

Article

Machine Learning and Cognitive Ergonomics in Air Traffic Management: Recent Developments and Considerations for Certification

Trevor Kistan ^{1,*}, Alessandro Gardi ² and Roberto Sabatini ²

¹ Technical Directorate, Thales Air Traffic Management, Melbourne, VIC 3001, Australia

² School of Engineering, Aerospace Engineering and Aviation, RMIT University, Bundoora, VIC 3083, Australia; alessandro.gardi@rmit.edu.au (A.G.); roberto.sabatini@rmit.edu.au (R.S.)

* Correspondence: trevor.kistan@thalesgroup.com.au; Tel.: +61-466-479-984

Received: 6 August 2018; Accepted: 26 September 2018; Published: 1 October 2018



Abstract: Resurgent interest in artificial intelligence (AI) techniques focused research attention on their application in aviation systems including air traffic management (ATM), air traffic flow management (ATFM), and unmanned aerial systems traffic management (UTM). By considering a novel cognitive human–machine interface (HMI), configured via machine learning, we examined the requirements for such techniques to be deployed operationally in an ATM system, exploring aspects of vendor verification, regulatory certification, and end-user acceptance. We conclude that research into related fields such as explainable AI (XAI) and computer-aided verification needs to keep pace with applied AI research in order to close the research gaps that could hinder operational deployment. Furthermore, we postulate that the increasing levels of automation and autonomy introduced by AI techniques will eventually subject ATM systems to certification requirements, and we propose a means by which ground-based ATM systems can be accommodated into the existing certification framework for aviation systems.

Keywords: cognitive HMI; machine learning; explainable artificial intelligence; computer-aided verification; air traffic management

1. Introduction

Several operational challenges underscore the need for increased automation to improve air navigation service provider (ANSP) productivity. By 2050, the air traffic will quadruple (double by 2025) [1] and aviation will contribute 6% of all human-induced climate change [2], while half of all air traffic will take off, land, or transit through Asia-Pacific [3]. Personnel costs already account for 60% of ANSP expenditure [4], while schedule predictability and flight delays are growing problems with significant direct and opportunity costs to national economies. Additionally, the exponential growth of unmanned traffic is expected to pose its own challenges and produce significant impacts on air traffic management (ATM) with clear consequences on both human–machine systems and infrastructure to support highly automated and resilient/trusted autonomous operations.

In this context, there seems to be little doubt that artificial intelligence (AI) and machine learning (ML) will be key enablers for advanced functionality and increased automation in the ATM system of tomorrow. Already we see the widespread adoption of AI and ML techniques in other industries driven by recent technological advances (graphics processing units (GPUs), cloud computing, big data, and deep learning algorithms) that are leading a resurgence in the field. AI is finally on the brink of realising its long-promised potential after several false starts and many governments and companies launched concerted AI initiatives.

For instance, as part of the transformation to a smart economy, the Australian government is investing ~\$30 million into AI and ML [5], realising the potential to develop new industries and increase the competitiveness of existing players.

AI and ML are important technologies for the digitisation of ATM—a very relevant topic today. The theme for the Civil Air Navigation Services Organisation’s (CANSO) 2018 Global ATM Summit, one of the most important ATM events globally, was “Air Traffic Management in the Age of Digitisation and Data”. The conference considered the likely impact of big data technologies on the industry, exploring aspects ranging from the operational, such as the migration of ATM systems to the cloud, the storage of voluminous space-based Automatic Dependent Surveillance–Broadcast (ADS–B) surveillance data, and cybersecurity, to the commercial, such as econometric analysis and forecasting, digital partnerships, and supplier relationships. As noted by CANSO Director General Poole, “Digitisation . . . has the potential to transform global ATM performance, bringing huge benefits . . . in terms of increased efficiency and enhanced safety” [6].

Despite this promising start, we contend that the ATM industry is inherently conservative and that the actual adoption of AI-driven automation, autonomy and human–machine teaming within the sector will likely be driven by specific use cases at the interface between ATM and adjacent aviation sectors that are either more innovative (e.g., unmanned aerial vehicles (UAVs) and unmanned aerial systems traffic management (UTM)) or less safety critical (e.g., air traffic flow management (ATFM)). We expect that the ATM sector will, in turn, contribute to developments in explainable AI, and the verification, qualification/validation, and certification (VQ&C) of systems incorporating AI techniques for these specific use cases before looking to adopt such techniques more generally.

The ATM industry remains interested in the medium-term convergence of several related technologies (AI, biometrics, cognition, trusted autonomy, variable autonomy, human–machine teaming, etc.), and the emerging societal, organisational, regulatory, and legislative reactions to them, as well as the possible cumulative impact of these diverse factors on the evolution of specific systems and procedures. Our main research questions build up to this impact:

- What VQ&C techniques are likely to be adopted for increasingly autonomous systems that incorporate AI techniques?
- To what extent do these VQ&C techniques and ongoing end-user acceptance (i.e., “trust”) of increasingly autonomous systems require that the AI techniques used be explainable?
- How can an explanation human–machine interface (HMI) component for explainable AI systems be developed on top of our existing work on cognitive HMIs?

These questions are tackled in the reverse order in this paper as we take a bottom-up approach. We begin by defining some basic concepts such as machine learning, neural networks, automation, and autonomy. The distinction between automation and autonomy warrants particular attention; however, before we can discuss any increase in the level of automation and autonomy, we first present several scales by which these levels can be measured. We introduce the concept of trust in order to consider how we may design autonomous systems that engender an appropriate level of trust. We then discuss how the AI and autonomy concepts presented may be applied in the aviation domain and ATM sector before entering the core discussion of explainable and cognitive HMIs and how they may be combined in a system. We consider the VQ&C implications for such a system, focusing our key findings on the specific use cases identified, although we outline future research directions in our conclusions.

2. Fundamental Concepts

The working definitions below are informed by standard references, but draw also on our practical experience.

Artificial intelligence is the ability of a system to partially simulate the workings of the human mind, although there is little evidence to suggest that this simulation is biologically plausible. In our context, we are concerned about “strong AI”-enabled autonomy for domain-specific tasks complex

enough to traditionally require a skilled human operator. We note, however, that the “artificial general intelligence” required for machines to perform with average human intelligence is yet to be developed. AI has many specialisations, and we are primarily concerned with machine learning in general and deep neural networks in particular.

Machine learning is an AI technique that learns patterns in data and evolves as that data keeps changing. Machine learning may be supervised or unsupervised, but both variants build probabilistic models in order to apply pattern recognition to new data.

Deep learning/deep neural networks are an advanced specialisation within machine learning that use a multi-layered model of human neural networks.

Fuzzy logic is an AI technique that uses “fuzzy” membership functions to remove hard distinctions, thereby accommodating ambiguity and imprecision when classifying data or relationships.

The adaptive neuro-fuzzy inference system (ANFIS) is a fuzzy logic system whose membership functions are tuned using neural network-like techniques.

Automation is the ability of a system to perform well-defined tasks without human intervention using a fixed set of “hard-coded” rules/algorithms to produce predictable, deterministic results. The automated tasks may be sub-tasks of a larger activity that involves human intervention—in which case, the overall activity is only partially automated to a lesser or greater degree.

Autonomy is the ability of a system to perform tasks without human intervention using behaviours, usually emergent, that arise from its interaction with the external environment. Such behaviours include reasoning, problem solving, goal-setting, self-adaptation/organisation, and machine learning, and may not be deterministic. In our context, the external environment may include human team members—in which case, the degree of autonomy exhibited by the system is dynamically variable as authority shifts between the system and the humans.

Emergent behaviour is common in biology and is a characteristic of self-organising autonomous systems interacting without central control to build more complex constructs. In systems theory, emergent properties arise spontaneously from the complex interactions of the various components of a system—i.e., they are not “programmed”. For example, the interaction of several autonomous, non-deterministic software agents can give rise to emergent properties—some unexpected, others anticipated as crucial to the performance of the overall system. In human-machine teaming, emergent properties can also arise from the collaboration amongst the human and machine team members. The degree of emergent behaviour is one of the characteristics that distinguishes autonomous systems from systems that are merely highly automated.

