, as described above, the approaching aircraft can follow the published missed-approach procedure without the need of vectoring on a non-standard GA trajectory. Conversely, if the DEP is airborne, the options to guarantee or re-establish the minimum separation are the same as in the present-day operations. For example, the ATCO could instruct the DEP to limit the climb below 4,000 feet, which is below the standard missed approach altitude of 5,000 feet, or to vector it on a non-briefed departure route (cf. Main Action 7b). Alternatively, the ATCO could decide to turn the arriving aircraft on a non-standard MAP trajectory (cf. Main Action 7a), although this option is generally less favoured to avoid putting pressure on the flight crew which is already dealing with a high-workload situation.

3.2.4 SafeOPS Display Design Evaluation

3.2.4.1 Eliciting User Feedback

The aim of the design evaluation was to identify design issues that had the potential to result in safety and usability problems, and thus inform the design to mitigate these. Common practice HF, when assessing a novel design, is to check the design against a series of design aspirations (requirements associated with effectiveness and safety) but also to focus on 'design heuristics' during the assessment. A heuristic is a mental 'shortcut' that eases the cognitive load of solving problems or making decisions. Heuristics are usually evaluated and enhanced in the design, with the aim of giving the user a perception of optimal usability. An example of commonly evaluated heuristics is as follows (a complete list is found in Appendix A):

> *Visibility of system status - The design should always keep users informed about what is going on.*
>
> *Match between system and the real world - The mental model users have of the system should match the way the system is laid out and operates.*

Once the relevant heuristics to SafeOPS had been established, heuristic 'probes' were written. These probes were questions that helped evaluate the heuristic in a discussion with ATCO and to ascertain if the heuristic was being supported by the design or not.

For example, for the first heuristic 'Visibility of system status', we could determine how well the design meets this heuristic by asking the following probing questions:

Is the design telling me about what is happening? Do I understand why it is telling me this? Do I understand what the current system status is? Does the system notify me of a change in status or a rate of change? Has the design given me appropriate feedback on my actions or any automated action? Was this feedback within a reasonable amount of time? Are all alerts immediately visible? Where on the screen am I notified of the system status? Does the design tell me the accuracy of the AI predictions or how conditions may be changing?

Table 4 is an illustrative example of the questions presented to the ATCOs about the design. A full list can be found in Appendix A, Table 6. Alongside the 'probe' question were relevant considerations for the SafeOPS tool and also the HF requirements associated with that aspect of the design.

Table 4. Example of the questions asked at the workshops with ATCO to investigate design heuristics.

	Display Design Question	Notes for WS facilitator
1	**Does the system help you to understand accurately the current aircraft status?**	Is all 'critical information' about the GA displayed on the radar screen? Should information be displayed on other systems?

For each heuristic there are many questions that can be asked about the system, and if the answer is 'no' or negative for the probe-question, then it is evident that the design is not supporting that heuristic. If the design does not support the heuristic, then we would conclude that usability is not optimal and safety is potentially being impacted. In this case we would seek some form of improvement in the design. This improvement can be achieved by referring to Human Factors best practice, design guidance and recommendations/requirements.

In total 9 ATCOs were interviewed in workshops that lasted 4 hours for each participant. The probe questions were presented in a semi-structured interview style, in which predetermined questions were asked of all participants whilst other questions and related discussion arose spontaneously in a free-flowing conversation. According to this the feedback was largely qualitative e.g. no rating scales or formal questionnaires were used. In this way the interview was exploratory and allowed the evaluation of many aspects of the design integration. A summary of the answers collected from the interviewees during the workshops is reported in Table 6 of Appendix A below each question. It should be noted that because the interviews took place in the form of an open discussion, not all ATCOs have been asked exactly the same questions and in the same order. In some cases the interviews evolved in such a way that made some of the questions to become irrelevant for the conversation Also, sometimes the respondents gave a personal answer which refer to them individually, and subsequently expanded it to claify why other colleagues which have a different “managing style” of air traffic might behave or think differently and motivating the differences. While we tried to capture the complexity of these answers in the following section and in Table 6, it should be noted that, for all these reasons, in some questions the percentage of the respondents do not add up to 100%.

3.2.4.2 Feedback Results

Following is a summary of the feedback, arranged in themes, elicited at the user workshops. The feedback directs the design on discrete design features but also introduces functional/operational aspects of the design that should be further explored when a more mature prototype is available.

As a basic requirement, the ATCOs want to be alerted to the presence of a GA, but do not want additional information. This is because too much information is distracting and the ATCOs do not have the spare capacity to process additional information, especially when the alert is close to the point that the GA will occur. They want to be told of the event happening and to simply acknowledge and accept the notification. Moreover, they do not want to monitor the evolving probability of the GA, as they will look at the display to check the aircraft parameters, including the GA status, only for a few seconds and will not attempt to remember any detailed information (e.g. changing probability percentage) of the aircraft after discretely viewing it.

Discussion on how the symbology notifies the ATCO of a change in status from predicting the GA to detection of an actual GA, led to the conclusion that the status of the aircraft, in these two differing states was not being represented in the design.

There was some suggestion that not every ATCO would recognise the GA symbology, the wording and the colour scheme, nor attend to it appropriately. In contrast, some ATCO’s have stated that the change in colour of the arriving GA aircraft might be too distracting, compared with the other arriving aircraft. They reported that they would like the right colour to attract their attention and information about the contents of the alert e.g. GA, but they want the alert to then remain as it is and not change its appearance according to whether the alert has remained above a ‘threshold’ level (of probability) or not.

In consideration of the system giving feedback to the ATCOs on any of their actions or any automated action undertaken by the system, the ATCOs suggested that they did not want any additional information related to these aspects as they would find it too distracting.

Concerning the timeliness of the prediction information, ATCOs claim that they would most likely check the GA status on first contact with the aircraft, which could be anywhere from 8NM to 4NM depending on the runway. But there was a general consensus that the sooner the GA is notified to them, the more useful the information would be. Moreover, they would like the information on the display until the point that the aircraft lands (if not a missed approach) or until the GA occurs.

