This article has been accepted for publication in IEEE Systems, Man, and Cybernetics Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MSMC.2021.3136983

# Frontiers of Brain-Inspired Autonomous Systems:
## How Does the Defence R&D Drive the Innovations?

M. Hou, Y. Wang, L. Trajkovic, K. N. Plataniotis, S. Kwong, M. Zhou, E. Tunstel, I. Rudas, J. Kacprzyk, and H. Leung

*Abstract*— **A brain-inspired Intelligent Adaptive System (IAS) is developed towards fundamental breakthroughs in the cognitive bottleneck of humans and the incompetence of AI under indeterministic conditions or with insufficient data. IAS has led to defence science and technology innovations for Interaction-Centered Design (ICD) methodologies, human-autonomy symbiosis initiatives, and a trust framework synergizing key strategies on intention, measurability, performance, adaptivity, communication, transparency, and security (IMPACTS) for trustworthy mission-critical autonomous systems. These paradigms of emerging technologies empower highly automated systems to think and behave like humans for generating collective intelligence. The IAS-based autonomous systems have not only fostered the development of a series of novel theories and methodologies such as brain-inspired systems and ICD approach, but also paved unprecedented paths to innovative applications in the defence and general industries.**

*Index Terms* — **Artificial intelligence, brain-inspired system, human-AI symbiosis, human-autonomy teaming, interaction-centered design, intelligent adaptive system, trust**

## I. INTRODUCTION

It is recognized that many fundamental theories and innovative technologies have been discovered or triggered by defence research and development including modern computers, Internet, and autonomous systems. Artificial intelligence (AI) has become ubiquitous and capable of rendering autonomous reasoning and decisions. It continuously changes the roles and responsibilities of human functions in human-machine symbiotic partnership. Human roles are becoming increasingly supervisory in nature and more contextual decisions are applied to technology [1].

However, this trend poses important questions about the limitations, liabilities, risks, privacy, ethics, and trust associated with increasing autonomous decision-making capabilities in safety and mission-critical applications such as self-driving vehicles, homecare and surgical robots, Industry 4.0 smart manufacturing systems, or remotely piloted combat drones [2], [3], [4], [5], [6]. Data-driven AI algorithms require big data to train that may not be suitable for many real-time applications. When facing insufficient data, incomplete information, indeterministic conditions, or inexhaustive solutions for uncertain actions, data-regression AI is unable to provide timely support in decisions regarding the what, where, when, who, and how associated with operational situations due to the absence of autonomous decision-making theories [7], [8]. With the increased contextual complexity and behavioral opacity of AI, it is even more challenging for humans to maintain sufficient situational awareness (SA) and safe responsibility transfer during the transition from “on-the-loop” to “in-the-loop” when AI and autonomy are coupled closely with humans [9]. The two fatal crashes of the Boeing 737 Max in 2018 and 2019 are typical examples of a failed responsibility transfer between an AI-enabled autonomous function and pilots due to the faulty assumption about human cognitive capacities and disregard for design principles when developing new technologies based on outdated designs created a few decades ago [10]. The investigation report on another airplane incident due to a machine failure, the emergency landing of US Airway 1549 in the Hudson River in 2009, revealed that the pilot needed sufficient resources (time and attention) to process information, assess the situation, make right decisions, and take over the control when the machine (aircraft) failed [9] [11]. These examples reiterate that the design of these emerging technologies needs to seriously consider the capability limitations of both human and machine intelligence from the onset rather than as an afterthought.

Further, networked systems with associated interconnectivity and interdependence accelerate the spiking global complexity. Thus, it is paramount to realize the imperative needs of guidance for understanding and mitigating the risks during the design, development, validation, certification, and exploitation processes for these socio-technical systems [3]. In fact, the entire spectrum of autonomous systems demands innovative technological solutions, enduring strategies, science-based design methodologies, and evidence-based process standards to ensure that these emerging technologies can be trusted and employed safely, effectively, reliably, legally, and ethically before AI or autonomy is integrated more widely into our systems, operations, and society [12], [13], [14], [15].

This article presents the latest defence research and development-driven autonomous systems and a novel brain-inspired Intelligent Adaptive System (IAS) as a technological solution to the aforementioned issues. An associated coherent body of Interaction-Centered Design (ICD) methodologies and an IMPACTS trust framework are elucidated as systematic strategies and structured guidance for designing trustworthy IASs where critical decisions are made by AI, humans, or both. Real-world brain-inspired IAS examples are demonstrated as strategy and paradigm validation studies. Future directions are suggested for broader ICD and IMPACTS related application, doctrine, and process endeavors in an IAS design, development, validation, certification, and exploitation life cycle.

## II. THE EMERGENCE OF BRAIN-INSPIRED INTELLIGENT ADAPTIVE SYSTEMS

It is commonly assumed that autonomous systems towards general AI may be achieved when brain-inspired cognitive capabilities for decision-making are available to exhibit humanlike behaviors of both intelligence and adaptability. Recent theoretical advances and technological convergence in cybernetics, system science, cognitive informatics, and augmented cognition offer possibilities of a brain-inspired and knowledge-based IAS that harnesses leap-ahead technologies such as the next generation of computing (quantum computing and cognitive computing) [9], [16], [17], [18]. Thus, a machine is able to "think" like humans in order to improve decision-making and generate anticipatory intelligence by mimicking the neurological processes of the human brain. With the collective human-machine intelligence, an IAS is able to effectively draw inferences from data including those that are incomplete, uncertain, or ambiguous and to learn from experience and users' interaction responses.

Therefore, an IAS is capable of changing its behavior in real-time as a function of a user cognitive state and the status of task, machine, and world (working environment). By optimizing human-machine interactions, an IAS may intelligently adapt to the capabilities, capacities, limitations, needs, and demands of a user and machine to attain and maintain safety, trust, effectiveness, and efficiency when humans perform various cognitively challenging tasks in complex and dynamic environments [9].

### *A. Architectural Framework of Intelligent Adaptive Systems (IASs)*

To achieve IAS objectives for its broad range of applications, a system needs to exhibit at least five fundamental characteristics: 1) tracking goals, plans, or intents, and the progress towards them; 2) monitoring and inferring the internal state of a user (behavioral, physical, cognitive, and emotional); 3) monitoring and inferring the external status of the world (environmental conditions, entities, and domain constraints); 4) monitoring the effects of machine status, automation, advice, and adaptation on user and world status (closed-loop feedback); and 5) customizing its human-machine interface (HMI) including brain-machine interface to handle the interactions and trust relations between a user and machine. To manifest these characteristics, a conceptual architectural framework has been developed as a basic IAS anatomy with critical components common to knowledge-based IASs [9].

As illustrated in Fig. 1, a generic IAS includes four modules pertaining to situation assessment, user state assessment, adaptation engine, and HMI. There are several knowledge models to support each of these modules.

1) *Situation Assessment* is concerned with the assessment of the "situation", and comprises functionality relating to the real-time analysis of the activities (required to achieve a specific goal), automation, and decision support. The module monitors and tracks the current progress toward a specific activity, goal, or status through the data sensing and fusion from internal (machine status) and external data sources by using task, goal, and situational knowledge. The knowledge is used for intelligence generation and the adaptation engine to decide appropriate strategies to assist a user through decision support and adaptive automation, or by adapting what information is presented to the user through HMI.