3. Autonomy

3.1. From Automation to Autonomy

Autonomous systems are distinguished from highly automated systems by their ability to respond to their environment and adapt their behaviour without being explicitly programmed to do so. Being non-algorithmic, they are often implemented using heuristics and non-deterministic AI techniques such as machine learning [7], deep neural networks, fuzzy logic, and genetic algorithms. Typical examples are the use of neural networks by autonomous cars to detect traffic signs [8] and the use of genetic algorithms and fuzzy logic to manage autonomous robotic systems [9].

Figure 1 outlines the fundamental differences between automation and autonomy in terms of emergent properties, whereas Table 1 compares the general characteristics of automation and autonomy when looking at a number of HMI and machine behavior/behaviour implications.

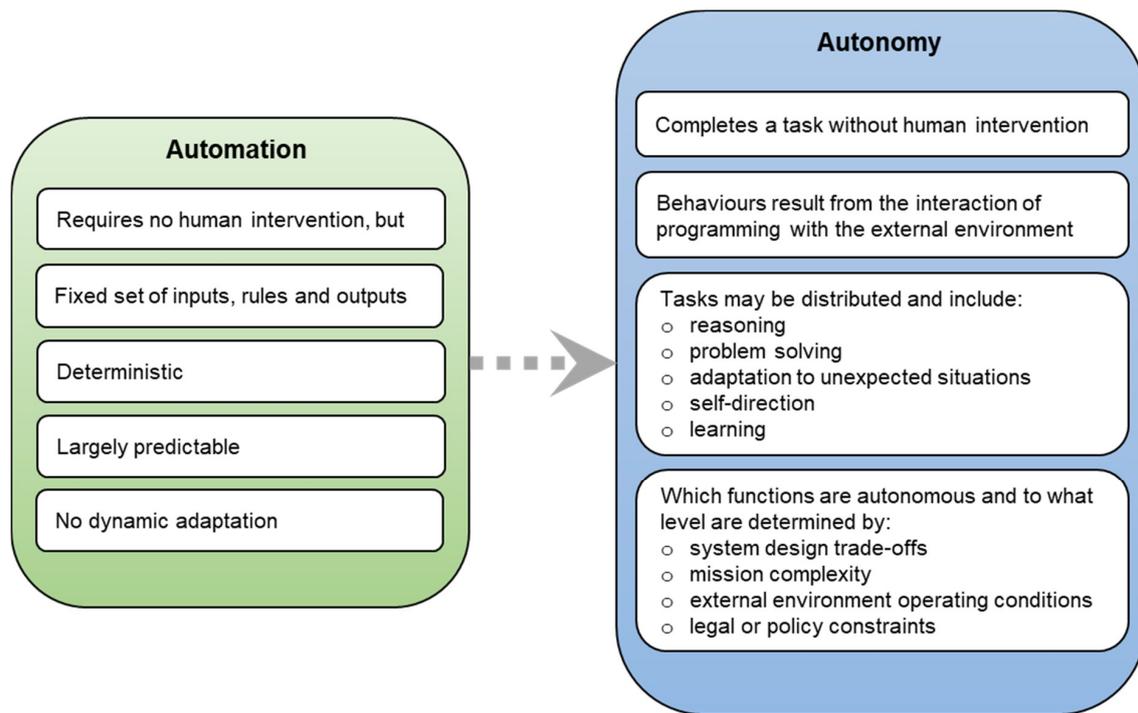


Figure 1. Contrasting automation and autonomy (adapted from Reference [10]).

Table 1. Characteristics of automation and autonomy (from Reference [11]). N/A—not applicable.

Characteristic	Automation	Autonomy
Augments human decision-makers	Usually	Usually
Proxy for human actions or decisions	Usually	Usually
Reacts at cyber speed	Usually	Usually
Reacts to the environment	Usually	Usually
Reduces tedious tasks	Usually	Usually
Robust to incomplete or missing data	Usually	Usually
Adapts behaviour to feedback (learns)	Sometimes	Usually
Exhibits emergent behaviour	Sometimes	Usually
Reduces cognitive workload for humans	Sometimes	Usually
Responds differently to identical inputs (non-deterministic)	Sometimes	Usually
Addresses situations beyond the routine	Rarely	Usually
Replaces human decision-makers	Rarely	Potentially
Robust to unanticipated situations	Limited	Usually
Adapts behaviour to unforeseen environmental changes	Rarely	Potentially
Behaviour is determined by experience rather than by design	Never	Usually
Makes value judgments (weighted decisions)	Never	Usually
Makes mistakes in perception and judgment	N/A	Potentially

From a practical perspective, however, autonomy relates to the level of independence that humans grant to a system to execute particular tasks in a specific environment. This implies some transfer of responsibility from the human to the machine based on an assumed or earned level of trust. Accordingly, defining and modelling trust became a topic of increasing research interest, as discussed in Section 5.

The term “increasingly autonomous” (IA; the acronym is not to be confused with intelligent agents—a popular way of implementing autonomous systems) systems is used to describe highly automated systems that are transitioning towards autonomy.

3.2. Measuring the Level of Automation/Autonomy

A meaningful discussion of increasing levels of automation or autonomy requires a measurement scale—a concept that became common knowledge with the advent of autonomous cars (Tesla’s autopilot is at level 2 of the Society of Automotive Engineers (SAE) J3016 scale (Table 3b)). As the variety of available scales outlined in Table 2 indicates, today, there is no consensus on how to measure varying degrees of automation or autonomy.

A scale that provides some indication of the tasks to be performed by both the human and the machine at the various levels is preferable to the high-level classification evident in the SAE’s J3016 scale (Table 3b). The authors currently use a version of Sheridan’s model of autonomy (Table 3a)—a simple linear scale, long used in aviation, that allows for intuitive judgements; for example, autonomy levels up to level 6 should be readily implementable in ATM systems provided that we supply a fool-proof HMI mechanism (highly visible, rapidly accessible, etc.) to abort the machine-proposed automation. Most such linear scales are categorical rather than cardinal—the difference between levels 2 and 3 cannot be considered to be the same as that between levels 8 and 9, nor does level 8 deliver “twice” the automation of level 4, for example.

Table 2. Various automation/autonomy scales (after Reference [10]). SAE—Society of Automotive Engineers; US—United States; OODA—Observe, Orient, Decide, and Act; DoD—Department of Defence; SESAR—Single European Sky Air Traffic Management Research.

Scale
Sheridan Model of Autonomy
Society of Automotive Engineers J3016
Clough’s Levels of Autonomy
US Navy Office Naval Research
Proud’s OODA Assessment
Clough’s Autonomy Control Level
Autonomy Levels Unmanned Systems
US DoD Defence Science Task Force
Billings’ Control-Management Continuum
SESAR Levels of Automation Taxonomy

Table 3. Prominent automation/autonomy scales (after Reference [15] and from Reference [16]).

a. Sheridan (aviation)		b. SAE J3016 (automobiles)	
1	Human does it all.	0	No automation
2	Machine offers alternatives and	1	Driver assistance
3	narrows selection to a few, or	2	Partial automation
4	suggests one, and	3	Conditional automation
5	executes it if human approves, or	4	High automation
6	allows human a set time to veto then executes automatically, or	5	Full automation
7	executes automatically and informs the human, or		
8	informs the human after execution if the human asks it, or		
9	informs the human after execution if it decides to.		
10	Machine acts autonomously.		

We note, however, that other scales used in aviation are also suitable and warrant further investigation. Billings’ early work [12] presented a control-management continuum for pilots that clearly enumerates the different automation and human functions to be performed at each of seven automation levels, assigning to each level a descriptive “management mode” identifier rather than a number. A more modern and ATM-specific measure is the Levels of Automation Taxonomy (LOAT) scale developed by Single European Sky ATM Research (SESAR) [13]. LOAT’s taxonomy is a matrix that aligns four sets of automation levels to four cognitive states that follow the decision-making process from information acquisition to action implementation. The cognitive states are closely aligned with those identified by Parasuraman as the various stages at which automation can be applied to a system [14].

