Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 64 (2022) 316–327

Transportation Research Procedia

www.elsevier.com/locate/procedia

International Scientific Conference „The Science and Development of Transport - Znanost i razvitak prometa"

Novel Artificial Situational Awareness System is Comparable with Human Situational Awareness in the En-route Air Traffic Control Domain

Kristina Samardžić[a,*], Tomislav Radišić[a], Ivan Tukarić[a], Ruth Esther Häusler Hermann[b]

[a]*Faculty of Transport and Traffic Sciences, Department of Aeronautics, University of Zagreb, Zagreb, Croatia*
[b]*School of Engineering, Centre for Aviation., Zurich University of Applied Sciences, Winterthur, Switzerland*

Abstract

AISA project introduces human-machine distributed situational awareness (SA) in en-route air traffic control. This paper is a result of the project's preliminary study. As a part of the project, a comparison between artificial and human situational awareness is performed. To investigate if the artificial SA could undistractedly monitor traffic, air traffic controller (ATCO) tasks are defined within Knowledge Graph (KG) system. KG system serves as a database with the ability to define relationships between data points. Human-in-the-loop simulations were performed to acquire needed data. ATCOs conducted different traffic scenarios which were later analysed and assessed regarding SA indicators defined in the paper. Tasks defined within the KG-based system can produce artificial SA that can successfully identify and complement all human SA degradation occurrences, contributing to Team SA. Therefore, defined tasks are sufficient to apply to the traffic data and to produce adequate machine SA.

Peer-review under responsibility of the scientific committee of the International Scientific Conference „The Science and Development of Transport - Znanost i razvitak prometa –ZIRP2022

Keywords: Situation awareness; Knowledge graph; air traffic control; monitoring tasks.

* Corresponding author. Tel.: +385 1 245 7745.
E-mail address: ksamardzic@fpz.unizg.hr

2352-1465 © 2022 The Authors. Published by ELSEVIER B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of the International Scientific Conference „The Science and Development of Transport - Znanost i razvitak prometa –ZIRP2022
10.1016/j.trpro.2022.09.036

1. Introduction

Situation awareness in air traffic control contains many different working tasks where degradation in any of them could lead to a violation of separation minima. Therefore, an ATCO (Air traffic controller) should create and maintain a mental image of traffic situations regardless of adverse occurrences, pilot non-compliances, or emergencies, Vidulich et al. (1994). To preserve situation awareness in a high task load, shared or team situation awareness is introduced. Team SA (TSA) is an overlap of each member's scope and level of situation awareness. In TSA, members simultaneously contribute and combine knowledge and perception. To have high team performance, each team member must have undistracted handling of their working tasks and maintain individual situational awareness (SA) that is comparable and concurrent to team SA, Endsley and Robertson (2000). In previous studies, Kaber et. al (2005) analysed how automation affects SA. Automation concerns different tasks, whereof automated information-processing functions enhanced ATCO performance strongest. Furthermore, Jipp and Ackerman (2016) investigated if the working memory and information-processing affect SA when introduced with the automation. The results indicated that at the high level of automation, performance did improve but the SA got worsened. The authors proposed the idea to enable automation to monitor ATCO when executing the tasks and to learn from those processes. The human-machine working environment is analysed in Svensson, (2020). She explored the TSA in-between only humans and applied the same concept to human-machine TSA. Two observed factors affecting TSA are boundary awareness (what are the limitations of the other team member) and implicit communication (unexpressed or latent cues used instead of the communication). Stanton et al. (2006) defined distributed SA as a group of knowledge objects that need to be activated to manage tasks. Relation between knowledge objects and their activation is inevitable for describing distributed SA environment, where the system requires information about their relation and activation to describe any given situation. Authors state that distributed SA increases workload and uncertainty as there are more knowledge objects activated that need to be addressed, managed, and tracked.

The system reported in this paper is Knowledge Graph (KG) based system in which the logic of ATCO tasks, which are used to produce SA, are translated into the system to form artificial SA. To improve operational performance, ATM should rely on human-centric Artificial Intelligence (AI). The idea is to add an AI component to the working environment and produce human-machine collaboration without excluding human from the loop. Therefore, introducing AI specifically for the ATCO monitoring tasks provides the ability to produce such interaction.