In relation to understanding the meaning of the number indicating the probability, the ATCOs suggested that as this number is situated next to call-sign, it might be misleading and be mistaken for the call-sign. Moreover, the number, as a percentage, was hard for them to understand and or have a comprehension of where this number is derived from. Not necessarily through the display design, but in general the ATCOs expressed a desire to understand more about where this number comes from. Some ATCOs went as far as saying that possibly the probabilistic information is redundant and unnecessary, depending on how this information is communicated, for example if the alert is clear as a non-numerical indication which, when the GA threshold is reached, changes colour, then no numeric indication is required.

The ATCOs suggested that any alpha-numeric information might be too distracting, therefore they would rather that the display is kept simple and no extra information is displayed. However, some expressed a desire for potentially displaying the cause of the GA, not necessarily on the radar screen though, but they postulated that the use of this additional information would differ between ATCOs. They stated that as soon as there was a GA, there was no need for any other information but before the GA happens, some additional information might be useful.

Some ATCOs expressed an interest in the option to switch off the GA alert system, if the system did not show high quality of prediction.

ATCOs suggested that this display should be harmonised with the current STCA alerts, in order to ensure that this display interface behaved like existing interfaces. Moreover, there was a suggestion that a 'red box' is placed around all the aircraft information once the GA is actually occurring, and that an orange colour code is used in the case that a GA is likely, this is consistent with the design of the STCA. Related to this, ATCOs suggested that if/when the GA is actually happening, the symbology should become red, as this is in alignment with the current design on the EFSS e.g. when a GA occurs, the input to the EFSS is displayed as a red frame.

Opinions differed on the point of being able to configure the system according to ATCO wants and needs. In terms of making an action or input through the display design, the ATCOs stated that they felt it was somewhat unnecessary and that any action could and should be done through already existing systems, including that of notifying other controllers of the GA. Moreover, it would require a large change to the current systems to enable this feature.

In relation to an incorrectly predicted GA, or even when the likelihood is less than 100%, the ATCOs suggested that their actions would depend on how accurate the system actually is, if it was an accurate 80% probability or an inaccurate 100% probability of GA. As they would look at the prediction for decisions made about the planning of outbound flights. However, some ATCOs state that as the current system gives no indication of a GA, and GA's are always a 'surprise', then a situation in which the aircraft actually lands when it has been predicted to GA, is "not a real problem". But they emphasised that they would prefer it if, when the system does recognize that something has changed, it would update the display instantly.

With regards to additional help, training or documentation for the new system, the ATCOs proposed that a useful document would be one in which different 'states' of the aircraft were defined according to how they are displayed on the screen e.g. meaning of different colours, symbology limits of the system etc.

4 Discussion and Recommendations

4.1 Guidelines on how to present the information

4.1.1 Human Factors Guidance for SafeOPS

According to Human Factors knowledge and research, there are guidelines on how best to present information to the user. These guidelines generate requirements for the SafeOPS system, which are discussed below and interpreted in terms of the experience and feedback from the users in the SafeOPS workshops. Indeed, the following requirements have been selected to increase usability and transparency, and to minimise the impact from the predictive tool when it is being less reliable/accurate in its predictions. These requirements should thus optimise the HMI to engender trust from the user.

The following are the requirements discussed in terms of the impact on the SafeOPS tool (with the referenced requirement ID), however the formatted list of requirements can be found in Appendix B.

i. *The design should always keep users informed about what is going on, through appropriate feedback within a reasonable amount of time. (REQ-SafeOPS 22-TS-HFSR-002-1)*

   It is noticeably apparent in the discussions with the ATCOs that they desire a display and tool which is very precise and accurate and that they require the GA alert to be present from as far out as possible. Therefore, in adherence with this HF requirement, it is recommended that a GA is only displayed when it is highly likely and this information is highly accurate. So there is no ambiguity in the display of GA information and there is no need for additional cognitive processing by the ATCO, especially in critical moments of the approach e.g. close to landing. Any information on a less likely GA or of information of questionable accuracy should not be displayed as it is not presenting a true and certain reflection of what is going on with the aircraft. A GA alert out at 8NM is desirable for the ATCO but this should be balanced against the accuracy of the information being provided and thus sacrifices should be made in the 'timeliness' component of the alert, with preference to accuracy.

ii. *All information that a controller needs to accomplish a task that is essential and time critical is located on a single page or in a single window (REQ-SafeOPS 22-TS-HFSR-002-2).*

   The controllers have very little additional capacity to process information and thus need something that gives them all the information they require in a single unit. It has become apparent through the workshops, that as long as the prediction is reliable, then simply being told that the GA will or will not occur is adequate and preferable. To adhere to this HF requirement, it is therefore recommended that the presentation of the GA alert on the radar screen should be complete in terms of its ability to notify the ATCO of a GA. Any additional information on other displays is thus additional information and not required for the proper notification of the GA to the ATCO. Somewhat contradicting the opinion of the majority of the interviewees, some ATCOs suggested that the display of the probability of GA occurrence (cf. also the discussion at point *x.* below) can be considered worthy not so much for its absolute value (which has to be high or the prediction is not useful) but as a way to track its evolution: "as time passes, is the GA becoming more or less likely?"

iii. *Critical information should never disappear from the screen without being deleted or suppressed by the controller (REQ-SafeOPS 22-TS-HFSR-002-3).*

In the approach, the display of the GA status should not switch between being present and then not being present without there being some notification to the ATCO. Numerous changes in status of the aircraft on the approach path is distracting, whether it's a change in the colour, symbology or presence/absence of the alert, especially because it's not linear data and gives no indication of the rate of change of the status (in contrast, change in speed or altitude of an approaching aircraft gives a rate of change and a direction of change).

iv. *Updated data (when something changes) are emphasised effectively so that it attracts the controller's attention (REQ-SafeOPS 22-TS-HFSR-002-4).*

It is apparent from the workshops that the ATCOs do not want a lot of information on changing parameters of the GA alert, only that it is likely or not to occur. The display needs to display this precisely and thus there should be a mechanism to show the user when an aircraft goes from no-GA likelihood status to very likely, with an attention getting characteristic, but one that does not draw the attention of the ATCO away for more important information on the display, e.g. the alert should sit appropriately in a hierarchy of ATC notifications which generally range from advisory, to caution, to warning.

v. *Essential ATC information is never blocked or obstructed by other information. Display clutter is not a problem (REQ-SafeOPS 22-TS-HFSR-002-5).*