3.3. Trust in Autonomy

Trustworthiness is a measure of how much someone or something should be trusted. In system engineering terms, it is related to reliability and is a quantity which we can readily subject to VQ&C.

EUROCONTROL's research into trust in future ATM systems determined quite early that trust is not the same as trustworthiness [17,18]. More contemporary work by Lee [19] notes that trust is distinguished from VQ&C and system/software assurance as it relates to the response to, and eventual adoption of, the system by the end user. Trust is, therefore, a social construct that becomes relevant to human-machine relationships when the complexity of the technology defies our ability to fully comprehend it. Trust is not an intrinsic characteristic of a system, but is rather an attitude based on human perception, experience, and prejudice that a system may or may not help to achieve a particular goal in situations characterised by uncertainty and vulnerability.

Following Lee [19], we can attempt the following working definitions:

$$\text{Overtrust} : \text{Trust} > \text{Trustworthiness}, \quad (1)$$

$$\text{Distrust} : \text{Trust} < \text{Trustworthiness}, \quad (2)$$

$$\text{Calibrated Trust} : \text{Trust} = \text{Trustworthiness}, \quad (3)$$

where our interest is in improving trust calibration via improved human-autonomy interactions.

Just as reliability is often used as a pseudonym for trustworthiness, so too is trust often associated with reliance. However, an operator may, at times, trust the machine, but not rely on it, while, at others, not trust the machine, but rely on it anyway. Following Finn and Mekdeci [20,21], we present firstly some hypotheses to assist in designing a system to overcome distrust.

$$\text{Performance} \propto \text{Trust}; \quad (4)$$

$$\text{Trust} \propto \text{Explanatory Ability}; \quad (5)$$

$$\text{Trust} \propto \frac{1}{\text{Unrelated System Failures}}; \quad (6)$$

$$\text{Trust} \propto \text{Meaningful Error Reporting}. \quad (7)$$

Then, using reliance as a proxy for trust, we follow these with hypotheses to detect scenarios likely to lead to either overtrust or unwarranted reliance.

$$\text{Reliance} \propto \text{Task Complexity}; \quad (8)$$

$$\text{Reliance} \propto \text{Workload}; \quad (9)$$

$$\text{Reliance} \propto \text{Convenience}. \quad (10)$$

Our rationale is to ensure or restore an appropriate level of calibrated trust in these situations via appropriate cognitive HMI measures.

4. Application in the Aviation/Air Traffic Management Domain

4.1. Autonomy in Air Traffic Management

Given the limited extent of automation present in ATM systems today, how realistic is it to expect the industry to embrace not just higher levels of automation, but increasing autonomy as well?

In 2013, the National Aeronautics and Space Administration (NASA) requested the National Research Council (NRC) to investigate autonomy in civil aviation. The goal was to set a research agenda that would include the following [11] (p. vii):

- Determining concepts of operation for interoperability between ground systems and aircraft with various autonomous capabilities.
- Predicting the system-level effects of incorporating IA systems and aircraft in controlled airspace.

The NRC report addresses the spectrum of autonomy between current ATM automation (e.g., safety nets and alerting, and decision support tools) and the type of adaptive, non-deterministic systems required to enable fully autonomous ATM systems in future. In general terms, it lists the possible uses of IA in ATM systems as follows [11] (pp. 23–24):

- Observe: Scan the environment by monitoring many more data sources than a human could.
- Orient.: Synthesise this data into information, e.g., as follows:
 - Monitor voice and data communications for inconsistencies and mistakes.
 - Monitor aircraft tracks for deviations from clearances.
 - Identify flight path conflicts.
 - Monitor weather for potential hazards, as well as potential degradations in capacity.
 - Detect imbalances between airspace demand and capacity.
- Decide: Identify and evaluate traffic management options and recommend a course of action.
- Act: Disseminate controller decisions via voice and/or datalink communications, where the effectiveness of IA in ATM systems would be enhanced by the presence of compatible airborne IA systems.

Of the numerous technological and regulatory barriers identified, the following were noted as being particularly challenging [11] (p. 5):

- Decision-making by adaptive or non-deterministic systems (such as neural networks).
- Trust in adaptive or non-deterministic IA systems.
- Verification, qualification/validation, and certification (VQ&C).

The report notes that existing difficulties with developing and maintaining a mental model of what automation is doing at any time could be exacerbated with the advent of advanced IA systems, particularly those exhibiting adaptive or non-deterministic behaviour. Of particular interest are the following recommendations [11] (p. 5):

- Determine how the roles of key personnel and systems should evolve as follows:
 - The impact on the human–machine interfaces (HMIs) of associated IA systems during both normal and atypical operations.
 - Assessing the ability of human operators to perform their new roles under realistic operating conditions, coupled with
 - the dynamic reallocation of functions between humans and machines based on factors such as fatigue, risk, and surprise [11] (p. 56)—which can be determined from biometric sensors and a cognitive model of human performance.
 - Developing intuitive HMI techniques with new modalities (such as touch and gesture) to [11] (p. 58) achieve the following:
 - Support real-time decision-making in high-stress dynamic conditions.
 - Support the enhanced situational awareness required to integrate IA systems.
 - Effective communication, including at the HMI level, amongst different IA systems and amongst IA and non-IA systems and their operators.
- Develop processes to engender broad stakeholder trust in IA systems as follows:
 - Identifying objective attributes and measures of trustworthiness.
 - Matching authority and responsibility with “earned levels of trust”.
 - Avoiding excessive or inappropriate trust [11] (p.58).
 - Determining the best way to communicate trust-related information.

4.2. Artificial Intelligence (AI) and Machine Learning (ML) in Aviation

Traffic collision avoidance systems (TCAS) are well-known airborne systems that prevent potential mid-air collisions by instructing one aircraft to climb and the other to descend. The TCAS code consists of over a thousand pages of explicit “if–then–else” rule statements [22]. The Federal Aviation Administration (FAA) is currently working on its successor, the airborne collision avoidance system (ACAS Xa). ACAS Xa is implemented as a deep neural network (DNN) with no explicit rule base. Instead, it is trained on millions of scenarios, including 180,000 real-life potential collisions, and it promises a 40% improvement over the latest version of TCAS [22].

Some important questions arise when considering the operational deployment of ACAS Xa and similar systems:

- Can you trust a non-deterministic DNN that can potentially deliver a different result each time that it is presented with the same scenario? (Note that for ACAS Xa, the potential for variability is moderated by filtering the generated solution set to find a TCAS-compatible resolution advisory and follow the same negotiation protocols as TCAS—interoperability is required to support mixed equipage. The situation is less clear for ACAS Xu, which supports vertical, horizontal, and merged manoeuvres to accommodate UAVs operating in controlled airspace and potential collisions with manned aircraft.)
- How do you know whether you are getting the right answer for the right reason?
- How do vendors verify such a solution, how does a regulator certify it, and how does an end user have confidence in its recommendations or autonomous actions?

The FAA was challenged on these and, in response, commissioned research for verifying DNNs [23]. This, and related work by Kochenderfer and Katz [23], is based on satisfiability modulo theories (SMT) and is a promising start for VQ&C; however, it still has a way to go when it comes to explaining the rationale for its decisions to end users—i.e., it does not address trust as defined above. Trust in automation is a long-standing issue in ATM. EUROCONTROL investigated conflict resolution assistants/advisories nearly 20 years ago [24], and they never achieved acceptance by air traffic controllers, with some early lessons discernible [17,18].

These and earlier studies highlight the need to consider the opinion of the end user. No less a body than the Air Traffic Control Association (ATCA) carried Nedelescu’s “A Conceptual Framework for Machine Autonomy” as the lead article of the winter 2016 edition of the *Journal of Air Traffic Control* [25]. Nedelescu argues that increased automation will only work in a tightly controlled environment; an unpredictable environment—such as one supporting point-to-point drone operations—calls for autonomy rather than high automation. Nedelescu notes that trust in autonomy requires a paradigm shift in VQ&C from a “once-off” design activity to a continuous operational activity—an autonomous machine can develop new behaviours while in operation. He is confident that safety can be achieved in a non-deterministic environment, arguing that one must make allowances for variations in how a machine achieves a valid outcome—particular solutions might contain an element of surprise, the outcome should not [25].

