Human Interaction and Emerging Technologies (IHIET 2024), Vol. 157, 2024, 57–67
https://doi.org/10.54941/ahfe1005464

A Human Centric Design Approach for Future Human-AI Teaming in Aviation

Barry Kirwan[1], Roberto Venditti[2], Nikolas Giampaolo[2], and Miguel Villegas Sanchez[3]

[1]EUROCONTROL Innovation Hub, Bretigny-sur-Orge, France
[2]Deep Blue SARL, Roma, Italy
[3]Skyway, C/Albasanz, 14 Edificio Verona, 28037 Madrid, Spain

ABSTRACT

The introduction of Artificial Intelligence (AI) in aviation has already begun, with Machine Learning systems supporting aviation workers in a number of areas. So far, such AI additions can be seen as 'just more automation', as the human - whether pilot or air traffic controller - remains very much in command and control, maintaining situation awareness and being the principal safety barrier against accidents. Future AI systems, however, are likely to have a higher degree of autonomy, and a collaborative relationship is foreseen. This relationship, in which the human will 'partner' with 'Intelligent Assistants', is termed Human-AI Teaming. But how will trust between human and AI be achieved? How can we develop satisfactory AI 'explainability' functions so the humans can understand its advice and choices? And how do we assess the human-AI interaction design, whether visual, verbal, or gestural, so that safe performance is assured? The European HAIKU project has developed a provisional Human Factors Assurance methodology and applied it to several aviation use cases that vary in terms of their AI autonomy and are consequently a reasonable testbed for evaluating new approaches. This paper outlines the methodology and illustrates its application via one of the use cases.

Keywords: Human factors, Artificial intelligence, Human-AI teaming, Aviation, Systems design

INTRODUCTION

Human Factors and Aviation have been effective partners for decades. Systematic research has delivered guidance, standards and regulations in areas such as cockpit design, air traffic control/display interface design, fatigue management and crew resource management. Such guidance has helped aviation maintain its record as the safest mode of transportation, particularly when dealing with increasing levels of automation, whether in the cockpit or in the ground control systems.

The introduction of Artificial Intelligence (AI) in aviation has already begun, with Machine Learning systems supporting aviation workers in a number of areas. So far, such AI additions can be seen as 'just more automation', as the human - whether pilot or air traffic controller - remains very much in command and control, maintaining situation awareness and being the principal safety barrier against accidents. With the advent of future

AI systems likely to appear in the next decade, however, this is likely to change. AI systems are envisaged that will have a higher degree of autonomy. A collaborative relationship is foreseen - generically known as Human-AI Teaming - in which the human will 'partner' with 'Intelligent Assistants'. This may include the AI deciding what to do and executing its own tasks, negotiating with the human crew, and even reconsidering its goals as part of the team. This may require Human Factors to raise its game in human-system performance assurance, either upgrading its existing approach and techniques, or else adding new ones.

A first step is to understand how far AI autonomy is expected to develop in aviation in the coming decade, both from a regulatory standpoint (what is envisaged as being permissible), and a more concrete viewpoint via Human-AI Teaming aviation use cases. This step is outlined in the following two sections. A Human Factors Assurance methodology is then outlined and applied to one of the use cases to illustrate the approach and its outputs.

LEVELS OF AUTONOMY IN FUTURE AVIATION HUMAN-AI SYSTEMS

The European Union Aviation Safety Agency (EASA) regulatory guidance on AI proposes six categories of future Human-AI partnerships (EASA, 2023), in increasing degrees of AI autonomy, interpreted by the authors as follows:

- Machine Learning support (1A), already existing today;
- Cognitive Assistant (1B), equivalent to advanced automation support;
- Cooperative Agent (2A), able to complete tasks as demanded by the operator;
- Collaborative Agent (2B), an autonomous agent that works with human colleagues, but which can take initiative and execute tasks, as well as being capable of negotiating with its human counterparts;
- AI Executive Agent (3A), where the AI is running the show, but there is human oversight, and the human can intervene; and
- Fully Autonomous AI (3B), where the human cannot intervene.

