Human Factors in Design, Engineering, and Computing, Vol. 238, 2026, 1602–1611
https://doi.org/10.54941/ahfe1008388

Intelligent Design Architectures for Human-AI Teaming in Safety-Critical Operations

Debra Henneberry[1], Dimitrios Ziakkas[2], and Thanasis Binis[3]

[1]School of Aviation and Transportation Technology, Purdue University, West Lafayette, IN 47907, USA
[2]School of Future Transport Engineering, Coventry University, Coventry, CV1 5FB, UK
[3]Hellenic Civil Aviation Authority, Athens, Greece

ABSTRACT

Intelligent systems increasingly influence operational decisions in transportation, yet their integration often occurs without considering the cognitive characteristics of the human operator. This paper proposes an intelligent design architecture that combines AI-enabled analytics, digital twins, competency modelling and systems engineering within a human-centred framework. Rather than optimising algorithmic performance alone, the proposed architecture dynamically adapts information presentation, decision support and workload distribution according to human performance requirements. Aviation case studies illustrate how computational intelligence can enhance resilience, explainability and operational adaptability while maintaining meaningful human authority.

Keywords: Human-AI teaming, Adaptive automation, Human digital twin, Competency modelling, Explainable AI, Meaningful human control, Intelligent design, Aviation

INTRODUCTION

Intelligent systems are moving from background analytics into the decision loop of safety-critical transportation operations. In aviation, research demonstrators now generate emergency diversion plans, recognise runways from camera imagery and advise crews during taxi (Airbus, 2023), while ground-based agents are being designed to support reduced-crew operations (Lachter et al., 2017; Vu et al., 2018). These capabilities are typically engineered to maximise algorithmic performance, and the human operator is considered late, as a user of outputs rather than as a teammate whose cognition shapes how those outputs are interpreted and acted upon.

This design sequence carries well-documented risks. Operators who are kept out of the loop lose situation awareness (Endsley, 1995), static levels of automation mismatch the fluctuating demands of real operations (Kaber and Endsley, 2004), and opaque recommendations make it difficult to judge when an intelligent agent should be trusted (Chen et al., 2018). Adaptive systems that change their behaviour according to operator state and task context have long been proposed as a remedy (Feigh, Dorneich and Hayes, 2012), but their

Received September 25, 2026; Revised November 3, 2026; Accepted November 13, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

design has rarely been integrated with the competency frameworks that define what operators must be able to do, or with the systems engineering processes that govern certification. Our previous work on cognitive assistance, fatigue risk management and competency-based decision-making in aviation points to the same gap (Henneberry and Ziakkas, 2025; Henneberry et al., 2026a).

This paper proposes a Human-Centred Adaptive Intelligent Design Architecture (HAIDA) for Human-AI Teaming (HAT) in safety-critical operations. Three research questions (RQs) guided the work: which architectural components are required for intelligent systems to adapt to human performance requirements (RQ1); how adaptation can be designed to preserve meaningful human authority and explainability (RQ2); and what aviation cases reveal about the feasibility and limitations of such an architecture (RQ3).

BACKGROUND

Adaptive Automation and Operator Cognition

Human performance in dynamic control depends on perceiving, comprehending and projecting the state of the environment (Endsley, 1995), and on the availability of limited, partly separable attentional resources (Wickens, 2008). Rasmussen's (1983) distinction between skill-, rule- and knowledge-based behaviour further shows that the information an operator needs differs with the familiarity of the situation. Adaptive automation responds to these characteristics by changing the allocation of functions, the scheduling of tasks, the style of interaction or the content of displays according to triggers such as operator state, task load or environmental events (Feigh, Dorneich and Hayes, 2012). Physiological triggers were proposed in the 1990s (Byrne and Parasuraman, 1996), and experimental work shows that adaptive allocation can improve performance and situation awareness relative to static automation, provided that operators understand the adaptation (Kaber and Endsley, 2004).