The implication here is that VQ&C cannot be confined to the factory. Just as human operators are periodically tested in the field, so too must adaptable machines. Moreover, the human–machine team needs to be tested as an integral unit to validate the parameters of variable autonomy; for example, has familiarity bred excessive or inappropriate trust?

This requires a deeper understanding of how automated systems engender trust in humans, particularly where the automation is black-box and/or non-deterministic, as is the case with most contemporary AI/ML techniques.

Ongoing verification during operations is not a new concept to ATM. The market has been requesting a predictive health and usage monitoring system (HUMS) capability for online diagnostics and prognostics of live algorithmic performance (e.g., is the conflict detection algorithm still performing

to speculation under current operational conditions?). The significance of this development is related to trust for human–machine teaming:

- Cognitive HMI: machine trust in the human;
- HUMS: human trust in today’s machine;
- Explainable AI: human trust in tomorrow’s IA machine.

5. Explainable AI and User Interface Design

Article 22 of the European Union’s General Data Protection Regulation (GDPR) requires that all AI algorithms be able to explain their rationale. AI can no longer be a “black box” and explainable AI (XAI) became a research topic of growing interest. Defence Advanced Research Projects Agency (DARPA) initiated an XAI program, and so too did major corporations such as Oracle and Amazon.

The DARPA XAI initiative [26] identifies a number of different explanation user interface design (UX) techniques, each with an associated explainable model clearly related to specific machine learning techniques such as Bayesian belief nets and decision trees. Significantly, each DARPA XAI explanation UX technique will be informed by the same psychological model of explanation, to be developed mainly by the Institute for Human and Machine Cognition (refer to Figure 2). Their approach is to extend the theory of naturalistic decision-making to cover explanation [26]. Several studies identified recognition-primed decision-making [27] as relevant for air traffic controllers.

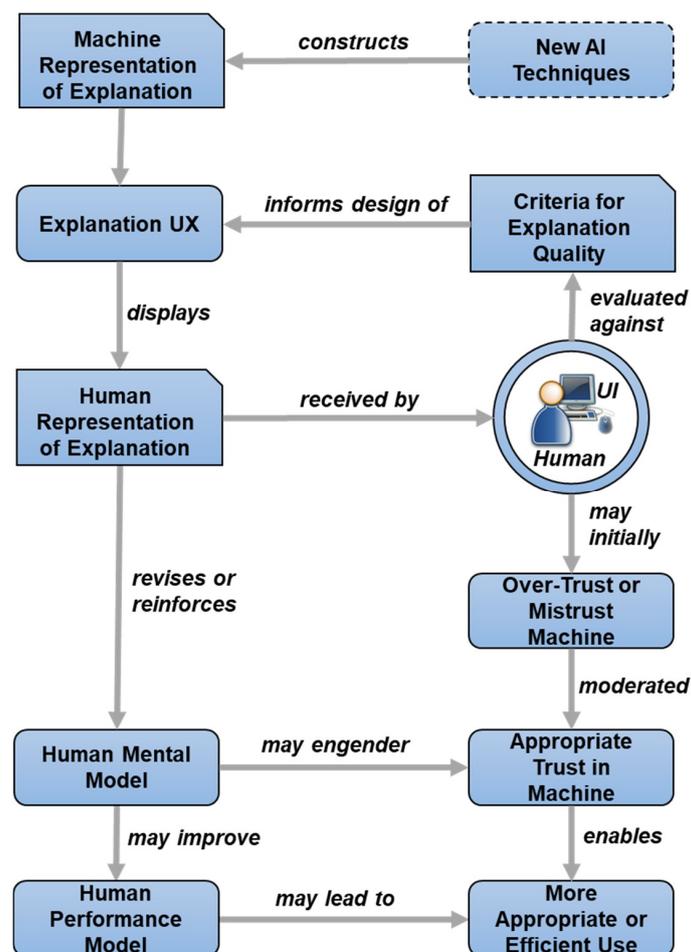


Figure 2. Example of a partial psychological model of explanation. Adapted from [26]. Note the operator interacting with the system via the explanation user interface within the circle.

Miller et al. [28] observed that a common problem with UX design is that it is often left to programmers rather than interaction designers. They posit that XAI is more likely to succeed if it adopts and adapts models from the existing body of research in philosophy, psychology, cognitive science, and human factors. Crucially, UX design decisions should be driven by the end user and validated via behavioural studies. This could imply some level of domain dependence when these techniques are deployed in different application domains. For example, consider which of the considerations for the general populace in Table 4 are also applicable to the following:

- Air traffic controllers in a busy approach environment,
- Military commanders in a command and control hierarchy, and
- Air traffic flow managers in a collaborative decision-making context.

Table 4. Some considerations for the user mental model (after Reference [28]). ATM—air traffic management; C&C— command & control; ATFM—air traffic flow management; XAI—explainable artificial intelligence; UX—user interface design.

Factor	Description	ATM	C&C	ATFM
Contrastive Explanation	“Why” questions are contrastive—they take the form “why P instead of Q”, where Q is a foil to P, the fact that requires explanation. If we can correctly anticipate Q, then we only need to contrast P and Q instead of providing a full causal explanation.	?	✓	✓
Social Attribution	Similar to the “belief–desire–intention” model used by intelligent agents, and it implies that we need a different explanation framework for actions that fail as opposed to actions that succeed.	✓	✓	✓
Causal Connection	People connect causes via a mental “what if” simulation of what would have happened differently if some event had turned out differently (a “counterfactual”). Understanding how people prune the large tree of possible counterfactuals (proximal vs. distal causes, normal vs. abnormal events, controllable vs. uncontrollable events, etc.) is crucial to efficient XAI.	✓	✓	✓
Explanation Selection	Humans are good at providing just enough facts for someone to infer a complete explanation. For causal chains with a number of causes, the visualisation techniques employed by the UX are crucial for allowing users to construct a preferred explanation.	✓	✓	?
Explanation Evaluation	Veracity is not the most important criterion people use to judge explanations. More pragmatic criteria include simplicity, generality, and coherence with prior knowledge or innate heuristics. A simpler explanation (with optional drill-down) may, therefore, be preferable if the primary goal is the establishment of trust as opposed to due-diligence completeness.	?	✗	✓
Explanation as Conversation	Explanations are usually interactive conversations. This may not be feasible in time-constrained situations; thus, UX design becomes crucial in minimising the need for interaction and ensuring that visual explanations conform to accepted conventions of conversation such as Grice’s maxims (paraphrased by Miller et al. as “only say what you believe; only say as much as is necessary; only say what is relevant; and say it in a nice way.” [28]).	✓	✓	✓

DARPA [26] identified that explainable AI requires both new machine learning processes and an explanation framework comprising both a psychological model of explanation and an explanation HMI. How one presents a machine-generated explanation to humans is crucial to its acceptance. Table 4 provides some insight into the factors that moderate how people provide explanations to each other, and engineers need to be cognisant of these when designing explanation HMIs. It is evident, however, that not only do these factors vary by application domain, but that they are also of a very high level. We believe that the construction of a high-fidelity explanation HMI relies on the participation of domain subject matter experts—in our case, air traffic controllers. Ultimately, only they can provide expert judgement on the quality or effectiveness of the explanation in the field. Table 5 provides an outline of the metrics that they might use in making this evaluation.

Table 5. Metrics of explanation quality [26].

Measure	Notes
User Satisfaction	<ul style="list-style-type: none"> • Clarity of the explanation • Utility of the explanation • Understanding individual decisions
Mental Model	<ul style="list-style-type: none"> • Understanding the overall model • Strength/weakness assessment • “What will it do” prediction • “How do I intervene” prediction
Task Performance	<ul style="list-style-type: none"> • Does the explanation improve the user’s decision, task performance?
Trust Assessment	<ul style="list-style-type: none"> • Artificial decision tasks introduced to diagnose the user’s understanding • Appropriate future use and trust
Correctability	<ul style="list-style-type: none"> • Identifying errors • Correcting errors • Continuous training

6. Cognitive Human–Machine Interface (HMI)

A cognitive HMI (C-HMI) is one which automatically adapts the information displayed and functions available based on an assessment of operator cognitive state and environmental conditions. The system may also use this assessment to execute actions autonomously along an escalating scale of automation (such as Sheridan’s).