Underpinning this module are task, machine, and world models. A task model contains knowledge pertaining to the tasks that a user is expected to perform and is represented as an organization of actions, goals, and plans. A machine model includes knowledge related to the machine itself, its abilities, and the means of assistance to support a user including advice, automation, and interface adaptation. A world model defines the external world according to the objects that exist in the working environment, their properties, and the rules that govern them.

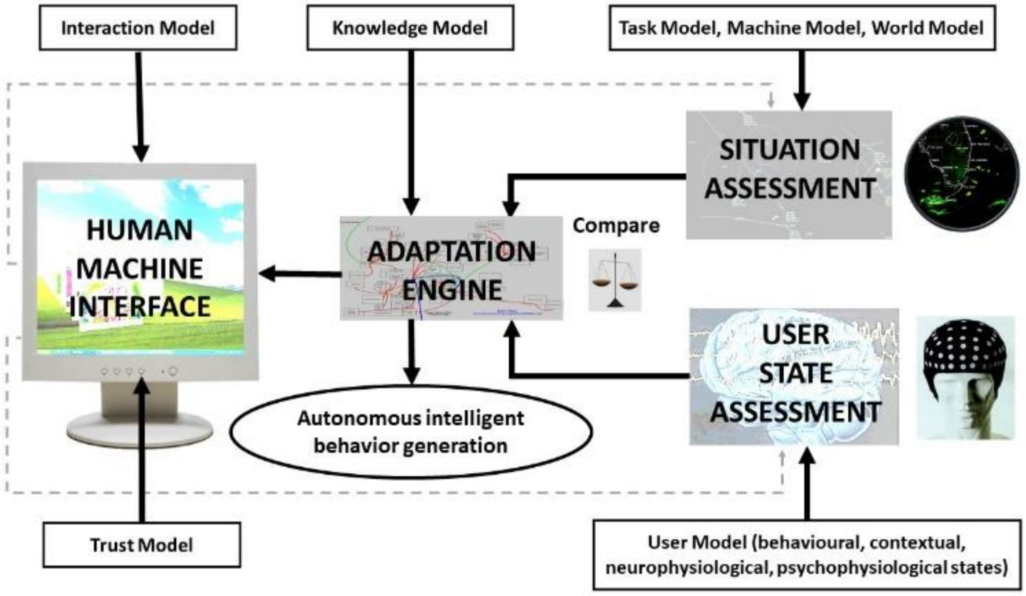


Fig. 1. The architectural framework of a generic intelligent adaptive system (IAS) with four modules and their supporting models.

This article has been accepted for publication in IEEE Systems, Man, and Cybernetics Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MSMC.2021.3136983

2) *User State Assessment* provides information and knowledge about a user's behavioral, contextual, neurophysiological, and psychophysiological states within the context of a specific work activity. A supporting user model incorporates the knowledge, skills, or behaviors that embody a specific user performing a specific task based on the mechanism of human cognition, control abilities, and communications. Through real-time analysis of a user's interactions, the module updates the system knowledge about the modeled state of a user's attention, engagement with the tasks, ongoing cognition (visual and verbal processing load), emotion, intention, performance, and competency. The knowledge provides a basis for the adaptation engine, which computes and compares the current user states with the built-in knowledge model of the user to assess the user's deficiencies (computational interaction for anticipatory intelligence). Then, the intelligent adaptation of the user assistance may be autonomously triggered to enhance and mitigate human information processing and decision-making capabilities and limitations.

3) *Adaptation Engine* employs high-level knowledge outputs from the user state and situation assessment modules to generate autonomous intelligent behaviors. It seeks to maximize the system performance with its functions to prioritize and optimize task allocations (to either the user or the machine or both) and information management (presenting only critical information in the right format and at the right time during periods of excessively high-workload levels through the HMI). A knowledge model supports this module with its abstract representations of the application domain, task, user, machine, and world. It provides baseline conditions for the comparisons between the expected outcomes and the current status of all these variables. Differences or deficiencies are used to drive the intelligent adaptation process to optimize human-machine interactions and assist in achieving overall system goals.

4) *Human-Machine Interface (HMI)* monitors, updates (the models), and communicates (with the machine) real-time user behavioral, neurophysiological, and psychophysiological changes such that the system may assess a user's state of overall cognitive resources including attention, engagement, emotion, and workload. Two models supporting this module are interaction model and trust model. An interaction model includes knowledge related to the mode of communications and interactions between a user and machine, as well as among system components. A trust model contains knowledge related to optimal states of variables affecting trust between a user and machine. Through real-time data collection about human-machine interactions, the changes of the user, task, machine, and world states and those trust factors can be computed (computational interaction for anticipatory intelligence). The computational results drive the adaptation engine to optimize human interactions with the machine through intelligent adaptation of machine behavior and information presentation. The computational results also update the trust levels and SA for a user and machine about the current system status such that decisions can be made for further actions to maintain, repair, or regain the desired trust (trust calibration and assurance).

Overall, the four modules operate within the context of a closed-loop system. In each module, a closed-feedback loop resamples user state and situation assessment to update all knowledge models following the adaptation of the machine or HMI. Thus, an IAS may adjust the level of adaptation such that optimal user states (attention, engagement, performance, and workload) and trust are attained, maintained, and assured. This conceptual framework provides a complete snapshot of the user, machine, task, and environment they interact with, which is crucial to the design, development, verification, validation, certification, and exploitation of complex brain-inspired IASs.

### *B. Interaction-Centered Design (ICD) for IAS Development*

Systems designers should be aware of user requirements and preferences when designing automation and interface as the two basic components of a human-machine system. Conventional automation and interface are designed following a technology-centered strategy based on task models, as illustrated in Fig. 2. To reduce user workload and increase task efficiency and productivity, designers use a task model to preplan and predesign automation functions with understanding of user requirements and preferences. With the emphasis on how technology advances may help allocate additional tasks to automation, technology acts as an assistant to a user. Meanwhile, the user does not take over the automation's tasks. Automation capability is derived from the leftover and compensatory principles. Basically, humans need to compensate with the functions that have not been automated or that could not be automated due to the machine limitations [19].

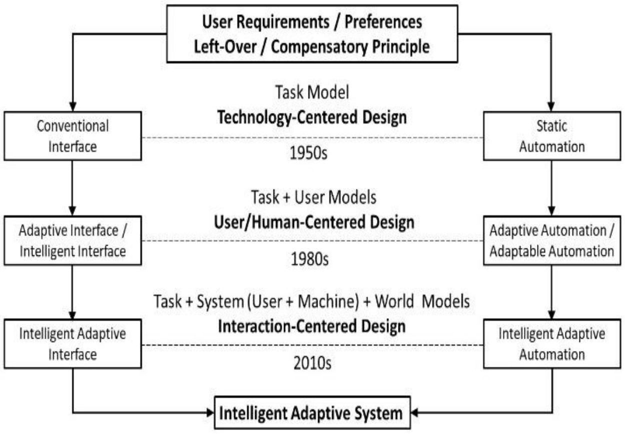


Fig. 2. Evolution of a design strategy for interface and automation technologies as two critical components of an intelligent adaptive system (IAS) from technology-centered to user-centered (UCD) and then to interaction-centered (ICD) design principles.