Human-AI Teaming, Transparency and Human Authority

HAT reframes the intelligent system as a teammate whose actions must remain observable, predictable and directable. Coactive design emphasises the interdependence between human and machine activity and argues that support for this interdependence must be designed explicitly (Johnson et al., 2014). The situation awareness-based agent transparency model specifies three levels of information an agent should communicate: its current actions and plans, its reasoning, and its projected outcomes and uncertainty (Chen et al., 2018). Explainable AI research provides computational techniques for generating such explanations (Gunning and Aha, 2019; Arrieta et al., 2020), although explanations must be tailored to the operator's task and time constraints to be useful. Meaningful human control adds a normative requirement: system behaviour should track the reasons of the responsible humans and remain traceable to them (Santoni de Sio and van den Hoven, 2018). EASA's proposed Issue 03 of its AI Concept Paper extends guidance

to advanced automation, reinforcing the need to define the human role explicitly as autonomy increases (EASA, 2026).

Digital Twins, Competency Modelling and Systems Engineering

A digital twin is a virtual representation of a physical system that is continuously updated with operational data to support monitoring, prediction and decision-making (Glaessgen and Stargel, 2012; Grieves and Vickers, 2017; Tao et al., 2019). The human digital twin extends this concept to people, coupling real-time measurements of behaviour, physiology and environment with human performance models (Miller and Spatz, 2022). Competency modelling supplies the reference against which human performance is interpreted: evidence-based training defines observable behaviours for pilot competencies and uses operational data to prioritise training (IATA, 2013). Cognitive work analysis provides a complementary means of deriving information requirements from work domain constraints (Vicente, 1999). Model-based systems engineering links these elements into traceable requirements and verification evidence (Madni and Sievers, 2018), and digital twins are increasingly used within it to validate designs against operational scenarios (Madni, Madni and Lucero, 2019).

METHODOLOGY

The study followed a design science research approach, in which the contribution is an artefact whose utility is argued from theory and demonstrated against cases (Gregor and Hevner, 2013). The architecture was constructed through conceptual framework analysis, which identifies, categorises and integrates concepts from multidisciplinary literatures into a coherent structure (Jabareen, 2009). Concepts were drawn from adaptive automation, HAT, explainable AI, digital twin and competency research, and were grouped into architectural layers according to their function in sensing, modelling, adapting and assuring. The architecture was then examined against four aviation cases chosen to represent successive generations of intelligent aiding: early cognitive assistants for the flight deck, NASA's HAT ground station research, the Airbus DragonFly demonstrator, and a biometric data collection framework for collegiate flight training developed by two of the authors. Each case was coded against the architecture's layers to identify which functions it demonstrates and which remain unaddressed. Because one case derives from the authors' own work, claims were restricted to published evidence and cross-checked between authors. Table 1 summarises the method.

Table 1: Research method of the study.

Element	Design Decision	Rationale
Research paradigm	Design science research (Gregor and Hevner, 2013)	Contribution is an architecture whose utility must be argued and demonstrated
Framework construction	Conceptual framework analysis (Jabareen, 2009); functional layering	Integrates multidisciplinary concepts into traceable layers

(Continued)

Table 1: Continued.

Element	Design Decision	Rationale
Case selection	Four aviation cases across successive generations of intelligent aiding	Reveals persistent design gaps over time
Data sources	Peer-reviewed studies, technical reports, manufacturer publications	Publicly verifiable evidence
Analysis	Coding of each case against architecture layers; cross-case gap analysis	Identifies demonstrated functions and unresolved requirements
Quality criteria	Source and investigator triangulation; reflexivity	Credibility of interpretive judgements

THE HUMAN-CENTRED ADAPTIVE INTELLIGENT DESIGN ARCHITECTURE

HAIDA comprises five layers linked by a closed adaptation loop (Figure 1). The sensing layer integrates operational, aircraft system, environmental and human state data, including physiological and behavioural measures where their use is proportionate and consented. The digital twin layer maintains two coupled models: a system twin representing the aircraft and operational environment, and a human digital twin estimating operator workload, fatigue and situation awareness (Miller and Spatz, 2022). The competency model layer translates task demands into the competency-referenced performance requirements that the operator must meet at a given moment, drawing on evidence-based training behaviours (IATA, 2013) and work domain constraints (Vicente, 1999).