We explored the application of C-HMIs to various problems in the ATM, ATFM, pilot/remote pilot, and UTM domains [29–35] using a general platform that integrates a variety of biometric sensors and interfaces with various simulators, as illustrated in Figure 3.

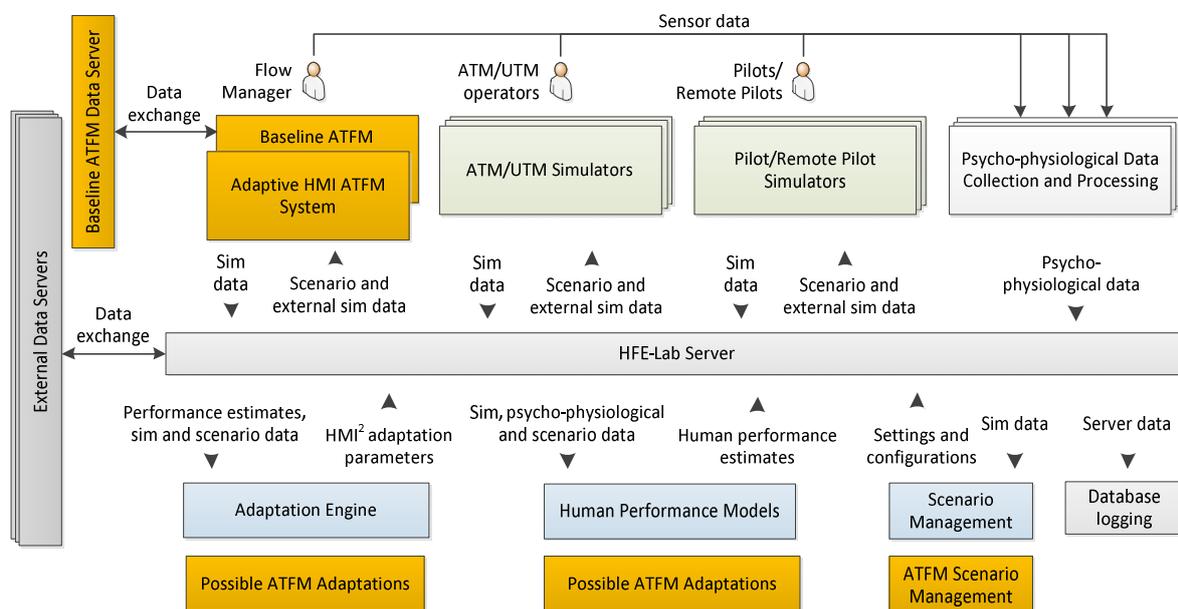


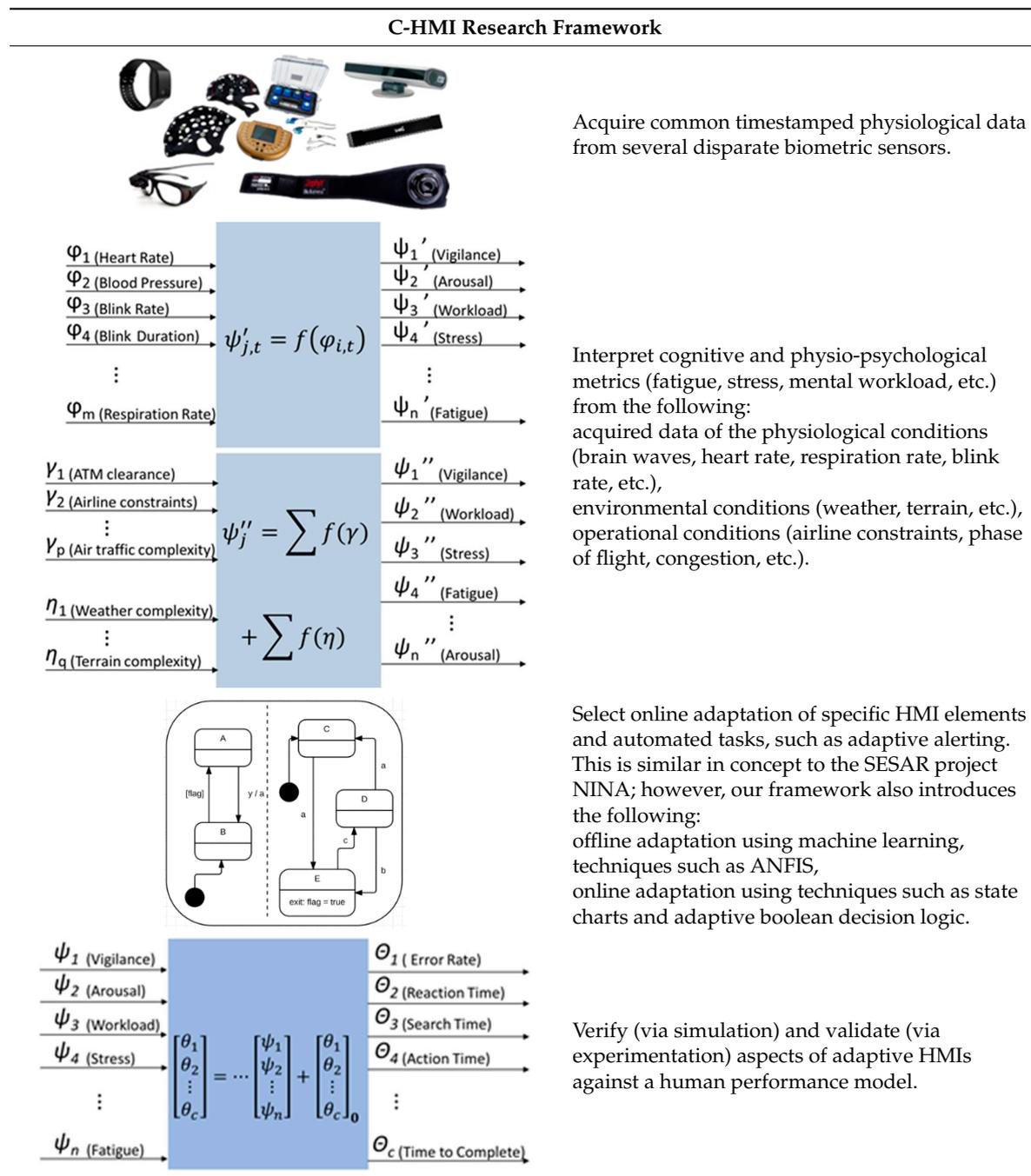
Figure 3. Architecture of the Royal Melbourne Institute of Technology (RMIT) air traffic management (ATM) laboratory (adapted from Reference [36]).

The laboratory test bench comprises several disparate biometric sensors (eye trackers, electroencephalograms (EEGs), and heart rate and respiration monitors) linked to a central data server for timestamping, consolidation, the computation of cognitive metrics, and human performance evaluation. The biometric sensors monitor the operators of the various networked simulators for different ATM (tower, en route, approach), ATFM, UTM, cockpit (pilot), and ground station (remote pilot) applications. Prior offline machine learning and adaptation sets performance baselines for each operator, and subsequently, environmental and situational data from the simulators contribute to the

online computation of the cognitive metrics and human performance evaluation, moderating the input of the biometric sensors. These metrics, in turn, help determine when and how aspects of the simulator HMIs should be dynamically adapted and for how long such adaptation should persist. Common scenario management allows specific simulation applications to interact—for example, an aircraft piloted by an operator in the cockpit simulator can appear as a track in the tower simulator; a track can be handed over from one air traffic control (ATC) position to another; etc.

The C-HMI framework developed thus far supports the functions outlined in Table 6.

Table 6. The cognitive human–machine interface (C-HMI) research framework (from Reference [37]). NINA—neurometrics indicators for ATM; ANFIS—adaptive neuro-fuzzy inference systems.