Human-AI Teaming raises a host of questions and challenges for Human Factors, such as how to achieve trust between human and AI, how to achieve satisfactory 'explainability' functions in the AI so the human can understand its advice and choices, as well as how best to design means of human-AI interaction, whether visual, verbal, or gestural. Such questions are often best met head-on during the assessment of use cases. Four candidate use cases are outlined next.

AVIATION HUMAN-AI TEAMING USE CASES

The European HAIKU project (https://haikuproject.eu/) has developed a provisional methodology and applied it to several 'use cases', four of which relate to aviation operational settings:

- AI support in emergencies to a single pilot in the cockpit;

Figure 1: Examples of 4 HAIKU use cases.

- AI support to flight crew who unexpectedly need to divert to a different airport;
- AI support to tower controllers dealing with arriving and departing aircraft;
- An executive AI 'manager' of pilot-less drone and sky-taxi traffic in urban environments.

These four use cases (see Figure 1) vary in AI autonomy, and so are a reasonable testbed for applying a new Human Factors approach, as outlined below.

A HUMAN FACTORS ASSURANCE APPROACH FOR HUMAN-AI TEAMING CONCEPTS

Brand new approaches or techniques may well be required during the next decade to deal with advanced Human-AI Teaming concepts, but currently we simply do not know what these new techniques will be. Therefore, a sensible place to start is to adopt an existing state-of-the-art framework, and to tailor and upgrade it to focus on the particular aspects of Human-AI Teaming that differ from today's systems. The recently launched HURID (Human Risk-Informed Design) platform (see Figure 2) from the European Horizon 2020 SAFEMODE Research Project (https://safemodeproject.eu/EhuridIndex.aspx) serves as a contemporary and comprehensive methodological toolkit for current and future systems validation.

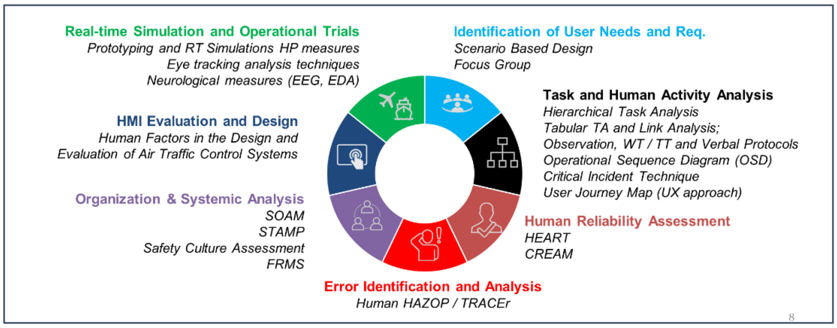

Figure 2: HURID human factors assurance toolkit (SAFEMODE project).

Figure 3 shows how this comprehensive toolkit has been adapted to focus on Human-AI Teaming concepts. Although it shows a 'once-through' flow of activities, in reality there are iterations and feedback loops as required by the design process. This paper focuses on the first three stages in the process: design requirements coming from Human Factors knowledge and guidance; analysis of the task to create a 'blueprint' of Human-AI interactions in time; and an analysis of the resilience of the system towards known failure possibilities, whether human, hardware, software or environmental in nature.

