Human-AI Hybrid Paradigm for Collaborative Air Traffic Management Systems

Duc-Thinh Pham[1], Hasnain Ali[1], Katherine Fennedy, Meng-Hsueh Hsieh, Sameer Alam, Vu Duong
Air Traffic Management Research Institute
Nanyang Technological University, Singapore
{dtpham, hasnain.ali, katherine.fennedy, menghsueh.hsieh, sameeralam, vu.duong}@ntu.edu.sg

***Abstract*—Even with the recent advancements in Artificial Intelligence (AI), incorporating AI-based systems into air traffic management (ATM) and air traffic control (ATC) poses significant challenges due to the extremely low tolerance for errors in ATM systems. Therefore, we propose the adoption of a novel human-AI hybrid (HAH) paradigm in ATM, emphasizing the collaborative aspect and high safety standards of human-AI interaction. In contrast to the substitution and augmentation concepts discussed within the human-AI teaming paradigm, which conveys that the roles of AI and humans are partitioned, we prefer the HAH paradigm, in which human and AI systems collaborate as integrated units to accomplish tasks. Under the HAH paradigm, ATM can significantly benefit from the complementary blend of ATCO judgment, intuition, and adaptability alongside the perceptual competencies, computational prowess, and tireless attention to detail that AI can offer. Some critical elements and design principles of HAH are investigated, and an example of HAH in air traffic conflict resolution, a typical human-centric and safety-critical task, is also presented for discussion. These contributions are fundamental prerequisites for successfully introducing HAH into ATM/ATC and will help create a framework for better understanding and supporting the effective use of AI systems for ATM/ATC.**

***Keywords*—Human-AI Hybrid, Human-AI Teaming, Air Traffic Management, Air Traffic Control, AI-based approach.**

I. Introduction

Air traffic management (ATM) and Air Traffic Control (ATC) are quintessential examples of safety-critical systems due to their direct impact on the safety of thousands of passengers and crew members daily. The system's failure can lead to catastrophic consequences, underscoring the paramount importance of maintaining rigorous safety standards and protocols. Furthermore, the complexity of air traffic management arises from intricate interactions between air traffic controllers (ATCOs), evolving technological systems, and varying environmental factors. These interactions usually require real-time decision-making under immense pressure and often with incomplete information, which is characteristic of many safety-critical domains.

On the other hand, over the past decade, Artificial Intelligence (AI) has made significant strides, surpassing human performance on benchmarks like image classification in 2015, reading comprehension in 2017, visual reasoning in 2020, and natural language inference in 2021 [1]. With AI and Machine Learning (ML) emerging as promising tools to optimize decision-making processes, their integration into ATM has become a topic of paramount importance and keen interest. This has also been recognized by international organizations such as the International Civil Aviation Organisation (ICAO) [2], the Federal Aviation Administration (FAA) [3], SESAR Joint Undertaking (SESAR JU) [4] and the European Union Aviation Safety Agency (EASA) [5]. Integrating AI into the highly regulated ATM presents substantial opportunities for multiple stakeholders. This includes scientists and researchers engaged in developing ML models, certification agencies responsible for ensuring adherence to safety standards, and the end-users, like ATCOs, who play a central role in implementing and relying on these technological advancements. Recently, EASA published the first set of technical objectives and organization provisions necessary to approve Level 1 AI applications (assistance to humans) [6]. EASA has also discussed the safety assessment and guidance for safety-related ML applications in terms of the initial safety assessment (during the design phase) and continuous safety assessment (based on operational data and in-service events) [7].

However, incorporating AI into this domain poses significant challenges due to the extremely low tolerance for errors in ATM systems. Many state-of-the-art AI systems, particularly based on deep learning models, can exhibit fuzzy and unpredictable behavior. More importantly, AI struggles with complex tasks that involve higher-order cognition and reasoning capabilities, like commonsense reasoning and planning [1], which are frequently encountered in time-critical and safety-sensitive scenarios. These tasks require a deeper understanding of semantics, relationships, and real-world knowledge to make inferences from visual/audio inputs, which remains an open challenge for AI systems [1]. On the other hand, humans possess an innate ability to perform commonsense reasoning and planning tasks with relative ease, leveraging their deep understanding of semantics, contextual relationships, and comprehensive real-world knowledge. This cognitive advantage over current AI systems becomes particularly crucial in time-critical and safety-sensitive scenarios frequently encountered in ATM/ATC. More recently, AI advancements have catalyzed the emergence of synergistic integration of human and AI capabilities. This evolution represents a paradigm shift, focusing not on replacing human ingenuity but on augmenting it through a complementary relationship with AI. Balancing AI capabilities with ATCO expertise and judgment is crucial

[1]Duc-Thinh Pham and Hasnain Ali contributed equally to this work.

to harnessing AI's potential advantages, such as improved efficiency and predictive capabilities, while ensuring safety and reliability standards.

Under the HAH paradigm, ATM can significantly benefit from the complementary blend of ATCO judgment, intuition, and adaptability alongside the perceptual competencies, computational prowess, and tireless attention to detail that AI can offer. By synergistically leveraging the strengths of ATCO and AI, these systems can enhance safety, improve operational efficiency, and reduce the likelihood of catastrophic failures. However, the successful implementation of HAH in ATM systems demands innovative approaches for human-centric design, dynamic adaptability, and human-AI integration and co-evolution while adhering to stringent safety requirements. As a use case, AI can handle repetitive, non-safety-critical tasks, such as routine communication and traffic flow analysis, allowing ATCOs to focus on critical decision-making and safety-critical scenarios. Additionally, HAH can improve real-time decision-making capabilities in which AI systems can aid controllers in planning and managing complex and congested air traffic scenarios. Furthermore, HAH can facilitate continuous monitoring and adaptation to changing environmental conditions. For instance, AI systems can constantly evaluate traffic patterns, weather conditions, and other variables, providing real-time updates and alerts to ATCOs. This enables a more proactive approach to managing air traffic, enhancing situational awareness, and enabling timely responses to emerging issues.