6.1. From Cognitive HMI to Explanation User Interface Design (UX)

Figure 4 illustrates current plans to develop and integrate an explanation UX with the cognitive UX (our C-HMI)—we believe that cognitive HMI elements would provide a machine with a predictive HUMS capability for humans, both for determining whether the human is still performing within acceptable parameters, before adjusting variable autonomy, and is accepting the machine’s XAI explanation, before adapting that explanation if needed.

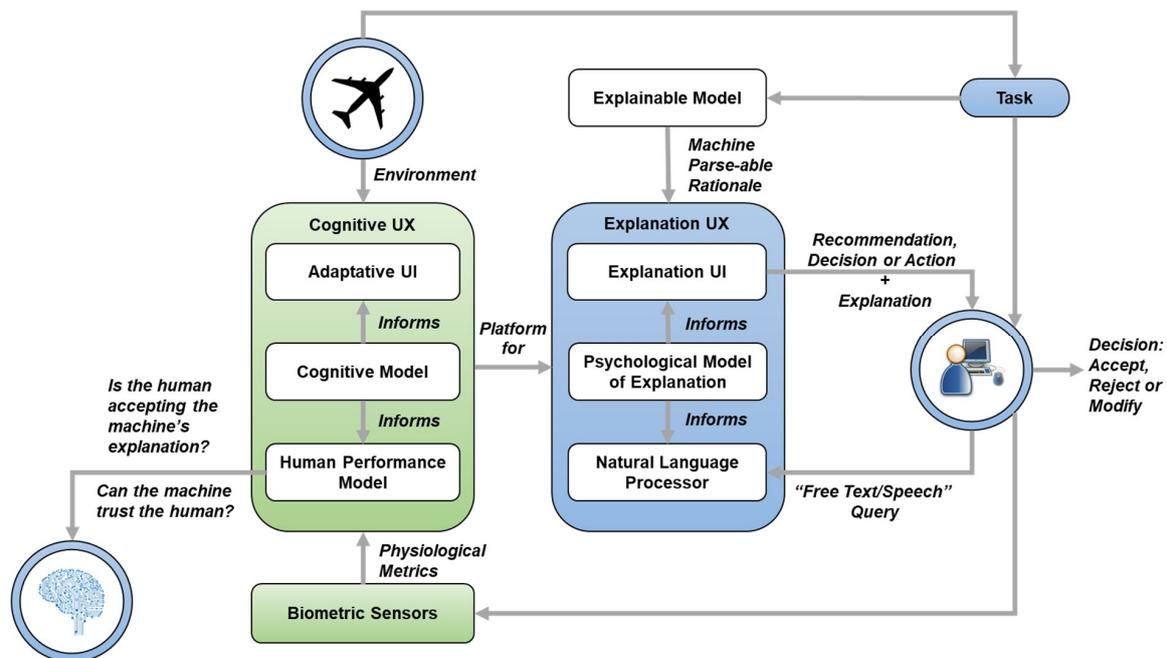


Figure 4. Combining cognitive and explanation user design interface (UX) principles (adapted from Reference [26]).

Of course, it is well worth questioning whether machine explanations should be so overtly “fine-tuned” to counteract human judgement. This call may need to be made on a case-by-case basis.

7. Regulatory Framework Evolutions: Certification versus Licensing

While there is a comprehensive regulatory framework for ATM (International Civil Aviation Organisation (ICAO), European Aviation Safety Agency (EASA), FAA, Civil Aviation Safety Authority (CASA), etc.), today, ground-based ATM systems are not required to be formally certified in the same manner as avionics. For example, ATM systems are not required to comply with either the Radio Technical Commission for Aeronautics (RTCA) DO-278 or DO-254. This gap in the standard has persisted for several decades, while ATM systems remain decision support tools with limited automation [38,39].

This is likely to change with increasing automation and emerging autonomy. However, the nature of this change may be unexpected. While the pursuit of a unified certification framework for integrated communication, navigation, surveillance and ATM (CNS+A) systems remains a worthwhile goal, for the IA components of ATM systems discussed in this paper, it could well be that they become subject to a process akin to ongoing personnel licensing rather than once-off type certification. This means that each deployment of such systems must be individually and regularly tested in the field as each could evolve/learn differently. Moreover, as discussed previously, they need to be tested in conjunction with the human members of their team (both existing and new controls).

Taking a cue from the automotive industry, computer-aided automated test techniques will be crucial for the pragmatic verification of autonomous systems. Just as it is impossible to explicitly

program an autonomous system to handle every possible scenario that it may encounter in the real world, so too is it impossible to use formal methods to verify that the system will behave correctly for every possible scenario. To proceed in an economically efficient manner, manufacturers of self-driving cars resort to automatically generating and testing a large, but finite, number of scenarios and put effort in ensuring that the scenarios generated, or retained after pruning, are realistic and relevant (e.g., both nominal and boundary cases such as state transitions) [40,41]. Each scenario needs to be run multiple times to check the extent of the variability in non-deterministic responses to the same stimuli.

However, ultimately, there will be restrictions on what initial certification will be able to achieve, and greater emphasis needs to be placed on the continuing certification process. Each new IA deployment will have to undergo a comprehensive, probationary shake-down as its human team members develop trust in its performance. Case-based scenarios written for such testing needs to accommodate some variability in IA response, and this is likely to be an education process for both vendors and customers. There is a role for the regulator in specifying the envelope of acceptable variability.

8. Key Findings

A review of the test plan for current C-HMI development activities already raised several VQ&C concerns, and resulted in the addition of state charts for online adaptation—a more deterministic approach that is acceptable for initial use in ATM. Originally ANFIS-based techniques, as used for offline adaptation, were considered. The use of AI/ML techniques for the offline/configuration/dataset/post-operations analytics aspects of ATM is not of particular concern, particularly since a technique like ANFIS has explanatory power and allows us to accommodate both the natural variability found across human operators, as well as the measurement noise produced by biosensors. Concerns arise when we wish to address the natural variability found in a single operator over the duration of his or her shift as operational conditions change. The concept of an ATM system modifying its own behaviour online during operations—potentially outside of specifications—is bound to raise safety concerns.

Blanket certification requirements are unlikely to be helpful. The regulators may focus first on those use cases where the ATM industry will face serious autonomy considerations, likely to be where it necessarily has to interface to external systems that themselves exhibit high levels of automation and increasing autonomy. A prime candidate for such an autonomy vector is the presence of unmanned aerial vehicles (UAVs) or unmanned aerial systems (UASs) operating in, or adjacent to, controlled or common traffic advisory frequency (CTAF) airspace and any UTM or fleet management systems in charge of them.

The National Research Council's 2018 consensus study report into In-Time Aviation Safety Management makes special mention of "Trust in Increasingly Autonomous UAS and Associated Traffic Management Systems" [42], noting the increased uncertainty from new entrants and emergent risks as a major challenge for air traffic controllers.

8.1. ATM-UTM Integration

The UTM and fleet management systems that the ATM system will connect to in order to exchange flight and mission plans and airspace reservation/geofencing data are likely to be highly automated and increasingly autonomous. What authority will be delegated to the UTM or fleet management system to autonomously change mission plans and geofences—or to ignore ATM instructions for local or commercial considerations? Human-machine teams may have "disparate or asymmetric goals, information, and abilities; understanding collaboration will involve accounting for and aligning these asymmetries" [43]. The implications on the ATM system and controllers have to be investigated.

8.2. Impact of UAS on ATM

Seasoned industry practitioners observed that systems developed for autonomous low-level operations may soon migrate to higher levels [44]. The National Air Traffic Services (NATS) already

refers to UTM as “unified traffic management” and Deutsches Zentrum fuer Luft- und Raumfahrt (DLR) just concluded a study into unmanned air freight operations, investigating several new operational concepts and their impact on the ATM system [45,46]:

- Relief Operations: the construction of segregated airspace corridors for unmanned relief missions. Unmanned freighters fly in formation and are separated from surrounding conventional traffic. During simulations, following aircraft in the formation showed a 15% reduction in fuel consumption, but controller taskload was higher than normal.
- Long-Haul Freight: Unmanned freighters are not segregated, but subject to “sectorless” control. Specially trained controllers monitor the unmanned freighters over long stretches of their route that cut across traditional sector boundaries.
- Airport Integration: Unmanned freighters are integrated into the arrival and departure sequences with consideration of their special requirements. ATM systems were enhanced to permit controllers to recognise the special characteristics of the unmanned freighters permitting, for example, standard surface operations such as towing to and from the runway with handover to a remote pilot. A designated engine start-up area may be required for drones to allow conventional traffic to pass them on the taxiway.