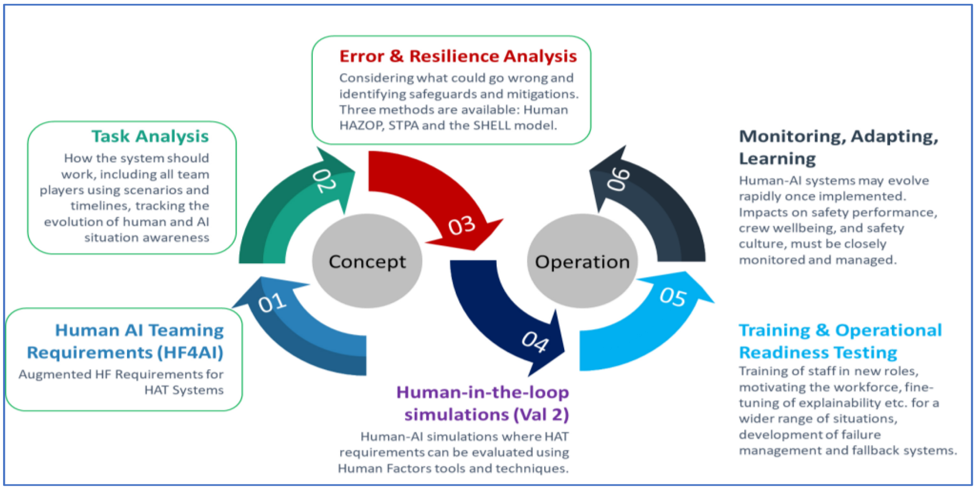

Figure 3: Human factors assurance process for human-AI teaming.

The assurance approach is led by a Human Factors practitioner, working with the Design Owner (the client), one or more operational users who have been able to explore / use the concept in a simulation or mock-up, and one or more of the AI developers, as illustrated in Figure 4. This 'collegial' approach helps ensure that the right expertise is around the table when doing analyses and determining whether design requirements are satisfied, or new ones are needed.

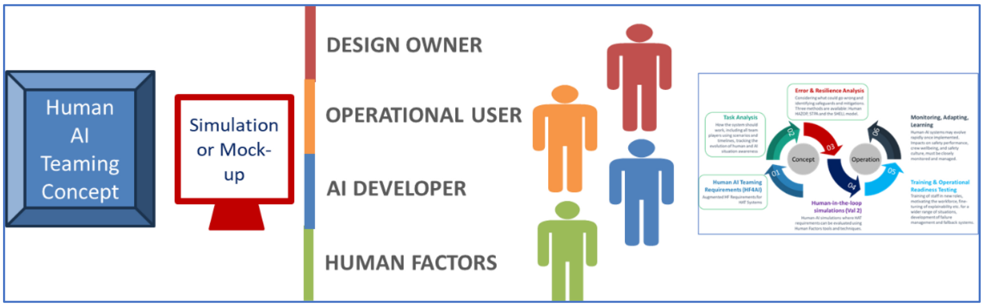

Figure 4: Key participants in human factors assurance process for human-AI teaming.

AN ATC TOWER HUMAN-AI TEAMING USE CASE

One of the use cases, for an air traffic control tower, is used to illustrate the Human Factors approach. In this use case, an Intelligent Sequence Assistant (ISA) is being developed to support and enhance decision-making for Air Traffic Controllers. ISA optimises runway utilisation in single-runway airports, providing real-time sequence suggestions for arriving and departing aircraft. For example, in Figure 5a, the original planned sequence of use for the runway is that the RYR lands first, then the KLM takes off, and then the BAW lands.

ISA computes the ordered sequence of aircraft that will use the runway. The order is displayed on the tower controller's Human Machine Interface (HMI) via numbers placed on the electronic strips of each aircraft (e.g. Figure 5b) in the controller's bay management area. If an event (e.g. the BAW accelerates faster than expected) triggers a resequencing, ISA updates the sequence in real-time, and the results are displayed on the HMI to the controller. ISA also provides explainability on-demand. For example, in Figure 5c, ISA signals that the take-off 'window' for the KLM is now too small (due to the BAW's increased speed), and so the BAW will land prior to the KLM take-off and takes position '2' in the strip.