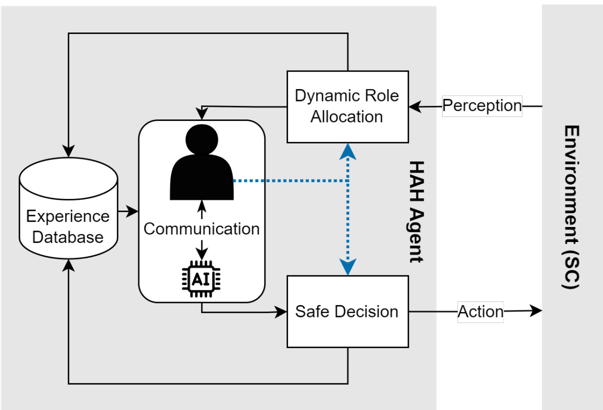

Figure 1. The concept diagram of HAH emphasizes human-centric design, human-AI integration, adaptation, co-evolution, and dependability. Within HAH agent, human and AI are unified unit with internal communication to maintain efficient shared situation awareness. The HAH agent firstly receives perception from the safety-critical environment and dynamic adjust the roles of human and AI based on human's guidance. Based on the allocated roles, the unified unit will process the perception to provide the safe decision. The data generated through the whole process is logged for improving human and AI performance through continuous learning and co-evolution.

This paper introduces a concept definition of HAH (depicted in Figure 1), which is essential for elucidating and accommodating the multifaceted landscape of HAH in ATM/ATC. Second, this paper examines the critical factors influencing HAH design and development. Finally, a potential application of HAH in air traffic conflict resolution, a typical human-centric and safety-critical task, is selected for discussion. These contributions are fundamental prerequisites for successfully introducing HAH into ATM/ATC and will help create a framework for better understanding and supporting the effective use of AI systems for ATM/ATC.

II. Human-AI Hybrid Paradigm

In ATM and other domains, various terms describe the interworking between human agents and AI-enabled systems, often interchangeably. This ambiguity arises from inconsistent terminology across disciplines and fragmented discussions in fields such as information systems, human-computer interaction, engineering, and management. Each definition encapsulates unique concepts with specific characteristics and requirements, making the selection of appropriate terminology essential for efficient design and development. For safety-critical systems like ATM/ATC, we argue that the HAH is particularly effective and must be precisely defined, designed, and developed.

A. *Definition of Human-AI Hybrid*

Definition: Human-AI Hybrid

HAH refers to a human-centric paradigm with emergent interaction and continuous adaption with shared knowledge between humans and AI systems as an integrated unit, leveraging their complementary strengths while mitigating weaknesses. This approach aims to enhance dependability in high-risk environments and maintain human authority in final decision-making while fostering co-evolution, adaptation, and trust within the overall system.

HAH systems in ATM must be dependable, human-centric, adaptable, evolving, and unified. The core characteristics of HAH definition are further elaborated below for better clarification of the discussed HAH definition.

- **Dependability**: Dependability is paramount in any AI system, including HAH systems, especially in safety-critical tasks, because it directly impacts operations' safety, reliability, and effectiveness [8]. A dependable system ensures that AI functions consistently and accurately under all conditions, even in the face of uncertainties, allowing ATCOs to trust their decisions and actions without constant oversight. It ensures that the human-AI partnership enhances, rather than endangers, safety.
- **Human-centric design**: A human-centric design in AI collaboration is essential for ensuring ethical decision-making, as it keeps ATCOs in control and accountable for AI-driven decisions. This approach enhances safety and reliability by prioritizing ATCO oversight, allowing ATCOs to monitor, intervene, and adapt AI systems to prevent risks and respond to changing conditions. It empowers individuals by augmenting rather than replacing ATCO abilities, ensuring that AI systems reflect and support ATCOs preferences and needs. This approach also

aligns with legal and regulatory requirements, ensuring compliance and fostering responsible AI development.

- **Adaptability**: The ability to adjust and optimize the division of roles, responsibilities, and control to ensure the highest levels of effectiveness and safety. It allows the system to respond to the unique demands of each situation, leveraging the strengths of both humans and AI [9]. The dynamic adaptation, on the one hand, ensures that AI handles routine, high-volume tasks, reducing ATCO cognitive load and allowing ATCOs to focus on more complex, critical decisions. On the other hand, it also enables rapid shifts in control during emergencies, where ATCO intuition and decision-making are paramount. This flexibility enhances safety by ensuring that the most capable entity—ATCO or AI—is in control at any given moment and improves overall system efficiency, responsiveness, and resilience in the face of evolving challenges.
- **Co-evolution**: Co-evolution, or continuous learning, is essential. It allows AI to adapt not only to changes in ATCO behavior and preferences but also requires human operators to evolve as AI capabilities improve, enabling them to leverage the increasingly sophisticated support that AI offers. This process necessitates both entities to dynamically update their respective mental models to maintain an up-to-date shared mental model. Through co-evolution, both humans and AI systems enhance their capabilities—humans develop new cognitive and operational skills, while AI refines its decision-making processes based on human input.
- **Human-AI integration**: Human-AI integration as a unified unit is crucial because it allows for the seamless blending of human intuition and ethical judgment with the computational power, precision, and speed of AI [10]. By functioning as a unified unit, this integration enables real-time adaptability and ensures that the strengths of both human and AI components are fully leveraged. In dynamic and complex environments like ATM, the ability of ATCOs and AI to operate as a single, unified system enhances decision-making, improves response times, and reduces the likelihood of errors. This deep integration also fosters a continuous exchange of information, enabling AI to learn from human input and vice versa, ultimately leading to more effective and efficient outcomes.
- **Trust**: Trust is a multifaceted concept that plays a pivotal role in ensuring HAH systems' smooth operation and adoption. Due to its unique characteristics, trust in a HAH system differs from traditional trust in human-to-human or human-to-technology relationships. Unlike static relationships, trust in HAH systems evolves based on the AI's performance, reliability, and ongoing human interactions. It is also highly context-dependent, meaning trust can fluctuate based on specific tasks or situations. One critical factor is the need for calibrated trust—humans must accurately perceive the AI's strengths and limitations. Miscalibrated trust can negatively impact system performance and safety, whether too high (leading to automation bias) or too low (resulting in under-utilization). Finally, trust in a HAH system is bi-directional—just as humans must trust the AI, the AI should also "trust" human inputs. This means that AI systems must incorporate human feedback, learn from operator corrections, and adapt their behavior accordingly. The interaction should be a collaborative, evolving process where both the human and the AI influence each other's decisions and actions.