8.3. Air Traffic Flow Management (ATFM)

At the other end of the scale from UAS, projected increases in traffic volumes mark ATFM as an increasing important sub-set of ATM. As technology enablers introduce true gate-to-gate operations we can expect to see many of the distinctions between ATFM measures and ATM techniques blur [47].

Nedelescu lists emergent patterns in ATFM as another potential autonomy vector [25], but ATFM is likely to remain an advisory service for some time and unlikely to be subject to certification, particularly if implemented in a multi-state, regional context. This provides more scope for the introduction of novel technologies. Parasuraman [14] has observed that there are four stages at which automation can be applied to a system:

- Data acquisition,
- Data interpretation,
- Decision selection,
- Action selection.

Table 7 illustrates how we can apply this classification to ATFM [48].

Table 7. Categories of automation applied to ATFM (from [48]). ADS-B—automatic dependent surveillance-broadcast; 4D—four-dimensional.

Data Acquisition	Data Interpretation	Decision Selection	Action Selection
Smart Sensors: <ul style="list-style-type: none"> • Space-based ADS-B • 4D weather cube • Biometrics 	Identification and prediction: <ul style="list-style-type: none"> • major traffic flows • workload • congestion • flight delays • arrival time 	Decision support: <ul style="list-style-type: none"> • scheduling • multi-agent flow control • sector planning • airport configuration 	Pre-tactical conflict detection and resolution: <ul style="list-style-type: none"> • hotspots • multiple flights or flows • weather

9. Conclusions and Future Research

An important enabler of future ATM operational concepts is the expectation that both the ATM system of tomorrow and the aircraft being controlled will be more automated and increasingly autonomous. In this context, how will air traffic controllers interact with increasing autonomy in both the ATM and external systems? What will be their changing role in a world where humans team with autonomous machines, exchanging functions dynamically as levels of trust shift between the two? A number of general problems need to be addressed:

- How do we establish appropriate scales and practical measures for both autonomy and trust in that autonomy?
- How do we determine the current trustworthiness of the humans and machines in the team, match authority with “earned levels of trust” and vary responsibility between them while avoiding excessive or inappropriate trust?
- By what criteria do we judge the quality of a machine-provided explanation and how do we present it on the HMI in a manner that the controller is more likely to trust to an appropriate level?

More specifically, using the lessons learned from recent research on cognitive HMI, how can we construct an explanation framework and, in particular, an explanation HMI that will do the following:

- yield immediate benefits where high degrees of ATM automation are already present (e.g., auto-completion of datalink uplink messages, and arrival sequencing) or already planned (inferring user intent, and re-routing flights),
- engender an appropriate level of calibrated trust, minimising both unwarranted distrust and overtrust,
- address the HMI requirements for variable autonomy in human–machine teaming,
- adapt when new explainable machine learning models become available?

Furthermore, what holistic VQ&C strategies can we propose to regulators for these new techniques and systems? Variable autonomy in particular is not only an important end-goal in its own right, but also a key transitory step towards the acceptance of fully autonomous systems. We believe that both the cognitive HMI (machine monitoring the human) and explanatory HMI (human monitoring the machine) aspects of our work will help to determine when, and to what extent, authority and autonomy should be dynamically shifted between the human and machine members of a mixed team. Our conclusions in this regard are that VQ&C activities need to cover scenarios across the variable autonomy spectrum, as well as the conditions (e.g., degree of trust) that trigger the transfer of authority and modulations in autonomy.

Much work remains to be done, but already we can see that a paradigm shift is required that calls for us to treat adaptable machines more like humans, accepting a degree of variability and potential fallibility within an envelope bounded by trust. In this respect, the transition from the conventional once-off system certification framework to an approach similar to ongoing personnel licensing, with regular testing, may prove more appropriate to accommodate IA in ATM systems and avionics.

Author Contributions: Conceptualization, T.K., R.S. and A.G.; Methodology, T.K., A.G. and R.S.; Software, A.G. and T.K.; Validation, T.K., A.G. and R.S.; Formal Analysis, A.G., T.K. and R.S.; Investigation, T.K., A.G., R.S.; Resources, A.G.; Data Curation, A.G.; Writing-Original Draft Preparation, T.K., A.G. and R.S.; Writing-Review & Editing, A.G. and R.S.; Visualization, T.K. and A.G.; Supervision, R.S. and A.G.; Project Administration, A.G.; Funding Acquisition, R.S.

Funding: The authors wish to thank and acknowledge Thales Air Traffic Management, Melbourne, Australia for supporting this work under the collaborative research project RE-02544-0200315666.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bisignani, G. *Vision 2050 Report*; IATA: Singapore, 2011.
2. Royal Commission on Environmental Pollution. *The Environmental Effects of Civil Aircraft in Flight*; Royal Commission on Environmental Pollution: London, UK, 2002.
3. IATA. *20 Year Passenger Forecast—Global Report*; IATA: Geneva, Switzerland, 2008.
4. Nero, G.; Portet, S. Five Years Experience in ATM Cost Benchmarking. In Proceedings of the 7th USA/Europe Air Traffic Management R&D Seminar, Barcelona, Spain, 2–5 July 2007.
5. Budget Paper No. 1 Budget Strategy and Outlook 2018–19. Available online: https://www.budget.gov.au/2018-19/content/bp1/download/BP1_full.pdf (accessed on 10 May 2018).