When more tasks are allocated to automation, humans have difficulty in addressing performance issues such as loss of SA (out-of-the-loop), loss of skills, overtrust (complacency), or undertrust (skepticism) [9], [20], [21], [22]. Thus, knowledge of a task model has been extended to include a user model of how and what humans are doing such that the system may

This article has been accepted for publication in IEEE Systems, Man, and Cybernetics Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MSMC.2021.3136983

provide flexible support. This advanced the design strategy from a technology-centered design approach to a user-centered design (UCD) or human-centered design (HCD) approach in the 1980s with the focus shift from technological capabilities onto human needs [23], [24]. The goal of HCD is to create designs (of products, services, workspaces, systems, procedures, organizations) that take into account the needs, capabilities, and limitations of those who are using or being impacted by the design for users' acceptance [24]. The need to assist humans in a flexible fashion has subsequently fostered the development of adaptive or adaptable automation, as well as adaptive or intelligent interface technologies [9], as shown in Fig. 2.

However, the human is only one of the many attributes of a broader human-machine system. The design should emphasize a system as a whole. Vicente [25] has posited that a UCD approach is not always ideal, arguing that a systems design perspective is more advantageous for correspondence domain applications (safety and mission-critical systems) [9]. For example, in the aviation, process control, and medical fields, as well as in warfare, a design flaw in a medical instrument or weapon system can have lethal and expensive consequences. The catastrophic disaster at the Three Mile Island nuclear power plant in Harrisburg, Pennsylvania, where a meltdown occurred in 1979, is a typical example that the design disregarded for appropriate design principles for correspondence domain applications [9] [26]. An issue with HCD has been further identified as local optimization that fails to consider the big picture and systems perspective [27]. Norman [27] then recommended modifications of insufficient HCD approach to address issues for correspondence domain applications with complex socio-technical systems.

As machines replace humans in a variety of tasks and slowly turn into independent entities, these issues regarding human-machine interactions come to the forefront [28]. Sheridan [1] suggests that as the frontiers between automation and humans blur, it becomes "increasingly critical" that automation designers realize that they are building not only technology, but also relationships. From a system of systems perspective, UCD or HCD is no longer sufficient to address broader human-machine interaction and relational issues, especially for domain applications with the socio-technical complexity [12], [25], [27]. Thus, the knowledge of task and user models has been broadened to include models of the intended machine and the world (i.e., working environment of an anticipated system) for fostering the development of both intelligent and adaptive automation and interface technologies guided by an Interaction-Centered Design (ICD) methodology [9], [29], as shown in Fig.2.

A primary goal of the ICD approach is to optimize human-machine interactions for IASs based on their joint capabilities, strengths, and limitations to maximize overall system performance, ensure safety, and enable trust. This requires the machine to be equipped with humanlike intelligence and behaviors so that the issue of human cognitive bottlenecks (limitations in attention, memory, learning, comprehension, visualization abilities, and decision-making) may be effectively addressed. With the humanlike cognitive capabilities of perceiving, reasoning, interpreting, and predicting the current and future status of a user, task, machine, and environment, the machine can predict human activities, awareness, intention, resources, and performance. It may then share responsibilities with its human partner and proactively assist in timely decision-making on task execution, automation adaptation and management strategy, system behaviors, and transfer of control and authority.

Shared responsibility exhibits functional integration of human and machine intelligence within human-machine symbiotic partnership, which is a key IAS humanlike characteristic [9], [29], [30]. Functional integration is extremely important in emergencies because it can create robust systems to handle unexpected events. For instance, safety redundancies should be built into aircraft control systems to allow an alternate course of action if a key component fails. If this strategy had been followed by the Boeing 737 Max design, the two catastrophic accidents in 2018 and 2019 could have been avoided even though the AI-enabled function failed.

The ICD methodology satisfies IAS design requirements and mitigates potential risks through detailed and comprehensive knowledge of a user, task, machine, and environment. IAS is essentially the unified evolution of an intelligent adaptive interface and intelligent adaptive automation into a hybrid system that features state-of-the-art automation and interface technologies. Fig. 2 illustrates the parallel evolution of a design strategy and principles for interface and automation technologies from technology-centered to UCD/HCD, and then to ICD. It also demonstrates the consistencies in their evolution and their eventual amalgamation into the IAS.

### *C. ICD Impacts on IAS Capability and Standard Development*

Over the past decade, the IAS framework, ICD approach, and a set of associated analytical and development methodologies have been applied to develop a variety of defence capabilities for the Canadian Armed Forces (CAF). One example is the first Canadian Intelligent Tutoring System (ITS) for Improvised Explosive Device Disposal (IEDD) operator training. An innovative intelligent adaptive learning system architecture was created based on IAS framework to guide the development of the ITS that enabled IEDD trainees to interact dynamically with training scenarios and receive real-time feedback on their questioning skill acquisition, resulting in an increased course success rate from 60% to 94% with reduced cost [31], [32]. Due to its novel IAS concepts based on ICD approach, this ITS technology has been filed for a patent application in Canada and the United States and exploited by CAE Inc. to create a commercial intelligent tutoring program for aviation training [33].

The ICD approach has also been applied for the development of the first Canadian Unmanned Aircraft System (UAS) Command and Control (C2) center that consists of a Ground Control Station (GCS) for supporting a Canadian major capital UAS acquisition project. A series of empirical studies

This article has been accepted for publication in IEEE Systems, Man, and Cybernetics Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MSMC.2021.3136983

conducted using this IAS capability has provided scientific evidence that informs the development of requirements for the project Request for Proposal, UAS GCS Workspace Optimization and Airworthiness Certification, and Operator Training Technology and Strategy [34] [35]. Another GCS has been deployed as a new trainer for the joint UAS operator training for Canadian Army, Navy, and Special Operations Forces. The ICD approach has also guided the development of Canadian Army Statement of Requirements for Micro UASs [36]. Thales Inc. has adopted related cognitive aspects of IAS in its AI program on Human Sensing Technology for intelligent adaptation based on operator intent and workload [30], [35].

The ICD paradigm has been recognized by the North Atlantic Treaty Organization (NATO) that adopts the systematic and structured IAS framework and ICD approach as a strategy with guiding principles for the development of NATO standards: “Guidance on Sense and Avoid for UAS” and “Human Systems Integration Guidance for UAS” to support the efforts of integrating UASs into the non-segregated civilian airspace [37], [38]. The related IAS design and development processes have been regarded as the validated best practices and advocated through the invited NATO Lecture Series on UAS technical challenges, concepts of operations, and regulatory issues [39], [40], [41].

## III. IMPACTS: A TRUST MODEL FOR HUMAN-MACHINE SYMBIOTIC PARTNERSHIP

IAS requires active human-machine interactivity at the highest level where both human and machine support each other proactively. It means that system control (authority) and responsibilities can be transferred safely between the partners whenever necessary. Thus, a trusted relationship needs to be built and maintained through dynamic interactions. Trust has then been identified as a key element and a “fundamental enabler” of a collaborative IAS decision-making capability [42].

Trust is a psycho-social relationship between entities or agents capable of acting. It is commonly understood as a cognitive process and a relational mediator for interactions between humans, human and organizations, and human-machine teaming. A variety of models examine trust factors and describe the development of trust in automation, cloud computing, blockchains, and Industry 4.0 smart manufacturing [4], [43], [44], [45]. However, these trust models, whereas comprehensive, may not consider dimensions related to the ever-increased AI capabilities and their impacts on the aggregated decision-making powers that reside in IASs.