This paper investigates if the development of the artificial SA when confronted with possible degraded human SA, could provide advisory and create a foundation for distributed or team SA. TSA requires that both members have access to the same data. Therefore, knowledge bases with all the relevant facts and rules are encoded in KG and formed to provide a connection between data for human and machine use, AISA Consortium (2020). The system was developed as a framework for building foundation for successful cooperation between human and machine is presented in the SESAR's project AISA (AI Situational Awareness Foundation for Advancing Automation). The project aims to research the impact of human-machine distributed situation awareness to automate monitoring tasks in en-route operations and to introduce artificial SA as a contributor to the TSA. The first step is defining artificial (machine) SA as capable of automating ATCO monitoring tasks and comparable to human SA. Human-in-the-loop experiment are performed to enable and test the proposed system's concept-of-operation. Experiment provided data for the human SA and artificial SA comparison, based on the ATCO monitoring tasks which are used to maintain and preserve SA. Monitoring ATCO tasks are defined in the ConOps of the AISA project, AISA Consortium (2020). Tasks are grouped into eleven disjunctive categories and each of these categories covers one aspect of the monitoring tasks. Data used for the comparison included recordings of the traffic situation and corresponding ATCO actions while being introduced with non-standard traffic occurrences. This comparison proves that artificial and human SA share the same logic and that artificial SA could be used as a part of TSA to identify human SA degradation. Only selected ATCO tasks which could be used to assess SA, later mentioned in detail, were applied to analyse and compare human SA with the artificial. SA indicators are defined and operationalized in relation to the scenarios investigated and occurrences of preserved or degraded SA are identified.

The work is achieved through the following sections. Section 2 describes the 3-level SA used to assess the human SA and links it with the ATC scope of work. Section 3 describes the monitoring ATCO tasks logic used for the comparison. In Section 4 the experimental environment and methodology are explained. In Section 5 the comparison of the artificial and human SA is performed based on the SA indicators. Lastly, Sections 6 and 0 present the comparison results and conclusion.

2. Situation awareness in ATC

According to Endsley (1995) there are three situation awareness levels. Each level is a prerequisite for the next level. Achieving one level is necessary but not sufficient for accomplishing the next higher level. From the perception (noticing two aircraft on the same altitude), through interpretation (becoming aware of their routes crossing), to projection (anticipate future time and position of the separation loss) ATCO processes the information.

On SA level 1, perception is acknowledging outer information and becoming aware of it. The issue arises when the selectivity of attention is challenged. ATCOs need to perceive information from labels and memorise information from radio communication with pilots. Constantly changing and updating labels requires a certain amount of attention. ATCOs need to prioritise and direct attention to essential information. On SA level 2, interpretation is giving perceived information a meaning. ATCOs should easily generate an expectation from a single situation. Based on their experience, ATCOs can build mental models about traffic situations that may occur and use them to generate expectations from a single situation. On SA level 3, projection is planning and foreseeing what could happen based on the current trends. Working memory is the key to maintaining the current mental model up to date. Due to its limited capacity, especially under high stress, information can get missed, be forgotten, or wrongly remembered. To reduce the impact of the working memory capacity, long-term memory stored mental models may be used.

Level 1 confronts ATCOs the problem of focusing attention only on relevant information. Selectivity could lead to inattentional blindness -steering the attention focus on the achievement of the goals rather than on essential information. Level 2 and 3 situation awareness assures that the information is acknowledged and understood. Attention should be driven to information cues that trigger expectations from SA levels 2 and 3. Preconceptions, although conducive for creating mental models, may lead to situational awareness level 2 errors. Furthermore, confirmation bias produces an error where one can be prevented from maintaining accurate situational awareness. Level 3 situation awareness is a challenging task regarding the current radar display types in use. Maintaining situational awareness in a dynamic environment by using a 2D radar display forces ATCOs to reassess traffic situations and implement changes if needed.

There are several methods for measuring SA. Subjective rating of the participant, either by ATCO or subject matter expert assesses SA, is the most direct method but lacks with validity as subjects cannot rate SA on aspects they are not aware of. SA. Endsley (1995b) describes subjective measurement rather as the opinion one has on his SA. The less subjective method is implicit performance measurement where SA is measured based on some predefined performance indicators, Salmon et al. (2009). Another method asks the participant questions about the current traffic situation and the one that will evolve. They are called probe techniques and can be performed in two versions. Probe technique with freezing includes stopping the simulation and asking questions while blinding information sources such as radar screens. Online probe technique could contribute to a higher workload or performance obstruction as questions are asked during task performance. As there is no one method without any limitations and shortcomings, a combination of subjective rating, online and "freezing" probe techniques are used for the scope of the AISA project.

Artificial SA assessment can be performed by neural networks. Bench T (2022) introduces neural network integration in the AI systems which would enable the system to learn from the previous events and make decision. Another approach includes expert's opinion on the generated artificial SA, Pullum (2021). A more objective approach involves the use of framework that, based on the defined conditions, evaluate the system's awareness level and ability of the system to assess its own performance, Jantsch and Tammemäe (2014). None of these methods use multiple sources to mitigate any inconsistencies caused by training data or system development. Svensson (2020) explored how teamwork consisted of only humans can be applied to the human-machine teamwork. She relied on the questionnaires, audio and screen recording to conduct simulator study and explore human-machine working environment. She measured the impact automation has on the human performance and introduced Boundary Awareness, the ability to know the current situational conditions that can affect automation performance, as a necessary in the future automation designing systems.