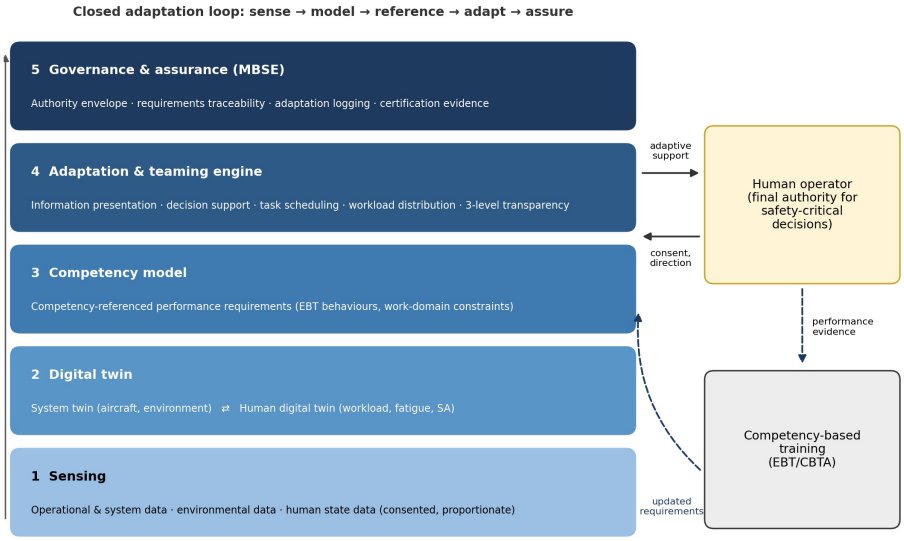


Figure 1: The Human-Centred Adaptive Intelligent Design Architecture (HAIDA).

The adaptation and teaming layer compares predicted human capacity with these requirements and selects among four adaptation types, namely information presentation, decision support, task scheduling and workload distribution (Feigh, Dorneich and Hayes, 2012). Its outputs are communicated through transparency at all three levels of the agent transparency model (Chen et al., 2018), so that operators can see not only what the system recommends but also why it has changed its behaviour. The governance and assurance layer, grounded in model-based systems engineering (Madni and Sievers, 2018), defines an authority envelope specifying which adaptations may occur automatically, which require operator consent, and which are prohibited, and records adaptation decisions to support explainability, investigation and certification. Operator performance and training outcomes feed back into the competency model, closing the loop between operations and training.

Five design principles govern the architecture. First, adaptation is referenced to competency-based performance requirements rather than to physiological load alone. Second, final authority for safety-critical decisions remains with the operator unless explicitly delegated. Third, the system explains its own adaptations as well as its recommendations. Fourth, degradation is graceful, with clear reversion to manual reasoning when data quality or model confidence falls. Fifth, operational data are fed back into training so that the human side of the team develops alongside the intelligent system.

CASE FINDINGS

The first case concerns early cognitive assistants. The Pilot's Associate programme developed a cooperative, knowledge-based system in which cooperating expert modules assessed the situation, planned responses and managed an adaptive pilot-vehicle interface informed by inferred pilot intent (Banks and Lizza, 1991). The Cockpit Assistant System (CASSY) modelled expected pilot behaviour to recognise intent and possible errors and communicated with the crew by speech dialogue, forming the basis for a theory of cognitive automation (Onken and Schulte, 2010). These systems demonstrated context-sensitive aiding decades ago, but lacked real-time human state sensing and a formal link to competency requirements.

The second case is NASA's HAT research for reduced-crew operations. Building on HAT tenets of transparency, bidirectional communication and operator-directed interaction (Shively et al., 2018), a ground station study compared a flight-following task with and without HAT features; participants preferred the HAT-enabled station and reported adequate situation awareness, reduced workload and efficient teaming with the automation (Brandt et al., 2018). The case demonstrates the value of the adaptation and transparency layers, while leaving human state modelling to future work.