B. Discussion on related terms of HAH in literature

Based on how humans and AI collaborate, these terms can be grouped into three general paradigms or definitions: Human-AI Collaboration (HAC) [11], Human-AI Teaming (HAT) [12], [13], and Human-AI Hybrids (HAH) [14]–[16]. Generally, both human-AI teaming and human-AI hybrids can be seen as subsets of human-AI collaboration. Furthermore, although human-AI teaming and human-AI hybrids share several characteristics, they possess distinct features that are important to differentiate. Their relationship is illustrated in Figure 2 and further discussed below.

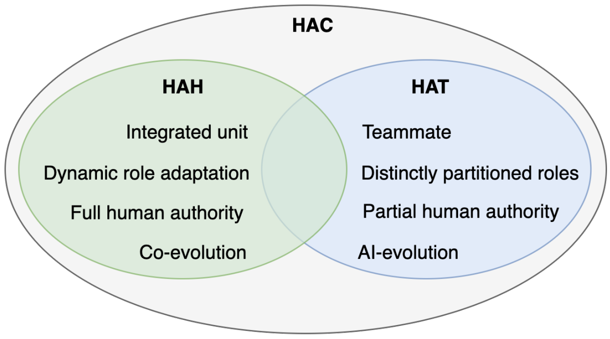

Figure 2. Illustration of the relationship between Human-AI Collaboration (HAC), Human-AI Teaming (HAT), and Human-AI Hybrids (HAH). Compared to HAT, HAH emphasizes human-AI integration, dynamic adaptation, human authority, and co-evolution.

Human-AI Collaboration is a widely discussed concept, broadly encompassing any cooperative interaction between humans and AI, where both contribute to a shared goal. The structure of this collaboration can vary, with roles and responsibilities shifting depending on the context and specific requirements. In some cases, humans oversee AI systems that act as aides, enhancing human capabilities through decision support. Alternatively, humans and AI can work together as equal teammates, leveraging their complementary strengths. AI might also serve as a moderator of human performance [17], ensuring that human actions remain within acceptable boundaries and reducing potential errors or biases. Overall, human-AI collaboration can range from substitution (where AI replaces humans) to augmentation (where humans and AI enhance each other) to assemblage (where AI and humans are dynamically integrated as a cohesive unit).

Another key term is **Human-AI Teaming**, which refers to scenarios where humans and AI work together as partners, each with distinct roles that enable mutual augmentation in task performance. In this model of collaboration, humans and

AI operate as separate entities that coordinate their efforts, typically with predefined roles—AI handles specific tasks while humans focus on others. The division of responsibilities is often fixed, with clear distinctions between the roles of AI and humans. For example, AI systems might manage repetitive tasks in air traffic control, while ATCOs handle more complex decision-making and oversee airspace traffic. Despite the clear role partitioning, there remains a level of interaction and communication between the two.

In contrast, the **Human-AI Hybrid (HAH)** is often seen as more effective in safety-critical systems like air traffic control, where seamless integration and real-time adjustments are crucial. Unlike in Human-AI Teaming, where roles are distinctly partitioned, the boundaries between human and AI roles in a hybrid system are fluid and dynamically adjusted [10]. The AI is not just a tool or partner but a vital part of a unified system where human and AI contributions are interdependent and complementary.

III. Design Principles for HAH systems

While HAH promises to outperform either humans or AI systems operating independently, merely combining human and AI capabilities within a team is insufficient. An effective HAH requires humans to (1) comprehend and anticipate AI behavior, (2) exert timely and appropriate control over the system, and (3) establish trust in the AI system [18], [19]. Researchers must, therefore, address these human-centric aspects alongside the question of optimal AI roles within HAH. Effective collaboration, adaptation, and trust are fundamental pillars of successful human-AI interaction, each playing a pivotal role in shaping teamwork, navigating dynamic environments, and building dependability. The list of design principles for HAH is presented in Figure 3 and further details are discussed below.

Human-Centric Design with Safety Priority: In ATM and ATC, the design must prioritize the safety of passengers, crew, and ground personnel. The system should be user-centric, ensuring that the AI aids ATCOs without increasing their workload or causing confusion. Safety protocols must be embedded into every aspect of the system. For instance, an AI system that assists air traffic controllers by suggesting optimal flight paths should include clear safety checks for potential conflicts (e.g., other aircraft, weather conditions) and present them alongside its recommendations. The AI might display a flight path suggestion, highlight potential conflicts, and suggest alternative actions or routes. Meanwhile, the critical role of human factors (HF) in safety-critical systems should be emphasized.