In this paper, SA is measured with the respect to indicators where each indicator has criteria and set of actions that identify preserved or degraded SA. This method, is based only on the performance, is used to measure and assess both human and artificial SA. Indicators with its criteria and actions are described in Section 5 of this paper.

3. Definition of the ATCO tasks in the KG-based system

The AISA project proposes a novel hybrid system (henceforth referred to as the KG-based system) that uses a knowledge graph, machine learning modules and a reasoning engine to achieve artificial SA. The KG contains aeronautical data collected from traffic situation simulations. These simulations were a part of the experiment whose detailed description is in Section 4 of the paper. Combined with machine learning module outputs, which are used to predict future traffic states, the integrated data is input into the reasoning engine to reach conclusions about the situation and produce artificial SA. The reasoning engine contains rules derived from expert knowledge of ATC procedures to simulate ATCO tasks. System structure and components are described in more detail in Concept of Operation, AISA Consortium (2020), while the architecture of the core system is explained in Proof-of-concept KG system, AISA Consortium (2021).

System tasks were then applied to the data to generate system outputs. The AISA project successfully automated 46 out of 57 en-route monitoring tasks defined in the project's Concept of Operation. Not all tasks are used for this study as the selection of applied tasks is conditioned by the choice of traffic situations that could be used for the SA assessment. These situations are referred to as "Situation awareness indicators". There are a few reasons why all 46 tasks weren't used for the SA comparison: (i) tasks that check if the a/c is maintaining the current state (FL, speed, heading) have the same system output repetitively and are hardly useful for the assessment, (ii) some monitoring tasks couldn't be used for the comparison as they don't define SA distinctly, (iii) the focus of the comparison was on the SA indicators, tasks that could identify those indicators were given a priority, (iv) over-delivery of the system outputs complicates the assessment and the comparison. Table 1 presents initially selected ATCO tasks that were used for the experiment analysis. ATCO task categories are extracted from article Gordon and Fitzpatrick, (2007). They identified and defined the en-route ATCO tasks with the following features: task name, task plan, operator, task description and task reference. From task plan it can be recognised how are tasks connected and which tasks are in common. Task categories suggest which ATCO tasks are related i.e., which tasks are connected in the task plan. Based on these tasks, five SA indicators are defined to assess and compare human and artificial SA.

Table 1 Selected monitoring ATCO tasks

ATCO task	Task category	Task description
Task 1	Conformance management	Check that A/C is climbing/descending towards cleared FL
Task 2	Conformance management	Check that A/C is flying at cleared speed
Task 3	Assess if exit conditions are met	Check that A/C will reach the exit point on the required FL
Task 4	Conflict management	Check all A/C pairs for conflict (ML (Machine Learning) module)
Task 5	Execute aircraft's plan	Detect A/C that have to climb/descend to exit FL
Task 6	Maximize quality of service	Detect direct-to candidates
Task 7	Transfer Aircraft	Check which aircraft need to be transferred
Task 8	Assume, Identify, and Confirm Flight	Check that aircraft is incoming
Task 9	Execute Aircraft's Plan	Check if planned trajectory passes through restricted airspace

3.1. The definition of selected ATCO tasks in the KG-based system

Task 1 Check that A/C is climbing/descending towards cleared FL

Task checks if the aircraft's current altitude is approaching the cleared altitude. Needed input data is: cleared FL, current FL, previous FL, current vertical rate.

To check if the aircraft is climbing toward cleared FL, the KG system checks several conditions: current FL is not cleared FL, previous FL is less than current FL, vertical rate (ROC (rate of climb)) is not 0.

The same methodology can also be applicable to descending traffic.

Task 2 Check that A/C is flying at cleared speed

Task compares if the current speed and cleared speed are the same. Needed input data is: cleared speed and current speed.

Task compares the two values from the input data. In en-route control Mach number is used to issue speed clearances while KG input data is in knots. To precisely recalculate Mach number to knots the speed of sound should be recalculated for the given atmospheric conditions. In the experiment, to minimise the error affected by recalculation, the buffer had been added.

Task 3 Check that A/C will reach the exit point on the required FL

Task checks if the vertical distance aircraft needs to pass is manageable regarding its current distance to the exit point. Several data are needed: cleared FL, exit FL, exit point position, current FL, current vertical rate, current position, current speed.