The third case is the Airbus DragonFly demonstrator. Flight tests on an A350-1000 managed a simulated crew incapacitation, generated a new trajectory accounting for airspace, terrain and weather, communicated with air traffic control and the airline operations centre, and provided taxi assistance through obstacle alerts, assisted speed control and airport map

guidance (Airbus, 2023). The case shows advanced system-side intelligence and a clear trigger for delegating authority, but publicly available information does not describe how adaptation is calibrated to the capability of a remaining crew member.

The fourth case is a framework for collecting biometric and environmental data during collegiate flight training (Henneberry et al., 2026b), building on research into workload, acute stress, emotion and fatigue in training (Henneberry, Keller and Ziakkas, 2024). It provides the human sensing foundation for a human digital twin and connects measurement to competency-based instruction, but does not yet close the loop to real-time adaptation. Related work on direct voice input for extended minimum-crew operations illustrates how new interaction modes must be assessed for usability and safety before adaptive functions are layered upon them (Ziakkas, Harris and Pechlivanis, 2023). Table 2 summarises the findings.

Table 2: Findings: aviation cases mapped to the HAIDA layers.

Case	Layers Demonstrated	Contribution to Human-Centred Adaptation	Gap Relative to HAIDA
Pilot's Associate and CASSY cognitive assistants	Adaptation and teaming; partial competency modelling via expected-behaviour models	Intent inference and context-sensitive aiding	No real-time human state sensing; no formal competency link
NASA HAT ground station (reduced-crew operations)	Adaptation and teaming; transparency	Transparency, bidirectional communication and operator direction preferred by users	Human digital twin and adaptive triggers not yet integrated
Airbus DragonFly demonstrator	Sensing; system digital twin; governance of delegated authority	Automated diversion and taxi assistance supporting crews	Calibration of adaptation to remaining crew capability not publicly described
Collegiate flight training biometric framework	Sensing; human digital twin foundations; competency linkage	Human state measurement grounded in training practice	Real-time adaptation loop not yet closed

Across the cases, no single system implemented all five layers. Intelligence has advanced fastest on the system side, while human state modelling, competency referencing and explicit authority envelopes remain the least mature. The cases nevertheless show that each layer is technically feasible, suggesting that the principal challenge is integration rather than invention.

DISCUSSION

In response to RQ1, the architecture identifies five required components: human and system sensing, coupled system and human digital twins,

competency-referenced performance requirements, an adaptation and teaming engine, and a governance layer rooted in systems engineering. In response to RQ2, meaningful human authority is preserved through an explicit authority envelope, three-level transparency that extends to the system's own adaptations, and graceful reversion. In response to RQ3, the aviation cases show that the components exist in isolation but have not been integrated, with human-side modelling lagging behind system-side autonomy.

The principal theoretical contribution is the use of competency models as the reference signal for adaptation. Earlier adaptive automation research typically adapted to workload or physiological state (Byrne and Parasuraman, 1996; Kaber and Endsley, 2004); anchoring adaptation to competency-based requirements links moment-to-moment support to what operators are trained and assessed to do, and allows operational adaptation data to inform training. The architecture also operationalises meaningful human control (Santoni de Sio and van den Hoven, 2018) as an engineering artefact that can be specified, verified and audited. From a resilience perspective, it supports the capacity to extend performance when events challenge normal boundaries (Woods, 2015), because the team can reallocate work rather than rely on a fixed division of labour. For regulators such as civil aviation authorities, the governance layer offers a structure for evaluating adaptive systems: the authority envelope, adaptation logs and competency linkage correspond to evidence that can be examined during approval and oversight. Advanced air mobility, with its expected single-pilot and remotely supervised operations, is a natural first application (Ziakkas, Vafadar and Henneberry, 2026).

The study has limitations. The architecture is conceptual and has not been implemented or experimentally evaluated; the cases rely on published descriptions that omit design details; and human state sensing raises privacy, consent and data governance questions that require dedicated treatment. Future research should implement the adaptation loop in a flight simulator using the collegiate training dataset, compare competency-referenced and workload-referenced adaptation experimentally, and work with regulators to define acceptable authority envelopes for adaptive functions.