These requirements are difficult to test when the system is very immature and representative screens/real-life scenarios are unavailable for evaluation with the new display symbology. However, it is recommended that, due to the fact that the GA alert is advisory and a decision making aid, the GA alert information display does not take priority over the current display of aircraft information on the radar screen. Therefore, a mechanism is required to occlude the GA information on the display when more important information is required in the same location.

vi. *When the meaning of the colour is critical, colour is used redundantly with another type of visual cue, such as shape, text, or size. For example, all yellow objects have a triangular shape (REQ-SafeOPS 22-TS-HFSR-002-6).*

Although some of the ATCOs suggested that any alpha-numerical data was not necessarily needed for the GA alert, only colour, it is recommended that this HF requirement is adhered to and a secondary 'visual cue' is included in the display design, to notify the user of the GA. This is because colour is critical to the status of the GA when it has been confirmed as highly likely. Moreover, the application of colour on the broader ATC display has significant meaning and is used to define most display elements, thus it's essential for the ATCO to interpret all colours correctly, relative to each other. To improve the accurate identification of a colour's meaning, it is thus useful to have a secondary visual clue, e.g. the letters 'GA'. It also increases the attention getting characteristics of the alert.

Notwithstanding this, if the letters 'GA' are placed in the aircraft information block, then it is essential to ensure that this information isn't misinterpreted as something else, e.g. part of the call-sign.

vii. *Colours are far enough apart in perceptual terms (e.g. distinctly different colours) that they are not able to be confused even when "washed out" by sunlight, if applicable. When the controller must distinguish between the colour of characters and symbols, small blue characters and symbols are not used as they have low luminance and visual resolution of fine detail is poor for blue. Saturated (i.e., vivid) red and blue are not presented next to each other (REQ-SafeOPS 22-TS-HFSR-002-7).*

The controller will not need to identify more than six colours (to interpret the meaning of the colour when it stands alone). E.g. no more than six colours should be assigned a specific meaning (REQ-SafeOPS 22-TS-HFSR-002-8).

The standard convention across aviation for red/yellow/green should be adhered to. Where the colour red is used only for warning/danger, yellow is used to indicate caution, green is used to indicate normal/ready status (REQ-SafeOPS 22-TS-HFSR-002-9).

It is critical that the correct colour is chosen for the GA alert according to the colour convention of the existing display system. Also, the colour should be distinctly different from the other colours, in all environments e.g. ambient light. If for example the wrong orange is selected, it can appear as a red in some lighting conditions.

viii. *High-priority alerts and other critical information are located within the central display area (i.e., the central 15 degrees of the area where the controller normally looks, given the normal viewing position) (REQ-SafeOPS 22-TS-HFSR-002-10).*

Spontaneous alerts that are positioned outside this central display area of 15 degrees are significantly less likely to be noticed, also depending on the workload of the ATCO at that particular point in time and the level of symbology being displayed on the entire radar display. This is sometimes mitigated when a user is trained to regularly scan a display or outside of the window view. However it is postulated that the GA alert is not high-priority, or critical information, in the same way that a potential aircraft conflict is. Therefore, it is not appropriate to draw the attention of the ATCO to the potential GA in the central display area, if the aircraft is at that time outside of this area. Although the GA symbology can be displayed outside of the central display area of 15 degrees, but the ATCO may only become aware of a likely GA, when he directly looks at that particular aircraft (or when the aircraft is within the central display area of the screen).

ix. *Interfaces should not contain information which is irrelevant or rarely needed. Every extra unit of information in an interface competes with the relevant units of information and diminishes their relative visibility (REQ-SafeOPS 22-TS-HFSR-002-11).*

It was emphasised a lot by the ATCOs that they do not want additional information on the display screen beyond that which is essential. Moreover, the display of a medium to low probability of a GA should not be displayed as its not essential information for the ATCO and takes up precious display ‘real-estate’ and attention away from the ATCO.

x. *The meaning of each icon is immediately apparent to the controller or it is labelled (REQ-SafeOPS 22-TS-HFSR-002-12).*

This is particularly pertinent to the meaning of a percentage-figure representing the probability of GA. Throughout the workshops it has been apparent that a percentage-figure does not give appropriate information to the ATCO on what it means. In fact, humans do not have an intuitive grasp of risk and are not good at interpreting probabilities accurately. There are cognitive biases and distortions associated with human interpretation of probabilities, for example two commonly identified distortions are *conservatism*, in which individuals interpret probabilities as being less extreme than the true values (tending toward 50%), and *repulsive* biases away from 50%, where extreme values are disproportionately interpreted e.g. a medium to high probability is interpreted as a high probability. Alternatively, the user could simply neglect the probability value, rather than incorrectly use the probability.

So, the presentation of a probability of likelihood, in a numerical form, is deemed to be too ambiguous and troublesome for the user, and it demands too much interpretation on the behalf of the user. This lends itself to using valuable attention of the user, causing a distraction and encouraging errors to be made in the decision making of the ATCO. As an alternative, a high likelihood of the GA should be represented by the colour coding and lettering or symbols representing the GA on the display.

xi. *Acronyms in the new display system have the same meanings as in the previous system. Users should not have to wonder whether different words, situations, or actions mean the same thing. The design should make use of the language that is clear to the intended user, e.g. use words, phrases, and concepts familiar to the user. Placement of standard data fields, and Formats used within data fields, is consistent from one display to another. Labels, terms, and abbreviations are used consistently across the display set (REQ-SafeOPS 22-TS-HFSR-002-13).*

Consistency of the new display symbology with the current system and current conventions is critical. Thus, to ensure adherence with this requirement, any novel design should be compared with the current system, for all aspects of the display. The new display design should not have a colour, symbology or other characteristics that have a meaning that is in conflict with the current definitions/conventions for these elements.

xii. *The 'mental model' users have of the system should match the way the system is laid out and operates, e.g. actions and options provided by the system should match actions performed by the user (REQ-SafeOPS 22-TS-HFSR-002-14).*

The display should, at the very least, match or represent the current procedures for the GA, e.g. the display should present to the user a confirmation of the GA, when a GA has been confirmed by the pilot.

xiii. *Alerts should have a low incidence of false alarms (REQ-SafeOPS 22-TS-HFSR-002-15).*

This is a critical requirement for the design. Acceptance of this kind of technology is almost entirely dependent on the accuracy and reliability of the information presented. Throughout the workshops, users have often expressed their needs and desires for the system, in the context of its accuracy in predicting a GA. They have stated that their willingness to use the tool would be completely dependent on its ability to accurately predict a GA. It is unlikely that trust in the system will occur if there is wavering reliability and many false alarms. Therefore, the effectiveness of the display design is wholly reliant on the trustworthiness of the tool.