### *A. IMPACTS Trust Model*

As machines evolve into highly autonomous functions of IASs with greater AI capabilities, the studies [42] and [46] have suggested that the human-machine trust relationship should more closely mirror the human-to-human trust to reflect the dynamic and complex nature of human-autonomy teaming. Trust is a function of capability and integrity for human relationships [42]. Human trust in technologies cannot exist without either of these two variables. Technologically, IASs with AI- enabled autonomous functions are more capable than their human partners in certain areas. To build up humanlike integrity, a model IMPACTS has been developed for IASs to exhibit seven characteristics to gain actual trust from humans [42]. As illustrated in Fig. 3, the defining integrity characteristics of the model are: shared *intention*, performance *measurability*, predictable and reliable *performance*, context *adaptivity*, bi-directional *communication*, optimal *transparency*, and protective *security*.

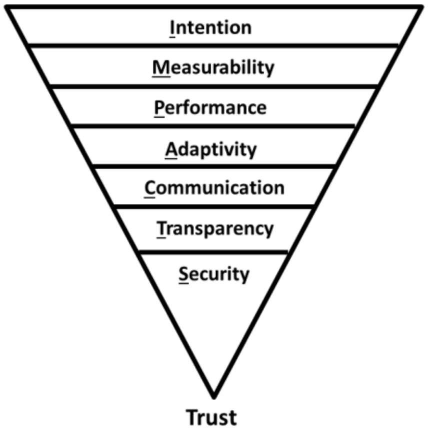


Fig. 3. An IMPACTS trust model to exhibit seven integrity characteristics of an IAS based on ICD methodology.

*1) Intention*: A collaborative partnership should have the desire for mutual support. The chosen desires are defined as intentions to which a machine commits efforts and resources for achieving the goals to help its human partners. The machine should know what humans are trying to achieve such that it may pursue actions that achieve its intention to help. Meanwhile, understanding a machine’s supportive intention towards a common goal serves as a starting point for humans to trust their machine partner. Thus, the design of IAS should allow the demonstration of shared intentions of humans and machines through their social interactions.

*2) Measurability*: The development of trust is a dynamic process and should not be undermined by a single instant. One may not trust words and may even question actions but should not doubt patterns. The same human-human relational philosophy can be applied to the human-machine partnership, where an AI entity may be complex and opaque. AI entities should be measurable such that their behaviors may be observed, their actions measured, and their behavioral patterns analyzed to gauge intentions. Thus, trust is developed through observable behaviors, measurable actions, or analyzable patterns through these types of computational interactions. The computational results and implications may then to be judged supportive or not.

*3) Performance*: Trust should be built, earned, maintained, and assured. The development of trust results from the reliable, consistent, and predictable performance of a partner over time [44]. In fact, performance is identified as the primary contributor for establishing trust for robots and significant for the establishment of trust for automation [47]. Performance entails a variety of attributes. For a machine to gain trust from

This article has been accepted for publication in IEEE Systems, Man, and Cybernetics Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MSMC.2021.3136983

its human partner, it must demonstrate performance that is valid (exhibiting as intended), reliable (being consistent over time), dependable (having few errors), and predictable (meeting human expectations).

*4) Adaptivity*: AI entities need to be capable of learning and understanding their human partners' intentions, the machine status and changes in the tasking environment, monitoring the human cognitive workload and performance, guarding the human resources and time, and changing their course of action to work with humans to achieve the team's common goals. An AI entity that exhibits these adaptive and intelligent characteristics may then be a trusted partner to build a truly collaborative human-AI symbiotic partnership. An effective IAS should enable machine adaptivity through dynamic interactions between humans and machines and thus foster the development, maintenance, and assurance of mutual trust.

*5) Communications*: Communication is a key to team success and needs to be bi-directional for humans and machines to learn and understand each other as partners. An effective and trustworthy IAS should enable AI entities and their human partners to clearly, fairly, and directly exhibit and justify their intentions, actions, and desired end-states and how they may help each other reach their common goals. These are critical characteristics that machines need to exhibit to build up human confidence and trust. An IAS also needs to be flexible and offers the types of feedback that the human partner would like to receive and accept such that effective communications happen at the right time, in the right format, through the right channel, and to the right recipient. If an IAS could facilitate such effective communications, an appropriate trust partnership would be enabled, maintained, and assured constantly.

*6) Transparency*: It has been said often that AI runs in a black box because it works in a fashion that humans do not fully understand and have no way of validating its intentions. The opacity of AI is problematic for trust development, maintenance, and assurance. Opacity has remained acceptable when the machine is reliable and is designed for simple tasks. Yet when AI is expected to perform complex tasks, involving multifaceted decisions in uncertain situations alongside humans with potentially vital consequences, transparency is crucial for humans to ascertain that AI entities' goals and methods for achieving those goals are aligned with shared intentions. It can be facilitated and optimized through the explainable HMI, which communicates real-time information to humans and machines regarding their intentions, goals, reasoning, decisions, actions, and expected outcomes.

*7) Security*: IAS capabilities are continuously enhanced by the growth of AI-enabled functionalities; however, it naturally increases system complexity. In engineering, complexity generally translates into uncertainty and risk, and this generalization applies to the design of IASs. The design should protect the system from accidents or deliberate threats (e.g., a cyber attack) and thus build human trust. A secured system must behave as designed and implemented following rules or laws even when it is under attack; otherwise, it cannot gain or assure trust. This insight requires designers and organizations to build confidence in IAS technologies by providing goal-directed explanations of security measures in place (i.e., at the right level of detail) to protect and ensure system safety and performance.

According to the IMPACTS model, trust is a vertex and a crucial ingredient for collaborative human-machine symbiotic partnership. It is the careful balance on which healthy relationships grow and are maintained and assured among partners when considering physical, intellectual, emotional, ethical, moral, relational, and even spiritual aspects of human beings. For technologies to be truly trustworthy, consistent, predictable, reliable, and demonstrate humanlike integrity with shared intentions through their adaptive behaviors and measurable performance, transparent communications, and secured protection are indeed IMPACTS that only humans can make with their AI partners. Researchers and practitioners developing IASs must carefully design them to inspire confidence and build trust, and the IMPACTS model is a conceptually practical tool to guide the design and development of collaborative and trusted human-AI symbiotic partnerships.

### *B. IMPACTS for Trustworthy IAS Development and Acceptance*

To address complex, lengthy, and error-prone target engagement processes, the IMPACTS model has been applied for the development of an IAS called Authority Pathway for Weapon Engagement (APWE) as an intelligent decision support system [48]. APWE automatically streamlines engagement processes and enables operators to visualize the dynamic engagement status intuitively through its intelligent, adaptive HMI, thereby reducing engagement times and errors while enhancing operators' SA.