To check if the aircraft will manage to climb to exit FL before passing the exit point, the task firstly calculates two distances. One to check what vertical distance the aircraft needs to pass and another to calculate the needed time from the current position to the exit point. With these two distances, the calculated vertical rate can be derived. If the calculated vertical rate is less or equal to the current vertical rate, the aircraft will reach the required FL.

Task 4 Check all A/C pairs for conflict (ML module)

Task 4 relies on the conflict detection (CD) Machine Learning (ML) module to identify aircraft pairs that will have separation infringement. CD ML module is trained on the historical data. It uses all aircraft current data (position, heading, altitude, speed, and vertical rate) and provides output about minimum distances between considered aircraft pairs. CD ML module output is then filtered by KG System to check if the predicted minimum distances are information of interest to the controller e.g., a predicted minimum distance of 25 NM would be filtered out. Furthermore, the CD ML module output provides information on when would predicted minimum distance happen. It calculates distance and time to minimum distance.

Task 5 Detect A/C that have to climb/descend to exit FL

Task 5 checks if the aircraft is on its planned exit FL. Input data needed for this task are: current FL and exit FL.

The task compares the value of the two input values. Note that exit FL is agreed upon in the Letters of Agreement between two different sectors and should be respected accordingly.

Task 6 Detect direct-to candidates

Task 6 uses information about Temporary Reserved Airspace (TRA) activation status and checks if the aircraft can be issued with a direct point. Input data needed is: trajectory list, current position, current FL, sector vertical and horizontal boundary, TRA volume boundary.

Task calculates the azimuth from the current position to the exit trajectory list point. If calculated azimuth crosses the TRA volume, another Task that checks TRA activation status is triggered. KG system is provided with the list of all possible sector exit points to check if considered aircraft is an exit point direct-to candidate.

Task 7 Check which aircraft need to be transferred

Task 7 uses information about the aircraft's current position and calculates if the aircraft has less than 2 minutes to the sector boundary. Input data needed is: trajectory list, sector boundary, current FL, current position, current speed and vertical speed.

Task calculates the time between the current position and sector boundary. If the calculated time is less than 2 minutes aircraft should be transferred. The task also considers if the exit conditions are met before leaving the sector.

Task 8 Check that aircraft is incoming

Task 8 compares previous and current aircraft positions to provide information if the traffic is incoming or not. Input data needed is: sector boundary, sector vertical range, current FL, current position, previous position, current speed and vertical speed.

Firstly, (i) the difference between sector boundary and aircraft position is calculated together with the (ii) distance between sector boundary and previous time-step aircraft position. Task also uses vertical speed to calculate if the aircraft is approaching a vertical sector boundary. If the distance (i) is less than distance (ii) and the aircraft is within vertical sector range, the aircraft is identified as incoming.

Task 9 Check if planned trajectory passes through restricted airspace

Task 9 uses information about Temporary Reserved Airspace (TRA) activation status and uses aircraft trajectory to check if the planned route crosses the TRA. Input data needed is: trajectory list, current position, current FL, sector vertical and horizontal boundary, TRA volume boundary and TRA activation times.

Task checks if the point from the trajectory list is affected by TRA activation. If it is affected, the estimated time over that point is compared to the TRA activation times. If the aircraft is expected to be over the affected point while TRA is active, planned trajectory passes through restricted airspace.

3.2. Situation awareness indicators

As mentioned before, previously selected ATCO tasks could assess situation awareness. AISA project focused on automating monitoring tasks. SA indicators were selected to imply the possible SA degradation. Selected indicators and KG tasks capable of issuing output are presented in Table 2. One SA indicator usually covers several ATCO working tasks which is why multiple tasks correspond to one SA indicator.

To check if the aircraft is transferred on time Task 7, 3 and 5 should be used. As Tasks 3 and 5 check if the aircraft will satisfy exit FL (Letters of agreement, LOA), Task 7 monitors if the aircraft is 2 minutes or less to the sector boundary.

Non-compliance to ATCO instruction or clearance is another SA indicator. Tasks 1 and 2 could identify any FL or speed deviation.

At this stage, the system cannot by itself check if the assumed aircraft is also assumed in the label as these two systems are not connected. At this stage, human effort should be used to check if the ATCO assumed aircraft on the frequency and on the label. However, Task 8. can identify incoming traffic which should be assumed.

Quality of service regarding issuing direct routes is covered by Task 6. To check if the current aircraft planned route crosses TRA, task 9 should be used. These two tasks had been used to maximize the quality of service and to safely manage aircraft with new routes.

The last considered indicator is conflict detection related. Task 4 is used to identify actions related to that SA indicator.