6. CANSO Global ATM Summit and 22nd AGM: Air Traffic Management in the Age of Digitisation and Data. Available online: <https://www.canso.org/canso-global-atm-summit-and-22nd-agm-air-traffic-management-age-digitisation-and-data> (accessed on 14 September 2018).
7. Xu, X.; He, H.; Zhao, D.; Sun, S.; Busoniu, L.; Yang, S.X. Machine Learning with Applications to Autonomous Systems. *Math. Probl. Eng.* **2015**, *2015*, 385028. [[CrossRef](#)]
8. Arcos-García, Á.; Álvarez-García, J.A.; Soria-Morillo, L.M. Evaluation of Deep Neural Networks for Traffic Sign Detection Systems. *Neurocomputing* **2018**, *316*, 332–344. [[CrossRef](#)]
9. De Figueiredo, M.O.; Tasinaffo, P.M.; Dias, L.A.V. Modeling Autonomous Nonlinear Dynamic Systems Using Mean Derivatives, Fuzzy Logic and Genetic Algorithms. *Int. J. Innov. Comput. Inf. Control* **2016**, *12*, 1721–1743.
10. Williams, A.P.; Scharre, P.D. *Autonomous Systems: Issues for Defence Policymakers*; NATO Communications and Information Agency: The Hague, The Netherlands, 2015.
11. National Research Councils. *Autonomy Research for Civil Aviation: Toward a New Era of Flight*; National Academy Press: Washington, DC, USA, 2014; ISBN 987-0-309-38688-3. [[CrossRef](#)]
12. Billings, C.E. *Aviation Automation: The Search for A Human-Centered Approach*; CRC Press: Boca Raton, FL, USA, 1997.
13. CAP 1377 ATM Automation: Guidance on Human-Technology Integration. Available online: <https://skybrary.aero/bookshelf/content/index.php?titleSearch=ATM+Automation%3A+Guidance+on+human-technology+integration&authorSearch=&summarySearch=&categorySearch=&Submit=Search> (accessed on 14 September 2018).
14. Parasuraman, R.; Sheridan, T.B.; Wickens, C.D. A Model for Types and Levels of Human Interaction with Automation. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **2000**, *30*, 286–297. [[CrossRef](#)]
15. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles J3016_201609. Available online: http://standards.sae.org/j3016_201609/ (accessed on 14 September 2018).
16. Sheridan, T.B. *Telerobotics, Automation, and Human Supervisory Control*; MIT Press: Cambridge, MA, USA, 1992.
17. Kelly, C.; Boardman, M.; Goillau, P.; Jeannot, E. *Guidelines for Trust in Future ATM Systems: A Literature Review*; Reference No. 030317-01; EUROCONTROL: Brussels, Belgium, 2003.
18. Kelly, C. *Guidelines for Trust in Future ATM Systems: Principles*; HRS/HSP-005-GUI-03; EUROCONTROL: Brussels, Belgium, 2003.
19. Lee, J.D.; See, K.A. Trust in Automation: Designing for Appropriate Reliance. *Hum. Factors* **2004**, *46*, 50–80. [[CrossRef](#)] [[PubMed](#)]
20. Mekdeci, B. Calibrated Trust and What Makes A Trusted Autonomous System. Emerging Disruptive Technologies Assessment Symposium in Trusted Autonomous Systems. Available online: https://www.dst.defence.gov.au/sites/default/files/basic_pages/documents/09_Mekdeci_UniSA.pdf (accessed on 14 September 2018).
21. Finn, A.; Mekdeci, B. Defence Science & Technology Organisation Report: Trusted Autonomy. Available online: <http://search.ror.unisa.edu.au/record/9916158006401831/media/digital/open/9916158006401831/12149369720001831/13149369710001831/pdf> (accessed on 14 September 2018).
22. Smarter Collision Avoidance. Available online: <https://aerospaceamerica.aiaa.org/features/smarter-collision-avoidance/> (accessed on 14 September 2018).
23. Katz, G.; Barrett, C.; Dill, D.L.; Julian, K.; Kochenderfer, M.J. Reluplex: An Efficient SMT Solver for Verifying Deep Neural Networks. In *Computer Aided Verification*; Majumdar, R., Kunčák, V., Eds.; Springer: Cham, Switzerland, 2017; pp. 97–117.
24. Kirwan, B.; Flynn, M. *Investigating Air Traffic Controller Conflict Resolution Strategies*; Rep. ASA, 1; EUROCONTROL: Brussels, Belgium, 2002.
25. Nedelescu, L. A Conceptual Framework for Machine Autonomy. *J. Air Traffic Control* **2016**, *58*, 26–31.
26. DARPA, Explainable Artificial Intelligence (XAI) Program Update, DARPA/I2O. November 2017. Available online: <https://www.darpa.mil/attachments/XAIProgramUpdate.pdf> (accessed on 19 May 2018).
27. Klein, G. Naturalistic Decision Making. *Hum. Factors* **2008**, *50*, 456–460. [[CrossRef](#)] [[PubMed](#)]
28. Miller, T.; Howe, P.; Sonenberg, L. Explainable AI: Beware of Inmates Running the Asylum. Available online: <https://arxiv.org/pdf/1712.00547.pdf> (accessed on 19 May 2018).
29. Liu, J.; Gardi, A.; Ramasamy, S.; Lim, Y.; Sabatini, R. Cognitive Pilot-Aircraft Interface for Single-Pilot Operations. *Knowl. Based Syst.* **2016**, *112*, 37–53. [[CrossRef](#)]

30. Lim, Y.; Liu, J.; Ramasamy, S.; Sabatini, R. Cognitive Remote Pilot-Aircraft Interface for UAS Operations. In Proceedings of the 2016 International Conference on Intelligent Unmanned Systems (ICIUS 2016), Xi'an, China, 23–25 August 2016.
31. Lim, Y.; Bassien-Capsa, V.; Ramasamy, S.; Liu, J.; Sabatini, R. Commercial Airline Single-Pilot Operations: System Design and Pathways to Certification. *IEEE Aerosp. Electron. Syst. Mag.* **2017**, *32*, 4–21. [[CrossRef](#)]
32. Lim, Y.; Gardi, A.; Ramasamy, S.; Sabatini, R. A Virtual Pilot Assistant System for Single Pilot Operations of Commercial Transport Aircraft. In Proceedings of the 17th Australian International Aerospace Congress (AIAC 2017), Melbourne, Australia, 26–28 February 2017.
33. Lim, Y.; Gardi, A.; Ramasamy, S.; Vince, J.; Pongracic, H.; Kistan, T.; Sabatini, R. A Novel Simulation Environment for Cognitive Human Factors Engineering Research. In Proceedings of the 36th IEEE/AIAA Digital Avionics Systems Conference (DASC), St Petersburg, FL, USA, 17–21 September 2017.
34. Lim, Y.; Gardi, A.; Ezer, N.; Kistan, T.; Sabatini, R. Eye-Tracking Sensors for Adaptive Aerospace Human-Machine Interfaces and Interactions. In Proceedings of the 2018 5th IEEE International Workshop on Metrology for AeroSpace (MetroAeroSpace), Rome, Italy, 20–22 June 2018.
35. Lim, Y.; Ramasamy, S.; Gardi, A.; Kistan, T.; Sabatini, R. Cognitive Human-Machine Interfaces and Interactions for Unmanned Aircraft. *J. Intell. Robot. Syst. Theory Appl.* **2018**, *91*, 755–774. [[CrossRef](#)]
36. RMIT. *Development of a Cognitive HMI for Air Traffic Management Systems—THALES Report 1*; Project 0200315666, Ref. RMIT/SENG/ICTS/AVIATION/001-2017; RMIT University: Melbourne, Australia, 2017.
37. RMIT. *Development of a Cognitive HMI for Air Traffic Management Systems—THALES Report 2*; Project 0200315666, Ref. RMIT/SENG/ITS/AVIATION/002-2017; RMIT University: Melbourne, Australia, 2017.
38. Batuwangala, E.; Gardi, A.; Sabatini, R. The Certification Challenge of Integrated Avionics and Air Traffic Management Systems. In Proceedings of the Australasian Transport Research Forum, Melbourne, Australia, 16–18 November 2016.
39. Batuwangala, E.; Kistan, T.; Gardi, A.; Sabatini, R. Certification Challenges for Next-Generation Avionics and Air Traffic Management Systems. *IEEE Aerosp. Electron. Syst. Mag.* in press. [[CrossRef](#)]
40. Straub, J. Automated Testing of A Self-Driving Vehicle System. In Proceedings of the 2017 IEEE AUTOTESTCON, Schaumburg, IL, USA, 9–15 September 2017; pp. 1–6.
41. Mullins, G.E.; Stankiewicz, P.G.; Gupta, S.K. Automated Generation of Diverse and Challenging Scenarios for Test and Evaluation of Autonomous Vehicles. In Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 29 May–3 June 2017; pp. 1443–1450.
42. National Research Councils. *In-Time Aviation Safety Management: Challenges and Research for an Evolving Aviation System*; National Academy Press: Washington, DC, USA, 2018; ISBN 978-0-309-46880-0. [[CrossRef](#)]
43. National Research Councils. *Intelligent Human-Machine Collaboration: Summary of a Workshop*; National Academy Press: Washington, DC, USA, 2012; ISBN 978-0-309-26264-4. [[CrossRef](#)]
44. Butterworth-Hayes, P. From ATM to UTM . . . and back again. *CANSO Airspace Magazine*, June 2018; p. 30.
45. Temme, A.; Helm, S. Unmanned Freight Operations. In Proceedings of the DLRK 2016, Braunschweig, Germany, 13–15 September 2016.
46. Luchkova, T.; Temme, A.; Schultz, M. Integration of Unmanned Freight Formation Flights in The European Air Traffic Management System. In Proceedings of the ENRI International Workshop on ATM/CNS (EIWAC), Tokyo, Japan, 14–16 November 2017.
47. Kistan, T.; Gardi, A.; Sabatini, R.; Ramasamy, S.; Batuwangala, E. An Evolutionary Outlook of Air Traffic Flow Management Techniques. *J. Prog. Aerosp. Sci.* **2017**, *88*, 15–42. [[CrossRef](#)]
48. Kistan, T. Innovation in ATFM: The Rise of Artificial Intelligence, ICAO ATFM Global Symposium, Singapore. November 2017. Available online: <https://www.icao.int/Meetings/ATFM2017/Pages/Session-10-Presentation-Innovation.aspx> (accessed on 14 September 2018).