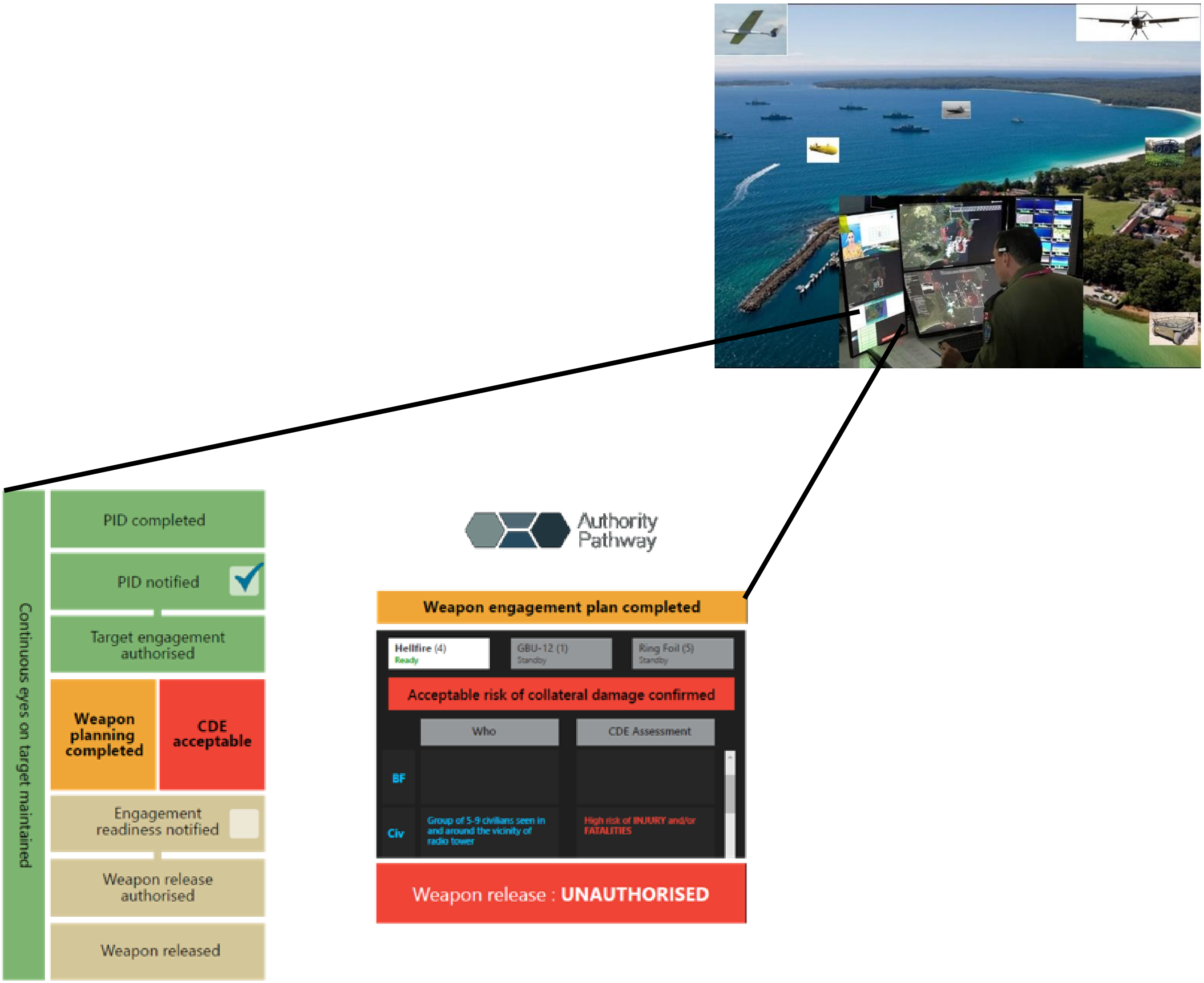
One of the critical computational interaction capabilities of APWE is its function that automatically generates system intelligence with gathered information and knowledge about operators' states, the tempo of mission tasks, assets, and entities in the battlespace. For example, it calculates, verifies, and advises whether a potential strike (i.e., lethal vs non-lethal) follows the appropriate legal and ethical policies (e.g., laws of armed conflict, rules of engagement, and standard operating procedures) before decisions are made by human authorities. APWE is an integral part of a joint Command and Control (C2) system that consists of a variety of highly automated emerging and disruptive technologies (EDT) from four nations [49]. The utility, effectiveness, and interoperability of these EDTs including APWE were assessed in a large-scale and complex international exercise with a context of a single operator controlling multiple heterogeneous unmanned systems in the air, on the ground, at the sea, and under the water, as illustrated in Fig. 4.

The evaluation results have revealed that military participants identify APWE as one of the top three strengths of the joint C2 system and a significant contributor to the success of the exercise. Its implementation has been considered as "exemplary, with major enhancements" because it "takes a lot of stress away from the operator". More importantly, it is the most trustworthy disruptive technology of the entire joint C2 system because APWE provides increased SA with reduced

This article has been accepted for publication in IEEE Systems, Man, and Cybernetics Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MSMC.2021.3136983

workload and potential human errors, according to [42], [49].

Given the dynamic, complex, and cognitively challenging nature of many mission-critical military systems with AI-enabled autonomous functions, the IAS framework, ICD approach, and IMPACTS model have been employed to support a series of CAF capability and concept development and evaluation activities. These activities provided various venues to validate the IAS paradigms for broader defence and civil applications such as autonomous transportation, homecare and surgical robots, and Industry 4.0 smart manufacturing [3], [4], [5]. Meanwhile, the innovative IAS framework and ICD approach are praised by academic experts, industrial practitioners, government authorities, and users from operational communities for their novelty and trend-setting initiatives of human-AI symbiotic partnership. The ICD principles are referred to as "a must read (consideration) for any serious professional in academia, government, or industry" and an excellent guide to the design of "twenty-first century human-computer symbiosis technologies" [9]. It is noted for setting the agenda for the coming years as human factors practitioners grapple with the demands that IAS will make on its operators and outlining how collaboration and partnership between human and AI can be achieved through ICD, according to [50]. IAS broad acceptance, significant impact, and exceptional contribution have been recognized by the Department of National Defence, Canada and the Professional Institute of Public Service of Canada with the prestigious Science and Technology Excellence Award and The President's Achievement Award in 2020 and 2021, respectively.



| | Who | CDE Assessment |
|---|---|---|
| BF | | |
| Civ | Group of 5-9 civilians seen in and around the vicinity of radio tower | High risk of INJURY and/or FATALITIES |

Fig. 4. APWE was integrated in an allied Command and Control system, demonstrated, and evaluated during an international exercise where one operator controlled multiple heterogeneous unmanned systems in the air, on the ground, at the sea, and under the water. The IAS framework, ICD design methodology, and IMPACTS trust model have guided the design and development of an AI-enabled decision-aid, APWE, to address the complex, lengthy, and error-prone target engagement process. The weapon cannot be released unless APWE completes all steps successfully based on built-in rules of engagement, international laws of armed conflicts, and standard operating procedures.

This article has been accepted for publication in IEEE Systems, Man, and Cybernetics Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MSMC.2021.3136983

## IV. FUTURE WORK IN SYMBIOTIC HUMAN-ROBOT TEAMING AND SYSTEM VERIFICATION REGIMES

One of the main strengths of the ICD approach is its systematic and structured process with stakeholder involvement and identification of system requirements and critical decisions with associated tasks for socio-technical systems in correspondence domain applications. It addresses a key challenge of system design: the need to elucidate, develop, and validate operational requirements that are obscured by the complexity of human-machine interactions and of the system itself. This has been demonstrated with the validation studies on a number of military systems discussed above and adopted by NATO standards. The next step is to systematically integrate the ICD paradigm into defence policies or acquisition processes and applying the systems engineering approach for defining, designing, developing, testing, acquiring, and employing IAS technologies.

The IMPACTS trust model is a practical, conceptual component supporting the HMI module in the IAS architecture as shown in Fig. 1. It is an integral part of the ICD approach and is being exploited as a systems engineering analysis and design tool for a trust management system (TMS) in a context of Soldier-Robot Teaming (SRT). Meanwhile, a TMS related mathematical matrix is also being developed to measure trust in real-time during human interactions with various autonomous systems. These mechanisms can then be implemented and integrated into SRT technologies for a series of military exercises. Once validated through these operational trials, these paradigms should make significant impacts on the systematic design and validation methodology for enabling trust.