Table 2 Selected SA indicators and corresponding ATCO tasks

SA indicator	ATCO task
Check if aircraft is transferred on time	Task 3, Task 5 & Task 7
Check for reaction to non-compliance	Task 1 & Task 2
Check if flight is assumed on the label	Task 8
Check if aircraft will fly through restricted airspace and if they can use previously restricted airspace	Task 6 & Task 9
Check if aircraft are in a conflict	Task 4

4. Experimental environment

The conduct of the experiment was conditioned by the need for data based on which comparison of human and artificial SA could be made. The experiment consisted of human-in-the-loop simulations. The experiment took place on the premises of skyguide where 20 licensed En Route ATCOs conducted a total of 120 exercises. Two exercises were conducted simultaneously in the same room. Different biometric parameters were measured while performing the run. An example of the experimental setup is presented in Fig.1. Since the ATCOs worked in a well-known sector but not in a known simulator system, they were accompanied by two subject-matter experts (SME) ATCOs who assisted them with all the ambiguities. Also, they oversaw supporting the ATCOs with all the technical difficulties and presenting them with the tools available on the simulator. The ESCAPE Light simulator platform ("EUROCONTROL simulation capabilities and platform for experimentation") used for the experiment does not contain all the safety nets that the skyguide current system has. Two pseudo pilots manipulated the aircraft and responded to ATCO clearances.

Fig.1. ESCAPE Light simulator platform ATCO interface

Since there was only an executive position without a planner, no coordination with neighboring sectors was possible. Furthermore, all working tasks of the planner were performed by the executive. In order to achieve a more realistic traffic situation, the incoming traffic before the sector boundary cannot be issued with flight level (FL) change instruction but could be turned. This adaptation was necessary because more direct routes are preferred, they reduce the excessive flown distance but can also resolve a conflict. While performing the exercises ATCOs were asked in regular intervals about their workload level.

The sequence of the experiment activities is shown in Fig.2 All ATCOs firstly did a training scenario where they could notice if they have system issues and assimilate with the experimental environment. The training scenario was followed by a light traffic scenario. The purpose of this scenario is to enable independent performance while traffic count being low. The Light scenario lasted 7 minutes and included 19 aircraft. All the other scenarios were designed based on real traffic situations and were performed in random order. Some traffic, with consulting SME, was implemented deliberately to create conflicts.

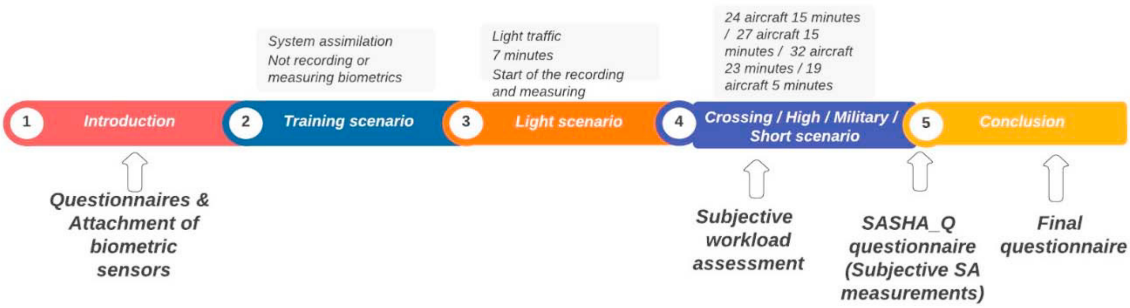

Fig.2. Human-in-the-loop experiment flow diagram

After two introductory scenarios, ATCOs were randomly given four more scenarios where each of them had a "specific occurrence". It is important to note that these "specific occurrences" are not used to assess ATCO performance. ATCOs are not rated here as "specific occurrences" can only provide subjective opinion on overall human performance. Furthermore, ESCAPE Light is lacking safety nets that are present in skyguide's Air situation Display (ASD) system tool skyvisu, it would be inappropriate to compare ATCO performance in the experiment with the actual operational performance.

The Crossing scenario included 24 aircraft and lasted 15 minutes. The purpose of this scenario is to produce several conflicts one of which is a triple conflict. The triple crossing would without any ATCO input trigger Short term conflict alert (STCA). The specific occurrence in this scenario was a flight level bust. The time of this occurrence depended on the pseudo-pilot decision, preferably while resolving the conflict. As flight level bust wouldn't initiate conflict with other traffic, no alarm or tool available in the simulator could identify it.

The High traffic scenario included 27 aircraft and lasted 15 minutes. Without any deliberately created conflicts, the focus in this scenario is on handling high traffic count for the short time. The specific occurrence herein is speed non-compliance. Initially, the following traffic with the second one faster is an undesirable situation and therefore must be corrected before transferring traffic to another sector. Pseudo-pilots wouldn't correct the speed on the first ATCO clearance, creating non-compliance. The specificity of this occurrence is that it didn't happen in every exercise. There were multiple solutions to that conflict situation so it couldn't be used every time for the analysis. There is another exit conflict, produced after issuing aircraft direct routes. It is important to notice that there is no available ESCAPE Light simulator tool that recognises speed non-compliance and created traffic situation is rarely present in the actual operations.