Transparency and Explainability with Traceability: The utilization of machine learning methods, notably deep learning (neural networks), in AI systems presents a significant challenge for humans to retain an accurate and contemporaneous understanding of the system's dynamic capabilities. As the AI system continuously learns and adapts, its decision-making processes and actions in any given situation become increasingly complex and dynamic, straining the ability of humans to keep pace. The requirements for transparency and explainability in AI solutions have increased to bridge this gap and compensate for the inevitable deficiencies in human mental models caused by the evolving nature of AI systems. These solutions should communicate the underlying logic and rationale driving the system's decisions, enabling humans to maintain a more accurate understanding of the AI system's decision-making processes over time. In dynamic, time-critical scenarios commonplace in safety-critical environments, real-time transparency becomes crucial for decision-making, while explanations primarily contribute to developing improved mental models for future scenarios. However, in situations with sufficient time for review and analysis, transparency and explainability can directly impact decision-making processes. For instance, if an AI system recommends diverting a flight due to predicted weather conditions, it should explain the data sources (e.g., weather radar, satellite data), the reasoning process, and how it weighs different factors. The system should log these decisions so that they can be reviewed in post-incident analyses or during routine audits.

User Control and Human-in-the-Loop (HITL): Human operators must maintain ultimate control over AI-assisted decisions [20]. HITL design ensures that human judgment is central, especially when quick and decisive action is required. In ATC, if an AI suggests re-routing an aircraft to avoid conflict, the human controller must be able to accept, reject, or modify the suggestion. The AI system should present its case but defer to the human operator's final decision. For example, if the AI recommends a reroute due to detected turbulence, the controller can override this if they have updated information or if the change would cause other issues. However, a significant concern arises in real-time safety-critical tasks when ATCOs become overly reliant on AI outputs. This over-reliance can lead to a phenomenon known as automation-induced complacency, where operators divert their attention elsewhere, potentially missing critical situations that demand immediate intervention. Humans should ideally maintain SA consistently or, at minimum, be able to intervene when necessary promptly. Mechanisms fostering heightened engagement, such as task divisions ensuring individuals retain meaningful roles and expertise, are crucial, along with the ability to exert meaningful control over operations [21]. Effective strategies for transferring control to humans, considering their need to re-establish SA, are imperative.

Adaptability and Context Awareness: The HAH system must adapt to different operational contexts and recognize and respond to changing situations. This adaptability is critical in environments like ATC, where conditions can shift rapidly. During an unexpected event, such as an aircraft losing communication, the HAH system could adapt by prioritizing this aircraft in its monitoring and providing the ATCO with specific protocols and suggestions based on historical data of similar incidents. The AI should also adjust its behavior based on the time of day, traffic density, or even the specific experience level of the ATCO on duty.

Reliability, Robustness, and Fail-Safes: In safety-critical systems, AI must be exceptionally reliable and robust. There

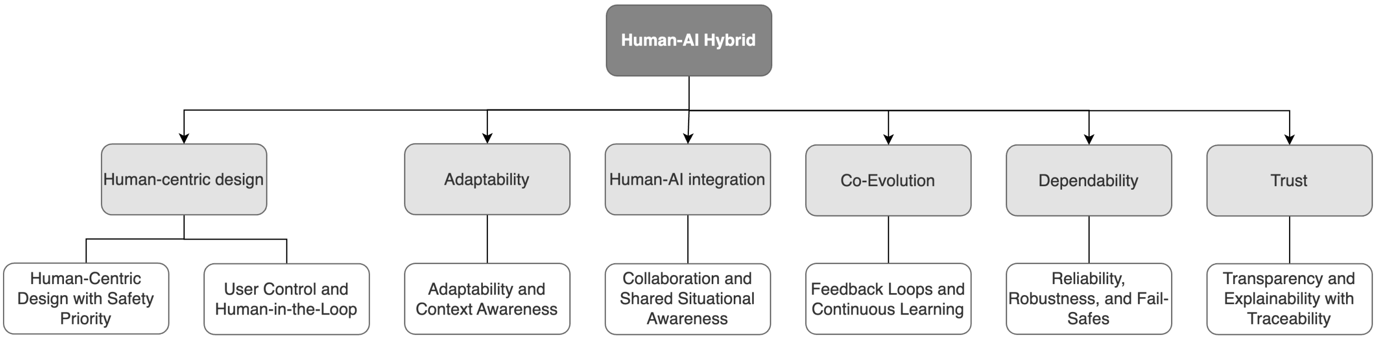

Figure 3. The list of discussed design principles for HAH and how it is related to the mentioned HAH definition

should be built-in fail-safes and redundancy mechanisms to ensure the system can handle unexpected inputs or failures without compromising safety. For instance, an AI system used for conflict detection in ATC should have multiple layers of redundancy. If the primary AI system fails, a backup system should immediately take over, using a different algorithm or data source to ensure continuous operation. Additionally, there should be manual override options that allow controllers to switch to traditional, non-AI-based methods if needed.

Feedback Loops and Continuous Learning: Continuous improvement is essential in safety-critical systems. A major focus lies in developing AI for effective HAH [12], [22] by transitioning towards bidirectional information exchange. AI systems must actively observe and analyze human performance, aligning their capabilities with the goals and objectives of their human counterparts. Crucially, this collaborative approach requires a closed-loop feedback system, where both objective system performance metrics and subjective human feedback are systematically integrated back into the AI system. This continuous cycle of learning and adaptation empowers AI to provide proactive, contextualized support tailored to human needs. Whether suggesting relevant tasks, sharing contextual information, or offering personalized recommendations, the AI must be designed to act as an intelligent partner, augmenting human capabilities through timely, goal-oriented assistance. Moreover, HAH systems must expose their biases and learning limitations—stemming from the generalizability constraints or non-causal patterns in the training data—to humans. As AI may encounter situations beyond its capabilities, it's crucial to identify and communicate the system's limits and biases to end-users.