A false prediction of GA is generally perceived as a low-risk situation. The users have suggested that if an aircraft actually lands when it has been predicted to GA, "it is not a real problem". Also in the reverse, where the aircraft does a GA instead of landing as predicted by the tool, would represent the current situation in which the GA is a last minute 'surprise'. Both situations, although they do not present an overly hazardous outcome compared to the current system, would render the tool unreliable in the eyes of the user and subsequently their trust in the tool would be degraded and use of the tool would no doubt decline.

xiv. *Its best if the system doesn't need any additional explanation. However, it may be necessary to provide help and documentation to help users understand how to complete their tasks (REQ-SafeOPS 22-TS-HFSR-002-16).*

The ATCOs expressed a desire for a design conventions document, in order to understand for example the states represented by the various colours. This would be useful documentation for the user in training, however this is relevant to the entire system. Therefore it is suggested that for the purpose of SafeOPS, the help and documentation needed is in the form of instructions/training in the new symbology.

4.1.2 General lessons that can be learnt about HF with AI-based tools.

The Human Factors guidance presented so far is made up of what initially appears to be quite discrete design requirements for individual aspects of the display design. However, these requirements are in fact written in a collective way in order to 'pull' the design into a place where transparency is paramount and where error is minimised. The motivation for this comes from the fact that, as previously mentioned, the three components of an effective HMI, in relation to engendering trust are transparency, meaningful error reporting and minimal system failures (the latter being more part of the technology design than the display design). Moreover, trust and reliance are key topics in the field of Human Factors related to AI. The following introduces these concepts, as they are pertinent to AI and machine learning, but it is recommended that ultimately these elements should be recalled and considered in the final design and evaluation of the SafeOPS concept.

Trustworthiness: A model is seen to be more trustworthy when the observable decision process of the model, matches the user's existing understanding on what this process should be. This means that the act of explaining what the system is doing (by the system) does not in itself enable trust. Trust is built not only when the user successfully comprehends the true reasoning behind the decision process of the model, but also that this reasoning matches the user's existing understanding of an adequate and agreeable level of reasoning. On this basis if the user has no existing understanding of what is an adequate decision process, then trust will not be encouraged, even if the AI is easy to understand and transparent. To mitigate this, proper education of the user, on what should be expected by the observable decision process of the model, should be accomplished but also the design should incorporate 'attributes' that uphold this understanding in the user. This is described as '*trustworthiness through explanation*' and comes from observing the inner workings of the model, by the user [12].

The other mechanism of engendering trust in a model is through the display of trustworthy behaviours, by the model. In this situation the trust is built more in the 'evaluation' abilities of the model, than in the decision-making process of the model. To encourage an 'evaluation scheme' to be trustworthy requires the demonstration that the identification of 'instances' (e.g. missed approach) during the model evaluation, matches the distribution of the true unseen 'instances'. This kind of trust can be encouraged in the user (for the system) through the contribution of expert (human) opinion on AI reasoning or behaviour, which can enable non-experts to gain trust in the AI. In addition to the review by the users, of post-deployment data, for them to witness the behaviour of the evaluation abilities of the AI.

Fundamentally though, the HMI should aim for the ideal balance between trust and trustworthiness, where trust is an inherent attribute of the user and requires a certain level of vulnerability and acceptance of risk, and trustworthiness is an ability of the technological system to deliver trustworthy characteristics or 'products' to the user.

Differences in decision making based on AI inputs: When the tool is used as a decision support tool, even if the predictions are accurate and consistent across all scenarios, this doesn't translate into consistent decision making across users. Individuals will make differing decisions, based on the same AI Input [13]. Indeed, humans have biases when making decisions, which is essentially the key justification for automation in AI, to mitigate these biases. However in the case that a high level of automation is not being used, and instead machine learning probabilistic information is being used by the ATCO as a decision making tool, then these biases become once again a factor in decision making. Therefore, it is important to be aware of the individual tendencies (or biases) of users when interacting with AI.

Over-reliance on AI: Whereas 'trustworthiness' is often reported as a challenge to AI user acceptance, the opposite can become a potential challenge, namely that users put too much or total trust in the AI solution and 'shifts the risk' associated with decision making, transferring it to the AI. In this situation, the user makes decisions on the AI information as if it were deterministic and not probabilistic. Therefore, although the reliability of the M/L prediction needs to be high, there still needs to be an understanding in the user and regulatory community that AI is most likely less accurate than traditional systems. Thus, the risk associated with any human decision based on the M/L output, must be fully articulated, explored, understood, and accepted by the user, regardless of the AI advisory.

Lack of predictability and explainability of the ML application behaviour: The probabilistic nature of an M/L model can lead to unpredictable outputs that may be difficult to explain. Therefore, in some cases, more insight and understanding of the basis of the prediction, is not available, useful and or able to provide adequate explainability [2]. This is applicable to some M/L models but this may not be a consideration for the SafeOPS tools as the output is binary, i.e. the aircraft is landing or conducting a GA. However it is important to stay mindful of point, in case in becomes relevant in the validation phase of the tool.

5 Conclusion and Outlook

Research on the Human Factors (HF) related to Artificial Intelligence (AI) is a relatively new field which is developing quickly to hasten the integration of AI tools into the operational environment. The integration into aviation, of automated systems which use deterministic data, has received much analysis and evaluation to date, and has provided us with a basis for understanding the impact on the human, of not being the only 'actor' in the flying of an aircraft or controlling of an airspace. However, the introduction of probabilistic information has added an additional dimension to the relationship between man and machine, which is producing the need for significant exploration for the purpose of all phases of technology integration, including design, validation, certification, operations and regulation.