The Military scenario is longer than the previous, 23 minutes with 32 aircraft altogether. Two temporary reserved areas (TRAs) were alternately activated. The specific occurrence in this scenario is the opportunity to improve the quality of service after the first TRA became available and rerouting traffic after the second TRA is activated. ATCOs were familiar with the information on when will and what will be activated. They were also permitted to use their standard rerouting procedures after TRA activation.

The Short scenario included 19 aircraft and lasted 5 minutes. In a brief time, much traffic needed flight level change. The Specific occurrence included flight level non-compliance. Pseudo-pilots deliberately didn't descend aircraft on the first issued clearance. This non-compliance would be easier to recognise as it would initiate a conflict with another traffic. Due to the shortness of the scenario, it ended before any conflict happened.

During the experiment eye tracking data and biometric data was measured and recorded. Additional questionnaires were asked before, during and after each exercise. This data could be used to evaluate human SA

5. Artificial and human SA degradation comparison

In this chapter detailed comparison of artificial degraded SA and human degraded SA occurrences is presented. Indicators that imply SA degradation are defined based on criteria that need to be met in order to conclude so. One indicator and its criteria contain several occurrences based on which SA degradation could be identified. The number of these occurrences are used for the validation and comparison of the overall human and artificial SA. The sum of degraded human and artificial SA occurrences together with the system outputs are in Table 3 which is adapted from the AISA Consortium (2022). In the next paragraphs, every SA indicator is separately described.

5.1. Check if the aircraft is transferred on time

Transfer of frequency should be achieved prior to the sector boundary. It is preferred to transfer aircraft 1-2 minutes before sector boundary if the traffic situation allows it. As there is no precise time or distance of when the handoff should happen, the occurrence of degraded human SA is not transferring the aircraft before the sector boundary. Transferred aircraft should also respect predefined exit FL. Not clearing aircraft to its exit flight level without prior coordination would also account for the degradation of SA.

Degraded artificial SA would mean that the KG system does not recognise the aircraft close to the sector boundary and not at exit FL.

5.2. Check for reaction to non-compliance

Non-compliance occurrences within this SA indicator could be regarding cleared speed, FL, vertical rate, and planned route. Herein, any non-compliance occurrence not corrected by ATCO after the 30s is considered as degraded human SA. The 30s are used as a buffer to allow ATCO to react in a given situation. The buffer value is agreed with the Subject matter expert.

Four different ATCO tasks are needed to identify non-compliance occurrences. If the KG-based system couldn't identify immediately non-compliance, the degraded artificial SA is considered to have happened.

5.3. Check if flight is assumed on the label

After the aircraft had made an initial call, ATCO should assume it on the label. If ATCO doesn't assume the label or assume the wrong label, his/her SA is degraded.

AISA system currently cannot identify if the label is assumed wrongly. To assess this occurrence, task which checks if the initial call has been made should notify which aircraft should be assumed. Without properly identifying which aircraft made the initial call, artificial SA degradation occurs.

5.4. Check if aircraft will fly through restricted airspace and if they can use previously restricted airspace

TRAs shouldn't be crossed in any circumstances. Five minutes prior to activating the TRA, ATCOs should reroute all traffic crossing it, including inbound traffic. Furthermore, after the TRA becomes inactive, all rerouted traffic should be issued a direct route to improve the quality of service. As it is hard to acknowledge when ATCO didn't clear aircraft to direct route and should have, herein only the traffic entering the TRA is an occurrence of degraded human SA.

For the artificial SA on the other hand, the KG-based system should identify all traffic that could use previously reserved airspace and, indicate which traffic needs to be rerouted.

5.5. Check if aircraft are in conflict

Aircraft in conflict would violate horizontal and/or vertical separation minima if they remain on the planned route. ATCO should recognise those occurrences and react with the appropriate actions to separate aircraft. The occurrence of degraded human SA is the activation of the Short-Term Conflict Alert (STCA) –safety net which notifies ATCO of a potential violation of separation minima 2 minutes before the violation time.

Degraded artificial SA would mean that the KG system does not recognise aircraft pair in conflict at all, or the predicted distances are greater than the separation minima which would indicate that the aircraft pair is separated sufficiently.