Collaboration and Shared Situational Awareness: Effective collaboration between humans and AI in safety-critical environments requires shared situational awareness [23]. The AI system should present information in a way that is easily understood by human operators and facilitates teamwork. For instance, in an emergency when an aircraft is experiencing engine failure, the AI system could provide ATCO with real-time analysis of the situation, including recommended descent trajectories and their possible outcomes [24]. The AI could also facilitate communication between different team members, such as the controller, pilots, and emergency responders, ensuring everyone has the same understanding of the situation.

IV. Air Conflict Resolution under HAH paradigm

A. *Related Work on AI-based for Conflict Resolution*

Air traffic conflict resolution is a dynamic, time-sensitive, and safety-critical aspect of air traffic control. It involves a complex interaction of humans, machines, and procedures. In current sector-based operations, where the airspace is subdivided into smaller geographical regions, ensuring safe separation between aircraft by resolving potential conflicts and maintaining an efficient traffic flow is the primary responsibility of air traffic controllers (ATCOs). This task is human-centric and safety-critical, demanding high cognitive effort.

Over time, ATCOs develop certain inherent preferences in managing air traffic conflicts, known as 'conflict resolution strategies.' These strategies help ATCOs manage air traffic while preserving their cognitive resources. However, existing AI methods [25], [26], typically do not incorporate ATCOs' conflict resolution strategies, resulting in low acceptance of these methods due to lack of conformance between ATCOs' perceptions of conflict resolutions and the advisories provided by automation tools. Recent research [27], [28] has demonstrated the advantage of using ATCOs' conflict resolution data to develop learning-based models for ATCO-conformal conflict resolution. These approaches aim to learn ATCO's mental model and mimic the ATCO's decision for given conflict scenarios (refer to Figure 4). Therefore, ATCOs can better understand and quickly validate the AI's resolution, leading to lower workload and higher ATCO acceptance.

While these conformal models encapsulate ATCOs' conflict resolution strategies and actions, they may also incorporate potential inefficiencies, such as excessive aircraft vectoring during maneuvers or higher separation buffers during conflict resolution. It is beneficial to devise a method that enables the model to adjust its balance between optimal and conformity dynamically. As discussed in [29], while ATCOs generally preferred conformal resolutions to resolve conflicts, there is also an inclination towards balanced conflict resolutions, which improve maneuver efficiency. ATCO feedback indicated that such a conflict resolution advisory mechanism could enhance controller decision-making. Those intelligent conflict resolution tools are consistently considered recommender systems

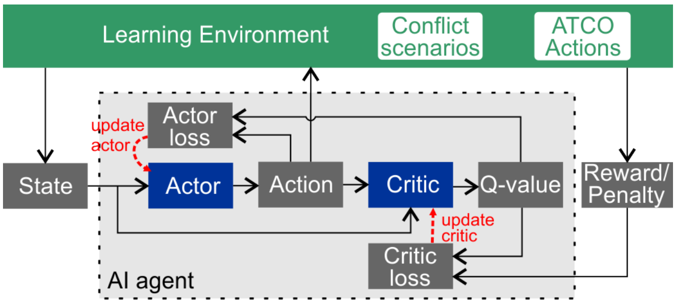

Figure 4. The concept diagram for learning air traffic conflict resolutions using reinforcement learning approach, sourced from [27]. The AI agent learned the ATCO's preferences in conflict resolution by considering the historical ATCO's actions for the critic loss to guide the convergence of the agent policy. Therefore, the trained policy can reproduce ATCO's resolution for similar air traffic conflict scenarios.

when integrated into an operational setting, offering advisories that ATCOs can accept or reject. ATCOs remain the final decision-makers in all situations.

Furthermore, addressing the long-term future air traffic demand requires a paradigm shift from current concepts of operations. As such, alternative concepts like flow-centric operations (FCOs) are being explored [30]. This new concept involves managing air traffic from an aggregated perspective rather than focusing on individual flights (as in sector-based operations), based on the formation and evolution of major traffic flows [31]. As a result, ATCOs' responsibility shifts from managing all traffic within a given sector to overseeing a specified number of aircraft throughout their flight segment within an airspace. ATCOs handling a flow must ensure safe separations between the flows (inter-flow) as well as between the aircraft within the same flow (intra-flow) [32]. Given the complexity of these scenarios, AI is expected to be essential in future air traffic control, particularly in conflict resolution tasks.

Key insights from these studies emphasize the importance of involving ATCOs in the AI development. Their domain expertise has been crucial in tailoring AI solutions to fit the operational context, with iterative feedback helping refine AI functionalities to make them more intuitive and effective. This collaboration between human expertise and AI has the potential to significantly improve efficiency and safety in air traffic management. However, by contracting with the HAH, these approaches have limitations in considering the evolution of ATCO behaviors during the task operation. Moreover, as discussed, AI will have more autonomy and tasks for handling traffic, so the mechanism for real-time adjustment of the level of automation should also be investigated. Some considerations for further studies on conflict resolution under the HAH paradigm are discussed in the following section, which is expected to provide the reader a better idea of what a HAH conflict resolution should have for safely and efficiently working with ATCOs.

B. A HAH Approach to Air Traffic Conflict Resolution

A HAH paradigm facilitates seamless collaboration between ATCO and AI, combining human expertise with AI's computational strength to optimize safety and efficiency. This contrasts with traditional systems, where static roles and limited feedback mechanisms often result in a lack of alignment and trust between humans and AI. Below, we explore how a HAH approach can enhance air traffic conflict resolution and a visual summary is presented in Figure 5.