The focus on human-machine trust relationships thus far has mostly been on the human's trust in the machine. However, in order to fully consider the safety property of a human-machine symbiotic partnership, additional trust relationships of machine-to-machine and machine-to-human (i.e., does the machine trust the human's decision-making or judgement) must be considered. A potential risk might be machine's blind trust in human decisions without knowing the decisions made under the impacts of logical and emotional trust attributes such as stress or bias [42]. From a computation perspective, machine learning may be rigorous without introducing bias if its algorithm is trustworthy. That is, cognitive and behavioral bias is often caused by interference inconsistency between machines and humans. Hence, overcoming bias potentially requires bidirectional communication and a comprehensive mechanism through overlaid interactions.

These mutual trust relationships are therefore suited for representing a more comprehensive and complete trust partnership. Thus, additional studies need to be conducted to understand and develop strategies for managing them simultaneously in real-time. For example, how does a brain-inspired machine learn to trust or assess confidence in human judgement or decisions, or what should be done when it does not trust the human while the human does not trust the machine? Trust mediation has yet to be addressed and should be studied.

Further, legal and ethical issues concerning the use of highly automated systems have been identified in [2], [12], including a sensitive topic when considering safety and mission-critical weapon systems [6]. These issues include the possibility that a system with autonomous functions may purposely and deliberately withhold information concerning a system failure, malfunction, or error.

The question has been raised as to whether certain trust repair strategies are ethical and more work is needed to address these issues. One area is to integrate both a TMS for measure of trust (MoT) and an ethical design review (EDR) in verification regimes such as formal systems engineering processes and industry production standards for autonomous transportation, homecare and surgical robots, and Industry 4.0 smart manufacturing [15] [51]. This may include the analyses of trust and ethics requirements during a system design process and then quantitative measures of the trade-offs (processing time, memory use, performance, potential misuse, bias, etc.) during a test and evaluation process.

Accordingly, a "trust certificate" and/or "ethics certificate" can be issued if test results are satisfactory through socio-technical verification and validation system-of-systems processes. Therefore, a framework for understanding trust and ethics in the context of verification regimes needs to be identified or developed to guide the implementation of the MoT and EDR in the verification process for optimal social benefits and impacts. To support this endeavor, the IAS framework and associated ICD approach and IMPACTS model have provided a sophisticated architecture, systematic methodology, and structured process for developing brain-inspired autonomous systems.

## V. CONCLUSION

At the frontier of brain-inspired autonomous systems, IASs have emerged to exhibit collective intelligence enabled by optimized human-machine interactions based on their joint capabilities and strengths to achieve shared goals. The IAS technology enables trustworthy solutions of human cognitive bottlenecks and AI's incompetency under indeterministic conditions or with insufficient data. The IAS framework, ICD methodologies, and IMPACTS trust model have been validated through a series of defence concept development and evaluation activities in large-scale international military exercises. These defence research and development successes have been recognized by military, academic, industry, and government authorities as well as international organizations for satisfying the growing demands for brain-inspired autonomous systems. Novel computing theories, process capabilities, and validation mechanisms will be developed for broader innovations in the defence and general industries.

## About the Authors

***Ming Hou*** (ming.hou@forces.gc.ca) received his Ph.D. from the University of Toronto, Canada in 2002. He is a Senior Defence Scientist and the Principal Authority of Human-

This article has been accepted for publication in IEEE Systems, Man, and Cybernetics Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MSMC.2021.3136983

Technology Interactions within the Department of National Defence (DND), Canada. He is responsible for delivering technological solutions, science-based advice, and evidence-based policy recommendations to senior DND decision makers and their national and international partners. Dr. Hou is an Integrator of the Canadian government $1.6B IDEaS program with responsibilities for guiding national R&D activities in AI, Robotics, and Telepresence. Dr. Hou is the Co-Chair of Human Factors Specialist Committee within NATO Joint Capability Group Unmanned Aircraft Systems.

***Yingxu Wang*** (yingxu@ucalgary.ca) received his Ph.D. in Computer Science from Nottingham Trent Univ., UK, in 1998. He is a full professor in Department of Electrical and Software Engineering and Hotchkiss Brain Institute at University of Calgary, Canada. He has held visiting professorships at Universities of Oxford (1995|2019-2022), Stanford (2008|2016), UC Berkeley (2008) and MIT (2012). He is the founding President of International Institute of Cognitive Informatics/Computing (I2CICC). He has published 600+ papers and 36 books and been recognized by Google Scholar as world’s top 1 in Software Science, Cognitive Robots, and Knowledge Science. He is a Fellow of IEEE, CCS, I2CICC, AAIA, and WIF.

***Ljiljana Trajkovic*** (ljilja@sfu.ca) received her Dipl. Ing. degree from University of Pristina, Yugoslavia, M.Sc. degrees in Electrical Engineering and Computer Engineering from Syracuse University, and Ph.D. in Electrical Engineering from University of California at Los Angeles. She is a professor at Simon Fraser University, Canada, and serves as Editor-in-Chief of the IEEE Transactions on Human-Machine Systems. She served as IEEE Division X delegate/director, IEEE SMC Society President, and IEEE Circuits and Systems Society President. Her research interests include communication networks and dynamical systems. She is an IEEE Fellow and Distinguished Lecturer of the IEEE SMC Society and Circuits and Systems Society.

***Konstantinos N. Plataniotis*** (kostas@ece.utoronto.ca) is a Professor and the Bell Canada Chair in Multimedia at the University of Toronto. His research interests are in machine learning and signal processing, and their applications in imaging systems, communications systems, and knowledge media design systems. Dr. Plataniotis is a Fellow of IEEE, the Engineering Institute of Canada, the Canadian Academy of Engineering / L’ Academie Canadienne Du Genie, and a registered Professional Engineer in Ontario. He is the General Co-Chair of the 2027 IEEE International Conference on Acoustics, Speech and Signal Processing.

***Sam Kwong*** (cssamk@cityu.edu.hk) received his B.Sc. degree from the State University of New York at Buffalo, M.A.Sc. in electrical engineering from the University of Waterloo in Canada, and Ph.D. from Fernuniversität Hagen, Germany. He is the associate editor of leading IEEE transaction journals, such as IEEE Transactions on Evolutionary Computation, IEEE Transactions on Industrial Informatics, and IEEE Transactions on cybernetics. He has coauthored three research books, eight book chapters, and over 400 technical papers. Dr. Kwong is an IEEE Fellow and serves as 2022-2023 IEEE SMC Society President.

***MengChu Zhou*** (mengchu.zhou@njit.edu) received his B.S. degree from Nanjing University of Science and Technology, China in 1983, M.S. degree from Beijing Institute of Technology, China in 1986, and Ph. D. degree from Rensselaer Polytechnic Institute, Troy, NY in 1990. He then joined New Jersey Institute of Technology and now is a Distinguished Professor. His research interests are in Petri nets, intelligent automation, Internet of Things, and big data analytics. He has over 900+ publications including 12 books, 600+ journal papers (500+ in IEEE transactions), 29 patents and 29 book-chapters. He is a Fellow of IEEE, IFAC, AAAS, CAA, and NAI.

***Edward Tunstel*** (tunstel@ieee.org) received B.S. and M.E. degrees from Howard University, Washington, DC and the Ph.D. degree from University of New Mexico, Albuquerque, NM in 1996. He served as a leader in robotics and autonomy with NASA Jet Propulsion Laboratory for nearly two decades, Johns Hopkins Applied Physics Laboratory for a decade and Raytheon Technologies Research Center for several years. He is CTO with Motiv Space Systems, Inc. His interests span research through deployment of robotic and autonomous systems, having authored 170 publications in those areas. He is an IEEE Fellow and served as 2018-2019 IEEE SMC Society President.