Accordingly, this current review aimed to introduce and evaluate the HF related to the design and integration of the SafeOPS tool, a machine learning tool which presents probabilistic information. To achieve the aim, this review delivered an overview of HF as it relates to probabilistic information, reviewed existing ATM 'safety' tools, presented a format for evaluating the HF of an early prototype of the SafeOPS tool, reported on user feedback elicited through several online workshops, and finally provided HF design requirements and guidance for the tool, according to user feedback, current HF best practice and up to date understanding of HF in AI. The main recommendations drawn from this analysis are: (*i*) High accuracy, i.e., it is preferable to only display predictions of highly likely Go Arounds, at the potential cost of a relatively high false-negative rate; (*ii*) timeliness, i.e., the users benefit the most the longer is the look-ahead time; (*iii*) the amount of information that is displayed should be minimised to avoid clutter or overload of information; (*iv*) colours and markings should be aligned with those currently in use in existing systems.

Based on this work it is therefore recommended that the proposed requirements, alongside the guidance notes, for the presentation of the Go Around information at the Human Computer Interface, should be used to direct the design. This should be through an *iterative* design and validation process with the user as an integral part of the cycle. As the design matures and the tool can be represented on the display with functional software at the rear, then deeper aspects of trust, human error, acceptance and response to false alerts should be assessed and analysed, in an increasingly more representative environment.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

6 References

[1] T. Kistan, A. Gardi, and R. Sabatini, "Machine Learning and Cognitive Ergonomics in Air Traffic Management: Recent Developments and Considerations for Certification," *Aerosp. 2018, Vol. 5, Page 103*, vol. 5, no. 4, p. 103, Oct. 2018, doi: 10.3390/AEROSPACE5040103.

[2] EASA, "EASA Artificial Intelligence Roadmap 1.0 A human-centric approach to AI in aviation Photocredits istock, Agustawestand," 2020.

[3] Eurocontrol, "Ground-based safety nets (STCA, MSAW, APW, APM)." https://www.eurocontrol.int/project/ground-based-safety-nets (accessed Mar. 30, 2022).

[4] J. Saint-Lot, J. P. Imbert, and F. Dehais, "Red Alert: A Cognitive Countermeasure to Mitigate Attentional Tunneling," *Conf. Hum. Factors Comput. Syst. - Proc.*, Apr. 2020, doi: 10.1145/3313831.3376709.

[5] SKYbrary Aviation Safety, "EUROCONTROL STCA Specification and Guidance Material." https://skybrary.aero/articles/eurocontrol-stca-specification-and-guidance-material (accessed Mar. 30, 2022).

[6] L. Giraudet, J. P. Imbert, M. Bérenger, S. Tremblay, and M. Causse, "The neuroergonomic evaluation of human machine interface design in air traffic control using behavioral and EEG/ERP measures," *Behav. Brain Res.*, vol. 294, pp. 246–253, Nov. 2015, doi: 10.1016/J.BBR.2015.07.041.

[7] B. Hilburn, "Evaluating Human Interaction with Advanced Air Traffic Management Automation," *RTO Meet. Proc.*, no. MP-088, pp. 1–12, 2002.

[8] B. Bakker, "EUROCONTROL Guidance Material for Area Proximity Warning EUROCONTROL Guidelines," Accessed: Mar. 31, 2022. [Online]. Available: www.eurocontrol.int.

[9] WIKIFACTA, "Investigate Minimum Safe Altitude Warning Systems (MSAW) |." http://wiki.ifatca.org/kb/wp-1999-88/ (accessed Mar. 31, 2022).

[10] SafeOPS Consortium, "D2.1 - Development of Use Cases , User Stories and Requirements," 2021. [Online]. Available: https://innaxis-comm.s3.eu-central-1.amazonaws.com/SafeOPS/SafeOPS_D2.1_v00.02.00.pdf.

[11] SKYbrary Aviation Safety, "Go Around." https://skybrary.aero/articles/go-around (accessed Mar. 30, 2022).

[12] A. Jacovi, A. Marasović, T. Miller, and Y. Goldberg, "Formalizing Trust in Artificial Intelligence: Prerequisites, Causes and Goals of Human Trust in AI ACM Reference Format," doi: 10.1145/3442188.3445923.

[13] P. Meissner and C. Keding, "The Human Factor in AI-Based Decision-Making," *MIT Sloan Management Review*, 2021. https://sloanreview.mit.edu/article/the-human-factor-in-ai-based-decision-making/ (accessed Mar. 31, 2022).

Appendix A Design Heuristics

A list of design heuristics commonly adopted in user-centred design to enhance users' usability is presented in Table 5. The complete list of questions that were asked during the workshops with ATCOs to identify the optimal display and timing of the GA information within the SafeOPS tool is shown in Table 6.

Table 5. Commonly-used design heuristics.

Design Heuristic	Description
Visibility of system status	The design should always keep users informed about what is going on.
Match between system and the real world	The mental model users have of the system should match the way the system is laid out and operates.
Users' language	The design should make use of the language that is clear to the intended user.
User control and freedom	Users should feel they are in control of the system.
Reversible actions	The system should allow the users to undo actions to recover from errors, and to learn from these errors.
Consistency and standards	Users should not have to wonder whether different words, situations, or actions mean the same thing.
Error prevention	The design and layout of the system should prevent users from making errors whenever possible.
Help users recognize, diagnose, and recover from errors	Error messages should be expressed in plain language, precisely indicate the problem, and constructively suggest a solution.
Recognition rather than recall	Minimize the user's memory load by making elements, actions, and options visible.
Flexibility and efficiency of use	The system should support a variety of users through customization and shortcuts.
Minimalist design	Interfaces should not contain information which is irrelevant or rarely needed.
Help and documentation	Its best if the system doesn't need any additional explanation.

Co-funded by the European Union

Table 6 (following page). List of questions for the semi-structured interviews of the ATCOs to identify the best design properties to deliver the GA information within the SafeOPS tool. A summary of the answers collected from the interviewees during the workshops is reported below each question, marked with the letter 'A'.