CONCLUSION

Intelligent systems will enhance safety-critical operations only if they are designed around the people who remain responsible for them. The Human-Centred Adaptive Intelligent Design Architecture combines AI-enabled analytics, digital twins, competency modelling and systems engineering so that information presentation, decision support and workload distribution adapt to human performance requirements rather than to algorithmic performance alone. The aviation cases show that the necessary components exist and that their integration, under an explicit authority envelope and transparent adaptation, offers a route to greater resilience, explainability and adaptability while maintaining meaningful human authority.

ACKNOWLEDGMENT

The authors thank colleagues at Purdue University, Coventry University and the Hellenic Civil Aviation Authority for discussions that informed this work. The views expressed are those of the authors and do not necessarily represent those of their institutions.

REFERENCES

Airbus (2023). Airbus tests new technologies to enhance pilot assistance. Press release, 12 January 2023. Toulouse: Airbus.

Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R. and Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, pp. 82–115.

Banks, S.B. and Lizza, C.S. (1991). Pilot's Associate: A cooperative, knowledge-based system application. IEEE Expert, 6(3), pp. 18–29.

Brandt, S.L., Lachter, J., Russell, R. and Shively, R.J. (2018). A human-autonomy teaming approach for a flight-following task. In: Baldwin, C. (ed.) Advances in Neuroergonomics and Cognitive Engineering. AHFE 2017. Advances in Intelligent Systems and Computing, 586. Cham: Springer, pp. 12–22.

Byrne, E.A. and Parasuraman, R. (1996). Psychophysiology and adaptive automation. Biological Psychology, 42(3), pp. 249–268.

Chen, J.Y.C., Lakhmani, S.G., Stowers, K., Selkowitz, A.R., Wright, J.L. and Barnes, M. (2018). Situation awareness-based agent transparency and human-autonomy teaming effectiveness. Theoretical Issues in Ergonomics Science, 19(3), pp. 259–282.

EASA (2026). EASA Artificial Intelligence Concept Paper, Proposed Issue 03. Cologne: European Union Aviation Safety Agency.

Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. Human Factors, 37(1), pp. 32–64.

Feigh, K.M., Dorneich, M.C. and Hayes, C.C. (2012). Toward a characterization of adaptive systems: A framework for researchers and system designers. Human Factors, 54(6), pp. 1008–1024.

Glaessgen, E. and Stargel, D. (2012). The digital twin paradigm for future NASA and U.S. Air Force vehicles. In: 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference. Honolulu, HI: AIAA.

Gregor, S. and Hevner, A.R. (2013). Positioning and presenting design science research for maximum impact. MIS Quarterly, 37(2), pp. 337–355.

Grieves, M. and Vickers, J. (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In: Kahlen, F.-J., Flumerfelt, S. and Alves, A. (eds.) Transdisciplinary Perspectives on Complex Systems. Cham: Springer, pp. 85–113.

Gunning, D. and Aha, D.W. (2019). DARPA's explainable artificial intelligence (XAI) program. AI Magazine, 40(2), pp. 44–58.

Henneberry, D., Keller, J. and Ziakkas, D. (2024). The role of mental workload, acute stress, emotion, and fatigue in aviation training. Paper presented at the International Conference on Cognitive Aircraft Systems (ICCAS), Toulouse, France, 16–17 May.

Henneberry, D. and Ziakkas, D. (2025). The role of artificial cognitive systems in the implementation of the aviation fatigue risk management systems. Paper presented at Intelligent Human Systems Integration (IHSI 2025), Rome, Italy, 24–26 February.

Henneberry, D., Ziakkas, D., Lekea, I. and Pechlivanis, K. (2026a). Harnessing emerging technologies to enhance decision-making in competency-based training and assessment. In: Intelligent Human Systems Integration (IHSI 2026): Disruptive and Innovative Technologies. AHFE International, 200.