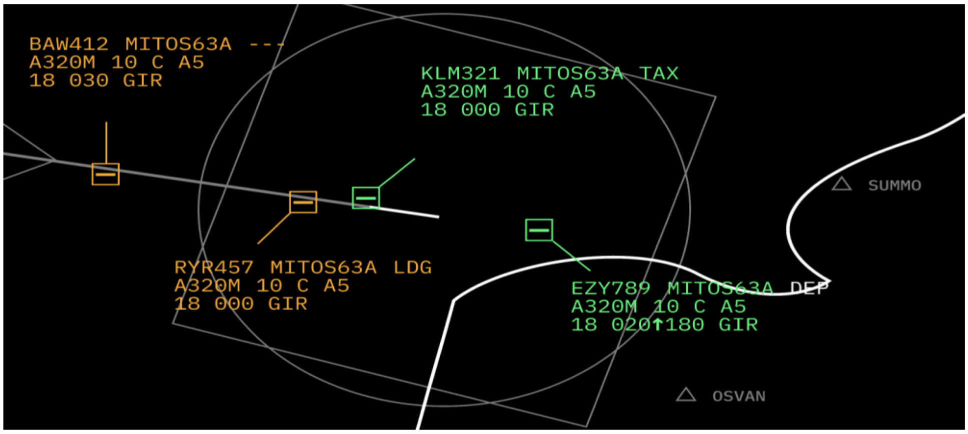

Figure 5a: Extract of tower controller's HMI showing aircraft positions in real time.

Figure 5b: Example of electronic strip with sequence on the left.

Figure 5c: Electronic strip with explainability after a resequencing process.

The real-time assistance provided by ISA ensures timely and accurate forecast updates, allowing Tower Air Traffic Controllers (ATCOs) to manage the traffic flow more efficiently, with more ‘look-ahead time’ as ISA can see further ‘upstream’. The expected benefits are improved decision-making, enhanced runway utilisation, increased operational efficiency, and a safer and more streamlined air traffic flow.

APPLICATION OF THE HUMAN FACTORS ASSURANCE APPROACH TO THE USE CASE

Human-AI Teaming Requirements Analysis

A framework for Human Factors Assurance used frequently in European Air Traffic Management is the SESAR Human Performance Assessment Process (SESAR-HPAP)[1], which has a comprehensive set of requirements in four over-arching areas:

- human limitations and capabilities,
- the human-machine interface,
- teamwork and communication,
- and transition from design into operation.

Based on a literature review of Human Factors and these requirements have been augmented to evaluate Human-AI Teaming scenarios and design projects, including additional questions related to trust, explainability and AI failure, etc. which are more specific to the use of AI. In total 160 questions are posed to design projects, in order to identify Human-AI performance improvement in nine areas, as illustrated in Figure 6.

The designer, operational user, AI developer and Human Factors practitioner (the Review Team) consider whether the Human-AI Teaming Concept of Operations (Conops) satisfies each of the requirements. Sometimes they do, sometimes the requirement may be judged irrelevant for this Conops and other times it may be too soon to judge and so must be returned to later when the design is more mature or when it is being tested in a validation exercise, for example. In all other cases, a design change needs to occur to satisfy the requirement. It is such design alterations that justify the exercise and show the added value of applying the approach. Some examples are given in Table 1 below for the ISA use case.

[1] https://ec.europa.eu/research/participants/data/ref/h2020/other/guides_for_applicants/h2020-sesar-er4-26-2019_hp-v1-3_16.06.05_en.pdf

Figure 6: Human factors requirement areas for human-AI teaming.

Table 1. Extract of human-AI requirements 'dialogue' within the review team.