Table 3 human and artificial degraded SA occurrences with AISA system output

SA indicator	Degraded human SA occurrence	Degraded artificial SA occurrence
Check if the aircraft is transferred on time	Aircraft not transferred before sector boundary	"Transfer aircraft" output not provided 2 minutes before sector boundary
	Aircraft is not cleared to the exit FL before sector boundary	"Aircraft will not reach the exit point on the required FL" output not provided for the aircraft that are not maintaining exit FL on the exit point
Check for reaction to non-compliance	Correction action not issued 30 s from non-compliance	"Aircraft is not following the 3D trajectory" output not provided when the aircraft deviates from the trajectory
		"Flight Level bust" output not provided after the cleared FL has been passed
		"Aircraft is not at cleared speed" output not provided after the cleared speed has been issued.
Check if the flight is assumed on the label	Aircraft not assumed 30 s after the initial call	"Aircraft has sent the initial call" output not provided after it occurred.
Check if aircraft will fly through restricted airspace and if they can use previously restricted airspace	The aircraft entered active TRA.	"Planned trajectory passes through restricted airspace" output not provided after the TRA activated and trajectory list crosses TRA.
		"Aircraft is not a candidate for direct-to" output not provided after aircraft could cross previously active TRA
Check if aircraft are in a conflict	STCA went off	For aircraft pair that would be in a conflict, the predicted minimum distance is greater than the actual.

6. Results

Out of 80 exercises performed in the experiment, 20 of them were randomly selected for the analysis. Six exercises from the crossing, high traffic, and military scenario and two exercises from the short scenario. For each exercise number of occurrences of each SA indicator is identified and counted. Fig.3. Occurrences of SA indicators for KG-based system and ATCO respectively depicts how many occurrences in each indicator category are. Preserved SA occurrences are all those instances where system output or ATCO performance didn't indicate potential degradation of SA i.e., conditions from Table 3 are not met. Detailed insight into the analysed data is presented in Table 4 Number of degraded and preserved SA occurrences with respect to absolute SA indicator valuewhere the number of occurrences for each SA indicator is presented with respect to absolute value.

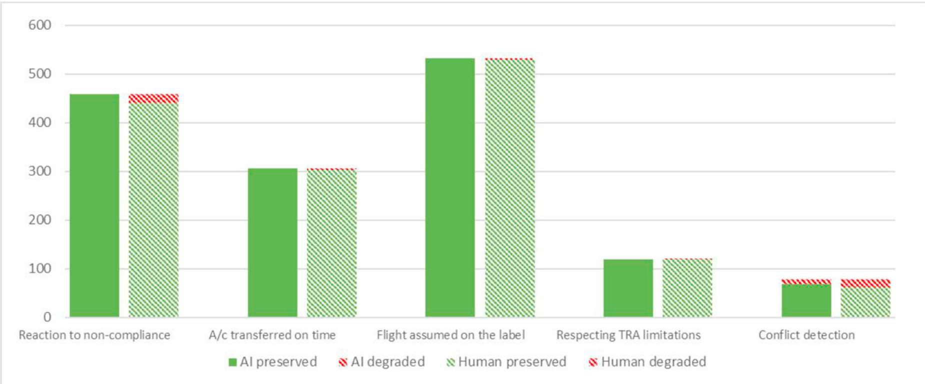

Fig.3. Occurrences of SA indicators for KG-based system and ATCO respectively

Artificial SA degradation is present only in the conflict detection indicator category as seen in Fig.3. The definition of the task herein doesn't produce an error. The error comes from the lack of conflict detection ML module accuracy. The task is defined only to use ML module prediction to deliver system output. The KG-based system can assess and track the ML module performance, but the accuracy is out of the scope of this paper. Although ML module performance could be analysed separately, conflict detection and resolution are a crucial part of assessing human SA. Therefore, SA indicator regarding conflict detection should be considered when comparing artificial and human SA.

From Fig.3. Occurrences of SA indicators for KG-based system and ATCO respectively, not reacting to non-compliance is the most frequent degraded human SA indicator category. As "specific occurrences" are deliberately set in scenarios to assess human and artificial SA, reason could be that ATCOs rely on many safety nets that could easily identify those occurrences. From Table 4 Number of degraded and preserved SA occurrences with respect to absolute SA indicator value again "Conflict detection" related indicator and "Reaction to non-compliance" have the highest degraded SA percentage. As there aren't any degraded artificial SA occurrences, the definition of the tasks and their applicability are confirmed to be correct. Also, defined ATCO tasks in the KG-based system apply to the collected data and the artificial SA could achieve high accuracy based on SA indicators.