1) Personalized Conflict Resolution System: A Conflict Resolution system would be personalized and adapt to individual ATCO preferences and workflows, ensuring the human-AI interaction is as seamless as possible.

- **Leveraging ATCO Data for Conformal Conflict Resolution**: Each ATCO has a unique working style, decision-making process, and experience level. By collecting and analyzing past ATCO decisions, preferences, and behaviors, AI systems can personalize the conflict resolution process. This tailored approach would ensure that recommendations and conflict resolution strategies align with the operator's individual tendencies, potentially reducing the cognitive load and enhancing decision accuracy. For instance, if an ATCO consistently favors certain types of conflict resolutions (e.g., altitude adjustments over course changes), the AI could prioritize similar solutions in its suggestions.
- **Customized Interface for Visualized Information and Communication Modes**: The amount and type of information visualized in an air traffic management system can vary depending on the situation and the ATCO's cognitive workload. Some ATCOs may prefer a more minimalistic interface with only the most critical data points, while others might want a detailed overview of all variables. Similarly, communication modes (text, visual alerts, or verbal communication) can be customized to fit the ATCO's preferences, making it easier for them to respond quickly and effectively. This adaptability helps ensure that ATCOs are neither overwhelmed nor underinformed during conflict resolution.
- **Customized Roles/Tasks and Level of Automation for Dynamic Adaptation**: Incorporating the fact that different ATCO may perceive the same conflict scenario differently, with varying levels of risk tolerance/appetite, historical data of ATCO behaviors should be analyzed to construct the workload and air traffic complexity models for each for each ATCO. Then, a personalized system could support each ATCO to set their preferred task allocation and levels of automation in conflict resolution tasks. For example, one ATCO might want the AI to provide only suggestions, while another may be more comfortable with the AI executing certain resolutions autonomously under specific conditions. This flexibility allows for a smoother human-AI interaction, ensuring that the ATCO retains control over the decision-making process while benefiting from AI support.

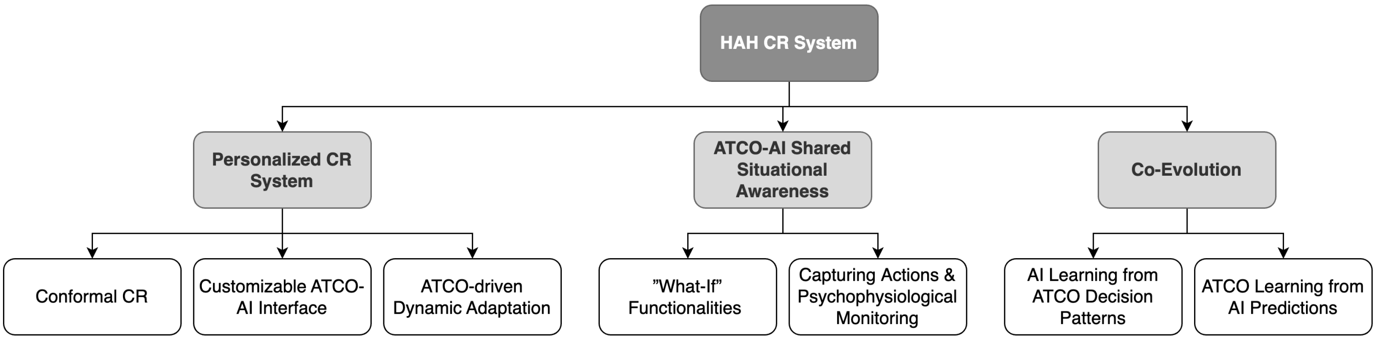

Figure 5. Considerations for design and development of HAH conflict resolution system.

2) ATCO-AI "Communication" for Shared Situational Awareness: Shared situational awareness is critical in ATC, especially in conflict resolution where split-second decisions must be made. More importantly, shared situational awareness allows dynamic adaptation, especially in emergencies, when autonomy is transferred back to ATCO for manual handling. Therefore, to maintain shared situational awareness, effective communication and information exchange between ATCOs and AI systems must be designed and developed.

- **Information flow from AI to ATCO - "What-If" Functionalities and Alternative Resolutions**: AI systems can offer "what-if" functionalities to help ATCOs assess the potential consequences of different actions. For example, if an ATCO is deciding whether to alter a plane's course or change its altitude, the AI can present the potential outcomes of each option, such as fuel consumption, impact on surrounding flights, or flight path safety. AI can also generate alternative resolutions, giving ATCOs a broader set of options to explore before selecting the best course of action. Since conflict resolution is time-critical, detailed explanations of AI decisions should be available only on demand, preventing information overload while offering deeper insights when needed. Besides, if the system detects unsafe behaviour from the ATCO, instead of 'policing' them, the HAH system could bring the risk to the ATCO's attention more sensitively [33] using polite, neutral language or subtle visual cues to gently alert ATCO without causing disruption or alarm.
- **Information flow from ATCO to AI - Capturing ATCO Actions and Psychophysiological Monitoring**: ATCOs influence the AI's decision-making by their actions and interactions with the system. By recording ATCO inputs (via keyboard, touchscreens, or voice commands), AI can continuously learn from these decisions and adapt its behavior to align more closely with ATCO's style. Additionally, monitoring ATCO psychophysiological data (such as eye-tracking or voice stress analysis) allows AI to assess the operator's real-time stress levels and workload. In high-stress situations, the AI could reduce non-critical alerts or take on more autonomous tasks to ease the ATCO's cognitive load. Or when AI detects that the ATCO's attention is not focused, AI can return autonomy to the ATCO to help improve their concentration level.