Design of the human-machine interface supports the human in carrying out their tasks		
Is all the required information presented to the user, in an uncluttered way?	Y	Yes, though some of the XAI levels were unclear to the users in terms of function and information organisation
Is the interaction medium used appropriate for the task and task context, e.g. keyboard, mouse/trackpad, touchscreen, voice (NLP), and even gesture recognition?	Y	Yes, it is just the same as the existing airport's current context
Do visual, tactile, and oral/auditory displays, controls and interaction media comply with detailed Human Factors guidance for such devices (e.g. colour coding, luminance, auditory range etc.)?	Y/N	Colour coding is based on the airport's current workstation, and the magenta colour was chosen to differentiate ISA from all other interface components. Though, it needs to be tested in terms of contrast, luminance etc. An aural component has yet to be considered for design (and may never be, if not needed)
Is the placement of a new AI screen or control/interaction system consistent with the operational workplace layout (e.g. cockpit) such that it supports rather than hinders or interferes with critical operations?	Y	Yes, it is in the same screen ATCOs use now. We must make sure that panels such as XAI don't hinder ATCO's views.
Is it ensured that any alerts, warnings or time-sensitive messages provided by the AI gain and direct the human's attention (without startling)?	Y	Yes, the alert is pretty clear.
Do any alerts / warnings given signify their priority and time-urgency?	Y	Yes, the blinking signifies the time urgency.
Do alerts and warnings follow Human Factors guidance and principles?	Y/N	We need to test this in the next Validation round.
Has the fatigue impact of sustained HAT interaction been evaluated?	N/A	No sustained performance required.
Is it made clear to the human(s) when the alerting situation raised by the AI is resolved, or if actions are not resolving the threat?	Y	Yes, the sequence is changed, and arrows appear on the electronic strip to signal the change of sequence, though they may need to be more apparent as some participants missed those during Val 1
Is the interaction with the AI seen as user-friendly?	Y/N	Pilots like several aspects but more work required (see above)

Task Analysis: Operations Sequence Diagrams

To create a 'blueprint' of Human-AI interaction, a task analysis format known as Operations Sequence Diagram (OSD) is applied, as it usefully shows what both the human and the AI think/calculate and do in real time as the scenario develops. An extract is shown in Figures 7a–7c (normally displayed on one line).

Time	Actual System State	Goal	Human1	Info sources (non-AI)
T0	Aircraft data suggest that an aircraft (BAW412) is going faster than expected and the sequence needs to be reshuffled.	Resequencing of the aircraft's order to relieve ATCO's workload.	ATCO is not aware of the possible changes at the moment.	Data from aircraft.

Figure 7a: OSD Part 1.

Operator believed system state	AI believed system state	AI solution	AI HMI	AI Rationale (XAI)
ATCO knows that the assistant will provide support if needed.	ATCO is unaware of the aircraft state.	AI will calculate the re-sequencing of the two aircraft.	The AI HMI remains the same at the moment displaying the sequence.	While resequencing is ongoing the explanation is being constructed.

Figure 7b: OSD Part 2.

H-AI Dialogue	Authority gradient	Decision / Action	Human Performance Impact:
ISA's resequencing will start automatically. The ATCO supervises and makes decisions based on these suggestions.	ATCO is in control supervising the state of the system and the sequence.	ATCO is unaware of need for resequencing at this time.	No negative impact on Human Factors in this scenario. The AI is simply monitoring the aircrafts and the ATCO is unaware.

Figure 7c: OSD Part 3.

The OSD shows the flow of information between human and AI and can highlight where something needs to be signalled to the human user, or where the human and AI may be ‘out of synch’. It gives context when carrying out the requirements analysis, and when evaluating resilience of the Human-AI Teaming Conops.

Resilience Analysis: Application of the SHELL Method

The Methodology in Figure 3 shows three candidates for evaluating the resilience of a Human-AI Teaming Conops: SHELL, HAZOP and STPA. In HAIKU, all three are being applied to various use cases, and for the ISA example both HAZOP and SHELL have been applied, with the SHELL method shown here.

The SHELL model, developed by Hawkins (1982), examines five key elements that interact to influence human performance: software (procedures, protocols), hardware (equipment design), environment (contextual factors surrounding human activity), liveware (human characteristics), and their interconnectedness. The analysis of these elements and their interactions reveals vulnerabilities that may lead to accidents. The challenge lies in pinpointing these vulnerabilities within the expanding operational landscape of aviation. This is where the integration with the OSD (see Figure 7a, b, c above) proves valuable, providing a structured map of the system's operational sequence. Integrating each operational step in the OSD with the SHELL analysis significantly reduces complexity in hazard identification. Examples of guidelines gained from the SHELL analysis are given below.

[SOFTWARE] Operational Resequencing

What Happens When There Is a Sequence Change During Nominal Operations? What Is the Dialogue Between the Human and the AI?