Table 4 Number of degraded and preserved SA occurrences with respect to absolute SA indicator value

SA indicator	Number. of occurrences	Preserved human SA occurrences		Preserved artificial SA occurrences		Degraded human SA occurrence		Degraded artificial SA occurrences	
Check for reaction to non-compliance	458	440	96,07%	458	100%	18	3,93%	0	None
Check if the aircraft is transferred on time	307	303	98,7%	307	100%	4	1,30%	0	None

Check if the flight is assumed on the label	533	529	99,25%	533	100%	4	0,75%	0	None
Check if aircraft will fly through restricted airspace and if they can use previously restricted airspace	120	119	99,17%	120	100%	1	0,83%	0	None
Check if aircraft are in a conflict	78	62	79,49%	68	87,18%	16	20,51%	10	12,82%

Statistical test was performed on the data from Table 3. To carry out the statistical test, it was necessary to present the data for each ATCO and for each indicator. The percentage of occurrences indicating preserved SA is expressed for both ATCO (human) and machine (artificial). As such, the data were subjected to an ANOVA ("analysis of variance") test which determines if there is a statistically significant difference between the corresponding population means. ANOVA test revealed that there was a statistically significance between group means for the indicator 1 where $F(1,38)= 4.098172$, $p= 0.012658$ and when all values from all five indicators are taken together $F(1,198)= 4.074$, $p=0.045$.

7. Discussion and conclusion

The main goal of the Artificial situational awareness system is to assist ATCOs in their perception of the air traffic situation. Based on the experiment conducted in an experimental environment a conclusion is made that the system is able to applicate collected data and produce artificial SA. The monitoring tasks proved to have high accuracy since the system reacted to all specific occurrences besides the conflict detection where its SA was degraded. However, conflict detection pertains to the ML module whose accuracy is out of the scope of this paper. The time needed for the KG based system to identify any of SA indicators is mainly less than ATCO's needed time. The immediate reaction to those occurrences proves system applicability and competence. As opposed to that, immediate reaction and notification can result in inappropriate usage of ATCO workload capacity. The time of notifying should be further discussed to agree upon the proper method without occupying ATCO's attention unnecessary. Further studies should include an analysis of a greater number of SA indicators for further confirmation of the level of SA and improve the accuracy of the ML module. The results of the preliminary study suggest that artificial SA could identify and complement potential human SA degradation and contribute to Team SA. The AISA System developed as a part of the exploratory research can contribute to the ATM environment where traffic situations demand continuous monitoring and timely response to different traffic occurrences.

Acknowledgement

This paper is part of the AISA project, which has received funding from the SESAR Joint Undertaking under grant agreement No 892618 under European Union's Horizon 2020 research and innovation program.

References

AISA Consortium, 2022. Report on human-machine distributed Situation Awareness 176.

AISA Consortium, 2021. Proof-of-concept KG system.

AISA Consortium, 2020. Concept of Operations for AI Situational Awareness (D2.1).

Can AI Systems Match Human-Level Situational Awareness? | Bench T, 2022. URL https://www.mouser.com/blog/can-ai-systems-match-human-level-situational-awareness (accessed 5.9.22).

Endsley, M.R., 1995a. Toward a Theory of Situation Awareness in Dynamic Systems. Hum Factors 37, 32–64.

Endsley, M.R., 1995b. Measurement of situation awareness in dynamic systems. Human Factors 37, 65–84.

Endsley, M.R., Robertson, M.M., 2000. Situation awareness in aircraft maintenance teams. International Journal of Industrial Ergonomics 26, 301–325.

Gordon, R., Fitzpatrick, M., 2007. Today's Operations Task Analysis - Human Factor Assessment.

Jantsch, A., Tammemäe, K., 2014. A framework of awareness for artificial subjects. Presented at the International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS), pp. 1–3.

Jipp, M., Ackerman, P.L., 2016. The Impact of Higher Levels of Automation on Performance and Situation Awareness: A Function of Information-Processing Ability and Working-Memory Capacity. Journal of Cognitive Engineering and Decision Making 10, 138–166.

Kaber, D.B., Wright, M.C., Prinzel, L.J., Clamann, M.P., 2005. Adaptive Automation of Human-Machine System Information-Processing Functions. Hum Factors 47, 730–741

Pullum, L., 2021. Verification and Validation of Systems in which AI is a Key Element. Oak Ridge National Lab., Hoboken, New Jersey, United States of America.

Salmon, P.M., Stanton, N.A., Walker, G.H., Jenkins, D., Ladva, D., Rafferty, L., Young, M., 2009. Measuring situation awareness in complex systems: comparison of measures study. International Journal of Industrial Ergonomics 39, 490–500.

Stanton, N.A., Stewart, R., Harris, D., Houghton, R.J., McMaster, R., Salmon, P., Hoyle, G., Walker, G., Linsell, M., Dymott, R., Green, D., 2006. Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology.

Svensson, Å., 2020. Human-automation teamwork: Current practices and future directions in air traffic control. Linköping University Electronic Press.

Vidulich, Dominguez, C.O., Vogel, E.B., & McMillan, G., 1994. Situation Awareness: Papers and Annotated Bibliography.