3) Co-Evolution and Enhanced Decision-Making: Traditional conflict resolution systems often limit human-AI interaction to predefined roles and tasks. AI provides recommendations or automation without adapting to the human's evolving needs. However, A co-evolutionary approach supports mutual learning, where ATCOs and AI systems continuously adapt to each other's behavior. This feedback loop improves decision-making over time as the AI learns from the ATCO's strategies and the ATCO gains deeper insights into how AI can assist in complex scenarios.

- **AI Learning from ATCO Decision Patterns**: In a co-evolutionary system, the AI continuously learns from the ATCO's real-time choices and interactions. For example, if an ATCO regularly overrides AI suggestions to prioritize safety in specific situations or to account for unique weather conditions, the AI can adjust its future decision-making process to reflect these nuanced judgments. This allows the AI to develop a model of conflict resolution that is informed not only by preset rules but also by human expertise, leading to more relevant and actionable recommendations.
- **ATCO Learning from AI Predictions**: AI can predict outcomes and simulate "what-if" scenarios more rapidly than an ATCO can process manually. For instance, the AI might calculate the future trajectories of multiple aircraft in conflict, showing the ATCO the potential impact of various decisions (such as changes in altitude or course). As the ATCO interacts with this data and sees the outcomes, they learn how to interpret and trust the AI's predictions more efficiently. This leads to faster, more informed decision-making and more effective collaboration between humans and machines.

In summary, future research directions for conflict resolution within the HAH paradigm should explore how to personalize CDR tools for individual ATCOs. This involves incorporating shared situation awareness and co-evolution to uphold system safety and efficiency. Based on the preceding discussion, several research topics related to human factors, algorithms, user interfaces, and system design are summarized in Table I.

TABLE I. Potential research topics for developing a HAH CR tool encompassing human factors, algorithms, user interfaces, and system design.

Research Topic	Description
Continuous Learning for Conformal Conflict Resolution	AI continuously adapts to ATCO preferences by updating its model based on recorded data, maintaining alignment with evolving ATCO behaviors. Model updates should occur based on discrepancies between ATCO and AI decisions, with real-time monitoring.
Predictability and Exploration of AI Agent for Enhanced Situation Awareness	AI systems can provide "what-if" functionality, enabling ATCOs to explore alternative resolutions and consider potential consequences before selecting the most appropriate solution.
Customized Interface for Visualized Information and Communication Modes	Includes on-demand explanations to improve transparency and usability.
Real-time ATCO Stress Levels and Workload Estimation	Real-time psychophysiological data can be used to assess ATCO workload, allowing for adaptive system responses.
Dynamic Task Allocation	Investigation of task decomposition and dynamic allocation between ATCO and AI, including transitions between various allocation strategies.
Fail-safe Mechanism for CDR	A fail-safe mechanism is required to monitor AI performance in real-time, alerting and allowing intervention as necessary. This mechanism relies on shared situation awareness and dynamic task allocation to ensure safe adjustments of AI tasks.

V. Conclusion

As AI continues to advance, there is growing interest in integrating it into ATM/ATC, where failures can have severe consequences. Leveraging HAH—a human-centric paradigm with emergent interaction and continuous adaptation with shared knowledge between humans and AI systems as an integrated unit—holds immense promise for real-time decision-making in ATM/ATC, as it capitalizes on the strengths of both humans and AI while mitigating their respective limitations. This study provides a comprehensive framework for understanding and supporting the effective use of AI systems in HAH for ATM/ATC. The insights and recommendations presented here can inform the development of more effective and reliable HAH systems, ultimately enhancing safety and performance in ATM/ATC. The ongoing journey towards integrating human and AI capabilities in ATM/ATC necessitates continued exploration and collaboration across disciplines. This collaboration will help address the remaining challenges and unlock the full potential of HAH in safeguarding lives and advancing the frontiers of ATM/ATC.

Acknowledgment

This research is supported by the National Research Foundation, Singapore, and the Civil Aviation Authority of Singapore, under the Aviation Transformation Programme. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore and the Civil Aviation Authority of Singapore.

References

[1] M. Nestor, F. Loredana, P. Raymond *et al.*, "The ai index 2024 annual report," *AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA*, 2024.

[2] ICAO, "Artificial intelligence," Available at https://www.icao.int/safety/Pages/Artificial-Intelligence-(AI).aspx, Accessed 10 Sept. 2024.

[3] FAA, "Roadmap for artificial intelligence safety assurance," Available at https://www.faa.gov/media/82891, 23 Jul. 2024.

[4] SESAR Joint Undertaking, "Ai in atm: transparency, explainability, conformance, situation awareness and trust," Available at https://www.sesarju.eu/sites/default/files/documents/reports/Sesar%20%20white%20paper%20AI%20in%20ATM.pdf, 2020.

[5] EASA, "Easa artificial intelligence roadmap," Available at https://www.easa.europa.eu/en/downloads/139582/en, Accessed 10 Sept. 2024.

[6] ——, "Concept of operations for safety-related artificial intelligence applications (easaconep)," 2023.

[7] ——, "Concept paper on operational aspects of a safety assessment of safety-related artificial intelligence applications (easaconep2)," 2024.

[8] J.-C. Laprie, "Dependability: A unifying concept for reliable computing and fault tolerance," *Dependability of resilient computers*, pp. 1–28, 1989.

[9] M. Johnson, J. M. Bradshaw, P. J. Feltovich, C. M. Jonker, M. B. Van Riemsdijk, and M. Sierhuis, "Coactive design: Designing support for interdependence in joint activity," *Journal of Human-Robot Interaction*, vol. 3, no. 1, pp. 43–69, 2014.

[10] A. Rai, P. Constantinides, and S. Sarker, "Next generation digital platforms: toward human-ai hybrids," *Mis Quarterly*, vol. 43, no. 1, pp. iii–ix, 2019.