***Imre J. Rudas*** (rudas@uni-obuda.hu) received his Ph.D. in Robotics and Doctor of Science from the Hungarian Academy of Sciences in 1987 and 2004, respectively. He received Doctor Honoris Causa degree from the Technical University of Košice, Slovakia, “Polytechnica” University of Timisoara, Romania, Óbuda University, and Slovak University of Technology in Bratislava. He is the founder and Professor Emeritus of Óbuda University. He has published 22 books and 850+ papers with 7000+ citations. His research interests include Computational Cybernetics, Robotics, and Computational Intelligence. Professor Rudas is a Fellow of IEEE and served as 2020-2021 IEEE SMC Society President.

***Janusz Kacprzyk*** (kacprzyk@ibspan.waw.pl) is a Full Professor at the Systems Research Institute, Polish Academy of Sciences, and Full Member of European Academy of Sciences and Arts, Spanish Royal Academy of Economic and Financial Sciences, and Bulgarian Academy of Sciences. He is the recipient of 6 honorary doctorates. His research interests include computational and artificial intelligence, and fuzzy logic in systems, decision, and control. He authored 7 books, 150 volumes, 650+ papers, including ca. 100 in WoS journals. He is the Editor-in-Chief of Springer’s 7 book series. He is listed in 2020 “World’s 2% Top Scientists”. He is a Fellow of IEEE, IET, IFSA, EurAI, AAIA, SMIA, and the President of the Polish Operational and Systems Research Society.

***Henry Leung*** (leungh@ucalgary.ca) is a professor of Electrical and Computer Engineering at the University of Calgary, Canada. He has over 300 journal papers in the areas of signal and image processing, data mining, information fusion, machine learning, IoT and sensor networks. He also holds more than 15 patents. Dr. Leung is the Editor of the Springer book

This article has been accepted for publication in IEEE Systems, Man, and Cybernetics Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MSMC.2021.3136983

series on “Information Fusion and Data Science”. He has been an Associate Editor of various journals such as International Journal on Information Fusion, IEEE Transaction on Aerospace and Electronic Systems, IEEE Signal Processing Letters, IEEE Circuits and Systems Magazine. He is a Fellow of IEEE and SPIE.

## References

[1] T. B. Sheridan, Humans and Automation: System Design and Research Issues, Santa Monica, CA, US: Wiley-Interscience, 2002.

[2] J.-F. Bonnefon, A. Shariff and I. Rahwan, "The social dilemma of autonomous vehicles," *Science* , vol. 352, no. 6293, pp. 1573-1576, 2016.

[3] A. Lacher, R. Grabowski and S. Cook, "Autonomy, Trust, and Transportation," in *AAAI Spring Symposium Series*, 2014.

[4] M. Harlamova and M. Kirikova, "Towards the Trust Handling Framework for Industry 4.0 , Ebook Volume S," in *Frontiers in Artificial Intelligence and Applications*, vol. 315: Databases and Information Systems X, Ebook, 2019, pp. 49-64.

[5] F. Mannhardt, S. A. Petersen and M. F. Oliveira, "A trust and privacy framework for smart manufacturing environments," *Journal of Ambient Intelligence and Smart Environments,* vol. 11 (3), pp. 201-219, 2019.

[6] H. M. Roff and R. Moyers, "Meaningful human control, artificial intelligence and autonomous weapons," in *Briefing paper prepared for the Informal Meeting of Experts on Lethal Autonomous Weapon Systems*, UN Convention on Certain Conventional Weapons, 2016.

[7] K. Bekiroglu, S. Srinivasan, E. Png, R. Su and C. Lagoa, "Recursive approximation of complex behaviours with IoT-data imperfections," *IEEE/CAA Journal of Automatica Sinica,* vol. 7, no. 3, pp. 656-667, 2020.

[8] Y. Wang, I. Pitas, K. Plataniotis, C. Regazzoni, B. Sadler, A. Roy-Chowdhury, M. Hou, A. Mohammadi, L. Marcenaro, F. Atashzar and S. alZahir, "On future development of autonomous systems: a report of the plenary panel at IEEE ICAS’21," in *1st IEEE Conference on Autonomous Systems*, 2021.

[9] M. Hou, S. Banbury and C. Burns, "Intelligent adaptive systems – an interaction-centered design perspective," CRC Press, Boca Raton, FL, 2014.

[10] Majority Staff of the U.S. House Committee on Transportation and Infrastructure, "Final Committee Report: the design, development, and certification of the Boeing 737 Max," 2020.

[11] N. T. S. Board, "Loss of Thrust in Both Engines, US Airways Flight 1549 Airbus Industrie A320-214, N106US," 2010.

[12] E. Awad, S. Dsouza, R. Kim, J. Schulz, J. Henrich, A. Shariff, J. F. Bonnefon and I. Rahwan, "The moral machine experiment," *Nature*, vol. 563, pp. 59-64, 2018.

[13] A. Girma, N. Bahadori, M. Sarkar, T. G. Tadewos, M. R. Behnia, M. Mahmou, A. Karimoddini and A. Homaifar , "IoT-enabled autonomous system collaboration for disaster-area management," *IEEE/CAA Journal of Automatica Sinica,* vol. 7, no. 5, pp. 1249-1262, 2020.

[14] Y. Wang, M. Hou, K. Plataniotis, S. Kwong, H. Leung, E. Tunstel, I. Rudas and L. Trajkovic, "Towards a Theoretical Framework of Autonomous Systems Underpinned by Intelligence and Systems Science," *IEEE/CAA Journal of Automatica Sinica,* vol. 8, no. 1, pp. 52-63, 2021.

[15] Well-being Metric for Autonomous and Intelligent Systems Working Group, *IEEE 7010-2020 - IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being,* IEEE Systems, Man, and Cybernetics Society and Standards Association, 2020.

[16] Y. Wang, F. Karray, S. Kwong, K. Plataniotis, H. Leung, M. Hou, L. Trajkovic, O. Kaynak, J. Kacprzyk, M. Zhou, P. Chen and S. Dilip, "On the philosophical, cognitive and mathematical functions of symbiotic autonomous systems," *Philosophical Transactions A Mathematical Physical and Engineering Sciences,* vol. 39, no. 2207, pp. 1-20, 2021.

[17] D. D. Schmorrow and A. A. Kruse, "Augmented cognition," in *Berkshire Encyclopedia of Human-Computer Interaction*, Great Barrington, Massachusetts U.S.A, Berkshire Publishing Group LLC, 2004, pp. 54-59.

[18] Y. Wang, "Cognitive informatics: A new transdisciplinary research field," *Brain and Mind,* vol. 4, pp. 115-127, 2003.

[19] P. M. Fitts, Human engineering for an effective air-navigation and traffic-control system, National Research Council, Division of Anthropology and Psychology, Committee on Aviation Psychology, 1951.

[20] M. R. Endsley and E. O. Kiris, "The out-of-the-loop performance problem and level of control in automation," *Human Factors,* vol. 37, no. 2, pp. 381-394, June 1995.

[21] R. Parasuraman and D. H. Manzey, "Complacency and bias in human use of automation: An attentional integration," *Human Factors,* vol. 52, no. 3, p. 381–410., 2010.