In this scenario, the Software Element of the SHELL model is explored, focusing on the sequencing/resequencing procedure. An aircraft suddenly picks up speed, leading to resequencing. Both compliant (ATCO follows ISA suggestions) and non-compliant (ATCO triggers resequencing due to disagreement) situations are explored. This scenario highlighted the need for ISA to adapt to the controller actions, and vice versa the need for the controller to maintain situation awareness about ISA.

[HARDWARE] Connection Error

What Happens to the Sequence When There Is a Technical Failure? What Is the Dialogue Between the Human and the AI?

This scenario investigates the consequences of a technical failure in the hardware component of the system, specifically the connection between the AI system and the HMI. This failure leads to sequence changes not being transmitted to the interface: the calculation is not updated, and the controller is not updated about a change in the sequence. This scenario emphasises the critical need for reliable data transmission between ISA and the HMI, and for ways to allow the controller to always acknowledge the status of the system and its connection. It also highlights that controllers can never be too over-reliant on the system.

[ENVIRONMENT] Handling Unexpected Traffic

What Happens When an External Factor Influences the Sequence? What Is the Dialogue Between the Human and the AI?

The scenario explores how ISA deals with an external environmental factor outside of the ATCO's control, such as unexpected traffic (e.g., unannounced Flight School operations). When an unforeseen operation starts and new aircraft appear on the radar, they don't come with an auto-generated electronic strip and are not automatically inserted into the sequence. This implies that controllers need to manually create a strip and they somehow need to integrate it into the current sequence. Thus, this OSD highlighted the need for ISA to adapt to unforeseen operations by having a function that automatically organises strips, even when added manually by controllers.

[LIVEWARE] Handover-Takeover Procedure

What Happens If There Is a Resequencing Process During the Handover Event?

This OSD studies how events unfold during a handover/takeover procedure (when a new controller comes to the sector and replaces one of the controllers in position). This event falls in the Liveware category as is about human-human interactions. In this case, the incoming controller must create a mental map immediately. This OSD highlighted the need to add a short briefing about ISA in the checklist of the handover procedure, and in general, the need for ATCOs to be able to immediately understand the current status of the sequence, even if they had just sat down.

The SHELL analysis, together with OSDs for each scenario, usefully brought to life several key scenarios that required design considerations, which could be added to those found via the Human Factors Requirements Analysis.

CONCLUSION

The approach outlined in this paper has been applied to two of the HAIKU use cases so far (the other one was cockpit-based) and has led to design changes and considerations welcomed by the design team and end users. The approach itself is relatively light, with each analysis taking typically 1-5 days, plus some preparation. The design maturity does not have to be high, making the approach useful for early consideration of requirements, though it does work best if there is a simulation or mock-up that licensed operational personnel can use and test to see how it works. The approach will be applied to two further HAIKU use cases later in the year. The use of a review team approach can also track whether the operational users feel their role and motivation might be affected by the introduction of the AI into their workspace, as an over-riding goal of HAIKU is to support human-centric AI.

The approach appears fit for purpose, but as AI develops and advances, it is probable that new techniques, new methodological frameworks and even paradigms may be required. Until then, the approach allows a useful and productive dialogue between designer, operational user, AI developer and Human Factors performance experts, helping ensure that new AI systems can be integrated safely and successfully.

ACKNOWLEDGMENT

This publication is based on work performed in the HAIKU Project which has received funding from the European Union's Horizon Europe research and innovation programme, under Grant Agreement no 101075332. Any dissemination reflects the authors' views only, and the European Commission is not responsible for any use that may be made of information it contains.

REFERENCES

EASA (2023) EASA Concept Paper: first usable guidance for level 1 & 2 machine learning applications. February. https://www.easa.europa.eu/en/newsroom-and-events/news/easa-artificial-intelligence-roadmap-20-published

Hawkins, F. H. (1993). Human Factors in Flight (H. W. Orlady, Ed.) (2nd ed.). Routledge. https://doi.org/10.4324/9781351218580