[11] A. Berdahl, C. J. Torney, C. C. Ioannou, J. J. Faria, and I. D. Couzin, "Emergent sensing of complex environments by mobile animal groups," *Science*, vol. 339, no. 6119, pp. 574–576, 2013.

[12] N. J. McNeese, M. Demir, N. J. Cooke, and C. Myers, "Teaming with a synthetic teammate: Insights into human-autonomy teaming," *Human factors*, vol. 60, no. 2, pp. 262–273, 2018.

[13] S. Berretta, A. Tausch, G. Ontrup, B. Gilles, C. Peifer, and A. Kluge, "Defining human-ai teaming the human-centered way: a scoping review and network analysis," *Frontiers in Artificial Intelligence*, vol. 6, 2023.

[14] E. Kamar, "Directions in hybrid intelligence: Complementing ai systems with human intelligence." in *IJCAI*, 2016, pp. 4070–4073.

[15] D. Dellermann, P. Ebel, M. Söllner, and J. M. Leimeister, "Hybrid intelligence," *Business & Information Systems Engineering*, vol. 61, pp. 637–643, 2019.

[16] S. Wellsandt, K. Klein, K. Hribernik, M. Lewandowski, A. Bousdekis, G. Mentzas, and K.-D. Thoben, "Hybrid-augmented intelligence in predictive maintenance with digital intelligent assistants," *Annual Reviews in Control*, vol. 53, pp. 382–390, 2022.

[17] M. R. Endsley, "From here to autonomy: lessons learned from human–automation research," *Human factors*, vol. 59, no. 1, pp. 5–27, 2017.

[18] S. L. Brandt, J. Lachter, R. Russell, and R. J. Shively, "A human-autonomy teaming approach for a flight-following task," in *Advances in Neuroergonomics and Cognitive Engineering: Proceedings of the AHFE 2017 International Conference on Neuroergonomics and Cognitive Engineering, July 17–21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA 8*. Springer, 2018, pp. 12–22.

[19] R. J. Shively, J. Lachter, S. L. Brandt, M. Matessa, V. Battiste, and W. W. Johnson, "Why human-autonomy teaming?" in *Advances in Neuroergonomics and Cognitive Engineering: Proceedings of the AHFE 2017 International Conference on Neuroergonomics and Cognitive Engineering, July 17–21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA 8*. Springer, 2018, pp. 3–11.

[20] J. J. Bryson and A. Theodorou, "How society can maintain human-centric artificial intelligence," *Human-centered digitalization and services*, pp. 305–323, 2019.

[21] M. Boardman and F. Butcher, "An exploration of maintaining human control in ai enabled systems and the challenges of achieving it," in *Workshop on Big Data Challenge-Situation Awareness and Decision Support. Brussels: North Atlantic Treaty Organization Science and Technology Organization. Porton Down: Dstl Porton Down*, 2019.

[22] National Academies of Sciences, Engineering, and Medicine, "Human-ai teaming: State-of-the-art and research needs." Available at https://doi.org/10.17226/26355, 2022. [Online]. Available: https://doi.org/10.17226/26355

[23] G. L. Zacharias, *Autonomous horizons: the way forward*. Air University Press Maxwell Air Force Base, AL, 2019.

[24] D. Delahaye, S. Puechmorel, P. Tsiotras, and E. Féron, "Mathematical models for aircraft trajectory design: A survey," in *Air Traffic Management and Systems: Selected Papers of the 3rd ENRI International Workshop on ATM/CNS (EIWAC2013)*. Springer, 2014, pp. 205–247.

[25] J. Mollinga and H. van Hoof, "An autonomous free airspace en-route controller using deep reinforcement learning techniques," *arXiv preprint arXiv:2007.01599*, 2020.

[26] D.-T. Pham, P. N. Tran, S. Alam, V. Duong, and D. Delahaye, "Deep reinforcement learning based path stretch vector resolution in dense traffic with uncertainties," *Transportation research part C: emerging technologies*, vol. 135, p. 103463, 2022.

[27] P. N. Tran, D.-T. Pham, S. K. Goh, S. Alam, and V. Duong, "An interactive conflict solver for learning air traffic conflict resolutions," *Journal of Aerospace Information Systems*, vol. 17, no. 6, pp. 271–277, 2020.

[28] Y. Guleria, D.-T. Pham, S. Alam, P. N. Tran, and N. Durand, "Towards conformal automation in air traffic control: Learning conflict resolution strategies through behavior cloning," *Advanced Engineering Informatics*, vol. 59, p. 102273, 2024.

[29] Y. Guleria, D.-T. Pham, and S. Alam, "Advancing beyond conformal conflict resolution in air traffic control: Balancing efficiency and conformance," in *Proceedings of the 10th International Conference on Research in Air Transportation (ICRAT)*, 2024.

[30] A. S. B. Jumad, K. Tominaga, C. X. Yi, V. N. Duong, E. Itoh, and M. Schultz, "Flow-centric air traffic control: Human in the loop simulation experiment," in *2023 IEEE/AIAA 42nd Digital Avionics Systems Conference (DASC)*. IEEE, 2023, pp. 1–8.

[31] M. Schultz, K. Tominaga, E. Itoh, and V. N. Duong, "Introduction of moving sectors for flow-centric airspace management," 2023.

[32] Y. Guleria, D.-T. Pham, and S. Alam, "An agent-based approach for air traffic conflict resolution in a flow-centric airspace," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2023, pp. 5619–5626.

[33] B. Kirwan, "The impact of artificial intelligence on future aviation safety culture," *Future Transportation*, vol. 4, no. 2, pp. 349–379, 2024.