	Display Design Question	notes for the WS facilitator
1	**Does the system help you to understand accurately the current aircraft status?**	*Is all 'critical information' about the GA displayed on the radar screen?* *Should information be displayed on other systems?*
A	9/9 ATCOs think the system helps understand the aircraft status and increased likelihood of GA. 2/9 ATCOs think that the only important information that is currently missing is an indication of the reason of the prediction, that is the cause of the predicted GA. However, ATCOs opinions diverge on this subject. Other 4/9 interviewees explicitly said that processing the additional information about the cause of the warning would be unnecessarily time consuming in a phase where time is critical, hence they would prefer to only have the situational awareness about the potential GA.	
2	**Does the display clearly notify you of a change in aircraft status?**	*At what point does the magenta turn back to yellow of the GA aircraft and should it be ATCO activated or system activate?* *When the situation changes, in the GA situation, is the ATCO informed through the design, in an attention getting way?*
A	9/9 ATCOs consider satisfactory the change in colour of the callsign and the aircraft diamond to indicate the likely GA. The opinions differ on the colour scheme to be used. 7/9 ATCOs think that a colour change from yellow (normal approach) to orange or magenta is acceptable and sufficiently clear. 2/9 ATCOs think that magenta might potentially be too distracting. 2/9 ATCOs consider red a possible option but this is in general discouraged by other ATCOs as red is the colour usually associated to an actual emergency or urgency and a GA prediction is not considered an emergency. Red might be considered to indicate actual GA detections.	
3	**Is the display of aircraft status immediately visible to you?**	*Is there any overlay of critical information by less critical information on the display?* *Are symbols on the radar screen easily discriminated from one another and symbology does not overlay each other?* *Is there something else on the display that notifies a missed approach apart for the colour changing of the aircraft marker, so there is some redundancy?* *Does the controller use a white or a black background?*
A	All 9 ATCOs said that the aircraft status is immediately visible. 1/9 ATCOs said that there is a possibility that not all ATCOs would recognise the status change in high traffic conditions and if the only change is in the aircraft marker and callsign change from yellow to orange. 1/9 ATCOs said that the colour change from yellow to magenta might be distracting.	
4	**Does the display 'get' your attention?**	*Consider if the display is attention getting enough when the GA is getting close to being a GA or when a GA is very likely.*
A	All 9 ATCOs replied affirmatively. 1 ATCO said that the colour change from yellow to magenta might be too striking and thus distracting.	

<table>
<tr><th></th><th>Display Design Question</th><th>notes for the WS facilitator</th></tr>
<tr><td>5</td><td>Should the display give you appropriate feedback on any of your actions or any automated action?</td><td>Is the system giving any confirmation or level of success of any interactions from the ATCOs?
No automated action should be performed without clear notification to the ATCO.</td></tr>
<tr><td>A</td><td colspan="2">The options differ on this point. 8/9 ATCOs decisively replied that they want to minimise the amount of information displayed on the screen, and therefore no feedback nor recommendations as they would be too distracting. 1 ATCO suggested that possibly additional information could be accessed, if time concedes, by interacting with the radar screen (e.g. by clicking on the callsign). All interviewees agreed that no automated action should be performed, and the ATCO must be in charge of any decision of how to manage the situation.</td></tr>
<tr><td>6</td><td>Are the display notifications timely?</td><td>When the aircraft status is updated, is it seemingly at the same time as real life aircraft status changes?
Does the display add benefit from providing ahead of time information?</td></tr>
<tr><td>A</td><td colspan="2">9/9 ATCOs think that the sooner the notification is displayed the better it is to form a mental plan of the traffic and how to manage it. Depending on the runways and airport layout, the timing to check the radar are different, but generally on initial contact the label will be checked when approaching aircraft is around 8 – 10 NM from the threshold at the earliest (in typical traffic conditions). Ideally, this is the earliest moment when the prediction is useful. At the workshops, the GA notification was displayed at 4NM, which all ATCOs agreed is close to the limit where this information can be exploited to adapt their plans. 3/9 ATCOs explicitly said that displaying the information is always useful for situational awareness. 2/9 ATCOs explicitly suggested to keep displaying information until touchdown or until GA is performed.</td></tr>
<tr><td>7</td><td>Do you understand the meaning of the number associated with the predictive tool and how it is changing?</td><td>Is it display appropriately? How does the ATCO interpret the number? How does this impact on their decisions?</td></tr>
<tr><td>A</td><td colspan="2">The opinions varied significantly and some ATCOs change their views in the course of the discussions. 2 ATCOs understand the meaning of the probability. 1 ATCO appreciate the possibility to monitor how such a probability change in time, because its evolution (increase or decrease) would give a better situational awareness than a fixed value. However, 6/9 ATCOs came to the conclusion that they would find less distracting to not have this information and only have the GA warning above certain threshold (typically identified at 90% or higher). 4/9 ATCOs think that the indication of the GA likelihood next to the callsign may be mistakenly read as part of the callsign itself and suggested to move it to a new line. 2/9 ATCOs said that the percentage is hard to understand and distracting, because a lot of attention would be dragged in trying to interpret what the number means in terms of chances of the GA actually happening.</td></tr>
<tr><td>8</td><td>Would you expect the display to show more contextually relevant information e.g. weather, fuel status, stability of approach?</td><td>Does the ATCO want the reason for the GA if available?</td></tr>
<tr><td>A</td><td colspan="2">While all ATCOs agreed that this information would be useful, 7/9 concluded that this should be shown only on request by interacting with the string to avoid distraction and cluttering. 2/9 replied that they would avoid all additional alphanumeric contextual information. 2/9 ATCOs said that before the GA happens contextual information might be useful to visualise but that as soon as the GA is actually initiated, displaying such information is irrelevant.</td></tr>
</table>