Henneberry, D., Keller, J., Wilson, M., Vhaduri, S. and Ziakkas, D. (2026b). Creating a framework for the collection of biometric and environmental data during collegiate flight training. Paper presented at the 17th International Conference on Applied Human Factors and Ergonomics (AHFE 2026), Istanbul, Türkiye, 20–24 July.

IATA (2013). Evidence-Based Training Implementation Guide. Montréal: International Air Transport Association.

Jabareen, Y. (2009). Building a conceptual framework: Philosophy, definitions, and procedure. International Journal of Qualitative Methods, 8(4), pp. 49–62.

Johnson, M., Bradshaw, J.M., Feltovich, P.J., Jonker, C.M., van Riemsdijk, M.B. and Sierhuis, M. (2014). Coactive design: Designing support for interdependence in joint activity. Journal of Human-Robot Interaction, 3(1), pp. 43–69.

Kaber, D.B. and Endsley, M.R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. Theoretical Issues in Ergonomics Science, 5(2), pp. 113–153.

Lachter, J., Brandt, S.L., Battiste, V., Matessa, M. and Johnson, W.W. (2017). Enhanced ground support: Lessons from work on reduced crew operations. Cognition, Technology & Work, 19(2–3), pp. 279–288.

Madni, A.M., Madni, C.C. and Lucero, S.D. (2019). Leveraging digital twin technology in model-based systems engineering. Systems, 7(1), 7.

Madni, A.M. and Sievers, M. (2018). Model-based systems engineering: Motivation, current status, and research opportunities. Systems Engineering, 21(3), pp. 172–190.

Miller, M.E. and Spatz, E. (2022). A unified view of a human digital twin. Human-Intelligent Systems Integration, 4(1–2), pp. 23–33. https://doi.org/10.1007/s42454-022-00041-x

Onken, R. and Schulte, A. (2010). System-Ergonomic Design of Cognitive Automation: Dual-Mode Cognitive Design of Vehicle Guidance and Control Work Systems. Studies in Computational Intelligence, 235. Berlin: Springer.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. IEEE Transactions on Systems, Man, and Cybernetics, SMC-13(3), pp. 257–266.

Santoni de Sio, F. and van den Hoven, J. (2018). Meaningful human control over autonomous systems: A philosophical account. Frontiers in Robotics and AI, 5, 15.

Shively, R.J., Lachter, J., Brandt, S.L., Matessa, M., Battiste, V. and Johnson, W.W. (2018). Why human-autonomy teaming? In: Baldwin, C. (ed.) Advances in Neuroergonomics and Cognitive Engineering. AHFE 2017. Advances in Intelligent Systems and Computing, 586. Cham: Springer, pp. 3–11.

Tao, F., Zhang, H., Liu, A. and Nee, A.Y.C. (2019). Digital twin in industry: State-of-the-art. IEEE Transactions on Industrial Informatics, 15(4), pp. 2405–2415.

Vicente, K.J. (1999). Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work. Mahwah, NJ: Lawrence Erlbaum Associates.

Vu, K.-P.L., Lachter, J., Battiste, V. and Strybel, T.Z. (2018). Single pilot operations in domestic commercial aviation. Human Factors, 60(6), pp. 755–762.

Wickens, C.D. (2008). Multiple resources and mental workload. Human Factors, 50(3), pp. 449–455.

Woods, D.D. (2015). Four concepts for resilience and the implications for the future of resilience engineering. Reliability Engineering & System Safety, 141, pp. 5–9.

Ziakkas, D., Harris, D. and Pechlivanis, K. (2023). Towards eMCO/SiPO: A human factors efficacy, usability, and safety assessment for direct voice input (DVI) implementation in the flight deck. In: Harris, D. and Li, W.-C. (eds.) Engineering Psychology and Cognitive Ergonomics. HCII 2023. Lecture Notes in Computer Science, 14018. Cham: Springer. https://doi.org/10.1007/978-3-031-35389-5_15

Ziakkas, D., Vafadar, A. and Henneberry, D. (eds.) (2026). Advanced Air Mobility: Human Factors, Ecosystem Readiness, and the Integration of eVTOLs. Boca Raton, FL: CRC Press. https://doi.org/10.1201/9781003685555