[22] P. de Vries, C. Midden and D. Bouwhuis, "The effects of errors on system trust, self-confidence, and the allocation of control in route planning," *International Journal of Human-Computer Studies,* vol. 58, no. 6, pp. 719-735, 2003.

[23] D. A. Norman and S. W. Draper, User Centered System Design: New Perspectives on Human-

This article has been accepted for publication in IEEE Systems, Man, and Cybernetics Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MSMC.2021.3136983

Computer Interaction, Hillsdale, N.J.: Lawrence Erlbaum Associates, 1986.

[24] *Human-centred design processes for interactive systems,* ISO 13407, 1999.

[25] K. J. Vicente, "Coherence- and correspondence-driven work domains: implications for systems design," *Behaviour & Information Technology,* vol. 9, no. 6, pp. 493-502, 1990.

[26] C. D. Wickens, Engineering psychology and human performance, 2nd ed., New York: Harper Collins, 1992.

[27] D. A. Norman, "Human-centered design considered harmful," *Interactions,* vol. 12, no. 4, pp. 14-19, 2005.

[28] J. Rasmussen, Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering, New York: Elsevier, 1986.

[29] M. Hou, H. Zhu, M. Zhou and R. Arrabito, "Optimizing operator-agent interaction in intelligent adaptive interface design: a conceptual framework," *IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews,* vol. 41, no. 2, pp. 161-178, 2011.

[30] M. Hou, "Intelligent Adaptive System: A Strategic Perspective on Delivering Science Technology and Innovation on Effective Human-AI/Autonomy Integration," DRDC-RDDC-2021-B020, Defence Research and Development Canada, 2021.

[31] M. Hou and C. M. Fidopiastis, "A generic framework of intelligent adaptive learning systems: From learning effectiveness to training transfer," *Theoretical Issues of Ergonomics Science,* vol. 18, no. 2, pp. 167-183, 2016.

[32] M. Hou, C. Kramer, C. Banbury, M. Lepard and K. Osgoode, "Questioning technique review and scenario specification for the CF counter-IED operator training course," TR 2013-061, Defence Research and Development Canada, 2013.

[33] M. Hou and CAE Professional Services. Inc., "QuestionIT (intelligent tutor for questioning technique)," 1416-12/016CA, Defence Research and Development Canada, 2014.

[34] M. Hou, R. D. Kobierski and M. Brown, "Intelligent Adaptive Interfaces for the Control of Multiple UAVs," *Journal of Cognitive Engineering and Decision Making,* vol. 1, no. 3, pp. 327-362, 2007.

[35] R. Arrabito, M. Hou, S. Banbury, B. Martin, F. Ahmad and S. Fang, "A Review of Human Factors Research Performed from 2014 to 2017 in Support of the Royal Canadian Air Force Remotely Piloted Aircraft System Project," *Journal of Unmanned Vehicle Systems,* vol. 9, no. 1, pp. 1-20, 2020.

[36] M. Hou, G. Ho, R. G. Arrabito, S. Young and S. Yin, "Effects of display mode and input method for handheld control of micro-aerial vehicles for a reconnaissance mission," *IEEE Transactions on Human-Machine Systems,* vol. 43 (2), pp. 149-160, 2013.

[37] NATO FINAS Sense and Avoid Specialist Team, *NATO STANREC 4811: Guidance on Sense and Avoid for Unmanned Aircraft Systems,* NATO Standardization Office, 2018.

[38] NATO FINAS Human Factors Specialist Team, *NATO STANREC 4685: Human Systems Integration Guidance for Unmanned Aircraft Systems,* NATO Standardization Office, 2021.

[39] M. Hou, "Context-based and Interaction-centred Design for UAS GCS, NATO Lecture Series AVT-274 on Unmanned Air Vehicles: Technological Challenge, Concepts of Operations, and Regulatory Issues," Defence Research and Development Canada, DRDC-RDDC-E17-0410-1725, 2017-2019.

[40] M. Hou, "Training Needs Analysis for UAS Operators, NATO Lecture Series AVT-274 on Unmanned Air Vehicles: Technological Challenge, Concepts of Operations, and Regulatory Issues," Defence Research and Development Canada, DRDC-RDDC-E17-0410-1722, 2017-2019.

[41] M. Hou, "Human Systems Integration and its Guidance for UAS GCS Certification, NATO Lecture Series AVT-274 on Unmanned Air Vehicles: Technological Challenge, Concepts of Operations, and Regulatory Issues," Defence Research and Development Canada, DRDC-RDDC-E17-0410-1723, 2017-2019.

[42] M. Hou, G. Ho and D. Dunwoody, "IMPACTS: a trust model for human-autonomy teaming," *Human-Intellignent Systems Integration,* vol. 3, no. Special Issue on “Human-Autonomy Teaming in Military Contexts”, pp. 79-97, 2021.

[43] P. Zhang, M. Zhou and Y. Kong, "A Double-Blind Anonymous Evaluation-Based Trust Model in Cloud Computing Environments," *IEEE Transaction on Systems, Man, and Cybernetics: Systems,* vol. 51, no. 3, pp. 1805-1816, 2021.

[44] P. Zhang, M. Zhou, Q. Zhao, A. Abusorrah and O. Bamasak, "A Performance-Optimized Consensus Mechanism for Consortium Blockchains Consisting of Trust-varying Nodes," *IEEE Transaction on Network Science and Engineering,* vol. 8, no. 3, pp. 2147-2159, 2021.

[45] J. H. Cho, K. Chan and S. Adali, "A Survey on Trust Modeling," *ACM Computing Surveys,* vol. 48, no. 2, pp. 1-40, 2015.

[46] T. B. Sheridan, "Individual Differences in Attributes of Trust in Automation: Measurement and Application to System Design," *Frontiers in Psychology,* vol. 10, p. 1117, 2019.

[47] K. E. Schaefer, J. Y. C. Chen, J. L. Szalma and P. A. Hancock, "A Meta-Analysis of Factors Influencing

This article has been accepted for publication in IEEE Systems, Man, and Cybernetics Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MSMC.2021.3136983

the Development of Trust in Automation," *Human Factors,* vol. 58, no. 3, pp. 377-400, 2016.

[48] D. McColl, K. Heffner, S. Banbury, M. Charron, R. Arrabito and M. Hou, "Authority pathway: intelligent adaptive automation development for a UAS GCS," in *Human Computer Interaction Conference on Engineering Psychology and Cognitive Ergonomics: Performance, Emotion and Situation Awareness,* Vancouver, Canada, 2017.

[49] J. Bartik, A. Rowe, M. Draper, E. Frost, A. Buchanan, D. Evans, E. Gustafson, C. Lucero, V. Omelko, P. McDermott, S. Wark, M. Skinner, J. Vince, C. Shanahan, M. Nowina-Krowicki, G. Moy, L. Marsh, D. Williams, H. Pongracic, A. Thorpe, H. Keirl, M. Hou and S. Banbury, "TTCP Autonomy Strategic Challenge (ASC) Allied IMPACT Final Report," TTCP Technical Report TR-ASC-01-2020, DRDC-RDDC-2021-N062, Defence Research and Development Canada, 2020.

[50] C. Baber, "Book Review: Intelligent Adaptive Systems: an interaction-centered design perspective," *Ergonomics,* vol. 60, no. 10, pp. 1458-1459, 2017.

[51] Engineering Methodologies for Ethical Life-Cycle Concerns Group, *IEEE 7000-2021 - IEEE Model Process for Addressing Ethical Concerns During System Design,* IEEE Standards Association, 2021.