	Display Design Question	notes for the WS facilitator
9	**Does the display contain information which is irrelevant or rarely needed?**	
A	2/9 ATCOs replied that potentially the probabilistic information is redundant. In particular, if it is clear that a non-numeric indication (e.g. the colour change) corresponds to a prediction above a fixed probability threshold, then numeric indication is required.	
10	**Are the words/symbols/ objects on the display familiar to you?**	*Are the symbols and colours chosen for the display intuitive so that the controller can interpret them quickly and accurately?* *Do the 'GA' signify something different currently on the radar screen or ATCO systems?*
A	The adopted symbology is familiar to all ATCOs and that the 'GA' indication is easily understandable as "go around".	
11	**Does the display interface behave like existing interfaces?**	
A	All ATCOs agreed that the display interface behaves in a sufficiently similar way to existing interfaces. 2/9 ATCOs noted that choosing the orange colour for GA predictions would align to the colour coding used in the STCA alert system.	
12	**How would you expect the display to communicate the information on GA's?**	
A	All ATCOs agreed that the change in colour of the aircraft marker and callsign and the display of the letters 'GA' (possibly not on the same line of the callsign to avoid confusion) is sufficiently clear. 2/9 ATCOs pointed out that highlighting the callsign with a box and particularly using the colour red should be avoided or eventually be used only when the GA is taking place.	
13	**Would you like to interact with the display design?**	*Consider what the current mechanism is for interacting with the display.*
A	4/9 ATCOs said that this possibility is unnecessary and distracting, and all interactions with the display should be done through the already-existing systems. However, the interviewees conceded that the possibility to interact with the display might be useful to interrogate for more information in case there is sufficient time.	
14	**Would you like to configure it to meet your needs?**	
A	4/9 ATCOs replied the same as the question 13. 1/9 ATCO said that having to configure additional settings would be considered deprecated because of the additional head-down time necessary for this operation.	
15	**Would you like to interrogate the display for more information?**	*Consider 'Progressive disclosure' is an interaction design pattern that sequences information and actions across several screens*

	Display Design Question	notes for the WS facilitator
A	9/9 interviewees agreed that this option might or might not be well-received depending on the traffic and weather situation. If there is time to react, e.g. because the approaching aircraft is 8 – 10 NM out with limited other traffic, having more information is useful to elaborate a better management plan. In high-workload situations, there would not be the time to get the information, so having the option would be irrelevant. 4/9 ATCOs replied that it also depends on the "management style" of the individual ATCO: some might be more interested in knowing the main causes of the GA predictions, others might be only interested in the dynamic evolution of the situation.	
16	**Would you like to make an action through the display design?**	*Does the ATCO require a mechanism to make an action, related to the GA through the display screen e.g. warn other Air traffickers of the GA?*
A	4/9 interviewees answered that this feature is not desired. 2/9 considered it useful the possibility to indicate through the system the go around to other ATCOs, however it is recognised that such feature would require a substantial change to the current systems.	
17	**Does the display design align with the design conventions used on the display screen currently (fonts, colour, capitalization, position, layout, terminology, sequence of actions)?**	
A	Considering the small differences that exist in the settings chosen by the individual ATCOs and the design conventions used in different airports, 9/9 ATCOs answered that the design options shown in the workshops reasonably align with the current display and colour conventions.	
18	**Does the new design align with the current procedures and training on go-around and missed approach?**	
A	7/9 ATCOs agreed that in general the new design does not seem inconsistent with the current procedures.	
19	**Is the display design behaving consistently with conventions on other platforms, e.g. the EFSS?**	*A colour needs to be chosen for the GA symbology that is consistent with the other displays (and with the ATCOs mental model) in terms of other notifications on GA.*
A	2/9 ATCOs interviewed together during one workshop pointed out that the magenta colour is currently used on the EFSS for initial contact, and thus its use for GA might be confusing. None of the other 7/9 ATCOs has raised this issue during other workshops. While this suggests that possibly this small discrepancy would be easily recognised and correctly interpreted by the ATCO in real-time operations, it also indicates that the colour-change to magenta is not the best choice for displaying the GA prediction.	
20	**How would you expect the display to perform when there is a go-around that is incorrectly predicted e.g. a go-around occurs which was not predicted?What are your actions in these situations?**	*Is there some notification of the tools reliability?*

	Display Design Question	notes for the WS facilitator
A	No clear indications came out from this question. 2/9 ATCOs answered that the display should not behave differently than usual, it will be the ATCO's job to monitor the situation and adapt to the evolving circumstances. In general, what is important is that the system recognises when the situation changes and eventually updates the notifications on the screen accordingly. 2/9 ATCOs pointed out that during training the ATCOs should be instructed on how to handle situations when the GA likelihood is less than 100%. 9/9 ATCOs agreed that the situation in which the tool would fail to predict a GA would not be dissimilar from the current operations where no indication of go around exists and therefore every GA comes as a "surprise". The opposite case in which there is a GA prediction and the aircraft lands anyway would not create a safety problem, but potentially only a capacity problems as the ATCO might have decided to hold a departure to avoid having two aircraft (the predicted GA and the departure itself) close in the air.	
21	**How often will you use the new system display?**	
A	7/9 ATCOs replied that it would depend on how reliable the system is. Should the predictions be accurate, there would be a clear benefit in displaying them. Otherwise, they would be distracting and the ATCOs would prefer to be able to turn off the system.	
22	**If you hadn't seen the interface in some time (e.g. there are no go-arounds occurring), would you recognise its meaning and how to use the information being presented?**	*Consider if the ATCO hasn't seen the GA symbology for a long time/infrequently. Symbols and colours chosen for the display are intuitive so that the controller can interpret them quickly and accurately.* *How many colours are currently used to represent aircraft in different states?*
A	2/9 ATCOs answered that the design is sufficiently intuitive that they would be able to recognise its meaning even if they had not seen it in some time.	
23	**Would any additional help, training or documentation be required to help you understand how to use the display in your role?**	
A	7/9 ATCOs agreed that it would be fundamental to make an official document available where different states are defined and what they look like (e.g. meaning of different colours, limitations of the system). 2/9 ATCO consider useful to have general guidelines on how to manage different situations when the GA is less than 100%. Should the system be able to display the GA probability while it changes in time, the guidelines should indicate the best practices when the GA probability exceeds different thresholds. On this point, however, the opinions differ and 1/9 ATCOs pointed out that a document that thoroughly describes all possible options and gives indications on how to react to different situations will soon become too voluminous and thus impractical.	

Co-funded by the European Union

Appendix B SafeOPS Human Factors Design Requirements

Table 7. Human Factors Design Requirements specific to the SafeOPS Tool

REQ ID	Title	Requirement	Rationale
REQ-SafeOPS 22-TS-HFSR-002-1	Visibility, Transparency and Timeliness	The design should always keep users informed about what is going on, through appropriate feedback within a reasonable amount of time.	When users know the current system status, they learn the outcome of their prior interactions and determine next steps.
REQ-SafeOPS 22-TS-HFSR-002-2	Window Management	All information that a controller needs to accomplish a task that is essential and time critical is located on a single page or in a single window.	'Window management' (e.g., accessing, sizing, moving, overlaying, and closing windows) takes time and effort. This can increase the user's workload and distract them from the primary task. Too many windows causes mental overload because the user must keep track of window locations and pertinent data items within any covered windows.
REQ-SafeOPS 22-TS-HFSR-002-3	Critical Information	Critical information should never disappear from the screen without being deleted or suppressed by the controller.	At no time, should the computer be solely in charge of how much information is presented on the display.
REQ-SafeOPS 22-TS-HFSR-002-4	Updated Data	Updated data (when something changes) are emphasised effectively so that it attracts the controller's attention.	The goal of visual alerts and coding techniques are to effectively direct the controller's attention to important information that might otherwise be missed or critical situations that might otherwise take longer to recognize. Both the physical and functional parameters of alerts should be considered in order to ensure that the alert is effective and useful.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

REQ-SafeOPS 22-TS-HFSR-002-5	Information obstruction and clutter	Essential ATC information is never blocked or obstructed by other information. Display clutter is not a problem.	This provides for maximum efficiency in task and minimises something critical being missed. It also minimizes the number of simultaneously displayed windows required for a task, or switching between display windows/systems, and minimises the presentation of data irrelevant to the task at hand.
REQ-SafeOPS 22-TS-HFSR-002-6	Coding Redundancy	When the meaning of the colour is critical, colour is used redundantly with another type of visual cue, such as shape, text, or size. For example, all yellow objects have a triangular shape.	Visual coding techniques, such as colour, brightness, or blinking, can be used to attract the controller's attention to unusual situations or potential problems. Visual coding alone, however, may not be sufficient, overuse of vis ual coding will be self-defeating. If the controller is focused on a particular situation, they may not detect unusual visual codes.
REQ-SafeOPS 22-TS-HFSR-002-7	Colour Perception	Colours are far enough apart in perceptual terms (e.g. distinctly different colours) that they are not able to be confused even when "washed out" by sunlight, if applicable. When the controller must distinguish between the colour of characters and symbols, small blue characters and symbols are not used as they have low luminance and visual resolution of fine detail is poor for blue. Saturated (i.e., vivid) red and blue are not presented next to each other.	The appearance of a colour can be altered by another colour next to it or another colour seen just before or after it.
REQ-SafeOPS 22-TS-HFSR-002-8	Number of colour codes	The controller will not need to identify more than six colours (to interpret the meaning of the colour when it stands alone). E.g. no more than six colours should be assigned a specific meaning.	Significant individual differences exist between people's ability to recognize colour. In research individuals were unable to discriminate more than six colours on a display. Other researchers found that advantages to color-coding decreased as more than five to six colours were employed. This is consistent with knowledge on

			human short-term memory characteristics, which indicates that memory-based codes limited to five to six dimensions present the least difficulty to humans.
REQ-SafeOPS 22-TS-HFSR-002-9	Standard Colour convention	The standard convention across aviation for red/yellow/green should be adhered to. Where the colour red is used only for warning/danger, yellow is used to indicate caution, green is used to indicate normal/ready status.	This is the stereotypical meaning, associated with red, amber/yellow and green, in western cultures.
REQ-SafeOPS 22-TS-HFSR-002-10	Location of High-priority alerts	High-priority alerts and other critical information are located within the central display area (i.e., the central 15 degrees of the area where the controller normally looks, given the normal viewing position).	Visual alerts will be recognized more quickly when placed in the area of the visual field that has the best visual acuity than when placed in the periphery. This optimal area can be thought of as a cone extending from the normal line of sight with a radius of 15 degrees. (The normal line of sight is the line 15 degrees below the line extending horizontally from the centre of the pupil.)
REQ-SafeOPS 22-TS-HFSR-002-11	Irrelevant information	Interfaces should not contain information which is irrelevant or rarely needed. Every extra unit of information in an interface competes with the relevant units of information and diminishes their relative visibility.	This keeps the content and visual design focused on the essentials, ensuring that the visual elements of the interface support the user's primary goals. It also prevents unnecessary elements from distracting users from the information they really need.
REQ-SafeOPS 22-TS-HFSR-002-12	Icons	The meaning of each icon is immediately apparent to the controller or it is labelled.	An icon is a graphical object that can be acted on, represent a command, or represent an application.
REQ-SafeOPS 22-TS-HFSR-002-13	Acronyms, language and labels	Acronyms in the new display system have the same meanings as in the previous system. Users should not have to wonder whether different words, situations, or actions mean	The idea is to maximize positive transfer of knowledge and skills from the old to the new system. Building on controllers' experience can benefit the workforce through simplified and/or reduced training,

		the same thing. The design should make use of the language that is clear to the intended user, e.g. use words, phrases, and concepts familiar to the user. Placement of standard data fields, and Formats used within data fields, is consistent from one display to another. Labels, terms, and abbreviations are used consistently across the display set.	particularly if familiar terminology and symbology are retained. Negative transfer is minimized from existing displays to new/different displays (Negative transfer occurs when an individual's learned response to an event is wrongly applied to a similar event in a different context).
REQ-SafeOPS 22-TS-HFSR-002-14	System arrangement	The 'mental model' users have of the system should match the way the system is laid out and operates, e.g. actions and options provided by the system should match actions performed by the user.	Capitalizes on people's existing knowledge and helps them easily learn an interface, with no need for training, but also minimises error.
REQ-SafeOPS 22-TS-HFSR-002-15	False Alarms	Alerts should have a low incidence of false alarms.	A high false alarm rate can lead to lack of confidence in the system and a slower response time to the alerts.
REQ-SafeOPS 22-TS-HFSR-002-16	Help documentation	Its best if the system doesn't need any additional explanation. However, it may be necessary to provide help and documentation to help users understand how to complete their tasks.	Interface help comes in two forms: proactive and reactive. Proactive help is intended to get users familiar with an interface while reactive help is meant for troubleshooting and gaining system proficiency.

EUROPEAN PARTNERSHIP

Co-funded by the European Union