HHAI 2025
D. Pedreschi et al. (Eds.)

doi:10.3233/FAIA250624

Advancing Human-Machine Teaming: Definitions, Challenges, Future Directions

Ruben S. VERHAGEN [a,1], Mark A. NEERINCX [a,b], X. Jessie YANG [c], and Myrthe L. TIELMAN [a]
[a] *Delft University of Technology, Delft, The Netherlands*
[b] *TNO, Soesterberg, The Netherlands*
[c] *University of Michigan, Michigan, United States*

Abstract. Humans and intelligent machines increasingly collaborate on complex tasks, although significant challenges remain before machines can function as effective teammates. The human-machine teaming research community attempts to address these challenges by developing and testing methods that identify and enhance the factors essential for successful teaming. However, this community suffers from a lack of requirements for effective research, numerous methods without centralized documentation, and a disconnect between research and real-world applications. These challenges hinder progress and limit the generalizability of research outcomes. To address these issues, we argue that the human-machine teaming research community should establish a more structured and systematic approach to studying and advancing the field. This paper identifies and discusses several key research directions and actionable outputs for such an approach. These include taxonomies and guidelines to streamline research, team design patterns to describe reusable solutions, modular testbeds to facilitate comparability and reuse, and study templates to foster creativity and encourage sharing. We believe that these elements can help formulate requirements for effective human-machine teaming research and foster the development of modular and well-documented testbeds. Achieving these goals can contribute to more ecologically valid human-machine teaming research and, thus, a stronger connection between research and real-world applications.

Keywords. human-machine teaming, research methodologies, future directions

1. Introduction

Humans and intelligent machines increasingly collaborate on complex tasks, ranging from firefighting to manufacturing. As the capabilities of intelligent machines grow, they will increasingly act as full-fledged team members instead of tools. The ultimate goal of human-machine teams is combining the strengths of both parties to accomplish tasks that neither can do alone [1]. The success of these human-machine teams can be affected by many factors, such as mutual trust, coordination, and co-adaptation [2,3].

Despite ongoing advancements in human-machine teaming, significant challenges remain before machines can function as truly effective teammates for humans [4]. Intelligent machines often operate as opaque “black boxes”, making their inner workings and

[1]Corresponding Author: Ruben S. Verhagen, r.s.verhagen@tudelft.nl

behaviors difficult for humans to comprehend and trust appropriately [5,6,7,8]. Furthermore, these systems still lack the necessary knowledge, skills, and strategies to manage interdependencies with humans effectively [4]. To address these challenges, the human-machine teaming research community plays a critical role by developing and testing methods that identify and enhance the factors essential for successful teaming.

Currently, this community suffers from a lack of requirements for effective research, numerous methods without centralized documentation, and a disconnect between research and real-world applications. Therefore, in this work, we propose that the human-machine teaming research community prioritize establishing a more structured and systematic approach to studying and advancing the field. More specifically, by identifying the requirements for human-machine teaming research and emphasizing reusable and comparable methods and testbeds. We discuss several concrete research directions and actionable outputs for the community to focus on, such as taxonomies and guidelines, team design patterns, modular testbeds, and study templates. Ultimately, we believe these efforts can accelerate progress in the field and contribute to a common platform with essential tools and resources for human-machine teaming research.

2. Human-Machine Teaming (Research)

2.1. Background and Definitions

Several terms and concepts related to human-machine teaming exist, such as hybrid intelligence [1], human-centered artificial intelligence (AI) [9], and human-machine interaction [10]. In this subsection, we review these concepts in detail. It is important to note that we do not intend to endorse any single definition over the others. Instead, we aim to capture emerging trends and highlight the nuances we observe. First, many human-machine alternatives, such as human-agent/AI/automation/autonomy/computer/robot interaction, are used. We believe human-machine interaction encompasses all these alternatives, as they specify the type of machine interacting with a human. This interaction can be of any kind, complexity, or modality and does not even have to be goal-oriented.

Human-machine teaming is a type of human-machine interaction and can be defined as at least one human and machine working together toward a shared goal [2]. Interaction becomes teamwork when there is a degree of 1) interdependence between the activities and outcomes of humans and machines and 2) machine agency involving independence of actions and proactivity [10,11,12,13]. Human-machine teams focus on augmenting human and machine capabilities by combining the unique strengths of both parties to accomplish what neither could do alone [1]. These teams are generally involved in cognitive and/or physical work, where the former consists of mental or information processing activities, while the latter relates to manipulating tangible objects in the world [14]. Team members usually have explicit roles during this work, such as supervisor, performer, or supporter [2]. These roles often result in different interdependencies between humans and machines, such as required or optional collaboration [2]. The application domains of human-machine teams include emergency response, healthcare, manufacturing, and defense. For example, firefighters that teleoperate explore-and-extinguish robots because of mutually exclusive dependencies to extinguish and navigate [15].

Hybrid intelligence is another type of human-machine interaction and can be defined as combining human and artificial intelligence to augment their isolated operations

[1,16,17]. This type of interaction is also described as symbiotic artificial intelligence in other works [18]. Most existing works on hybrid intelligence adopt a technology-centric perspective when augmenting human and/or machine intelligence to achieve human or machine goals [16,19,20]. Hybrid intelligence systems are primarily involved in cognitive work, such as supporting human learning [21] and computer vision [22]. Another example is combining machine processing with human understanding and reasoning to extract arguments from opinions [19].

In contrast to this technology-centric perspective, recent research on hybrid intelligence emphasizes the collaboration between humans and AI toward shared goals [16,17,23]. This team-oriented perspective even frames hybrid intelligence systems as hybrid intelligence teaming, providing an overlap with human-machine teaming [10]. However, hybrid intelligence teams are primarily involved in cognitive work, such as complementing expertise for joint object identification [23] or determining temporary navigation destinations. Since hybrid intelligence teams are rarely involved in physical work, we consider them a subset of human-machine teaming.

Another type of human-machine interaction is human-centered artificial intelligence, which can be defined as augmenting human capabilities with embedded AI methods while ensuring human control [9]. This perspective places humans and their goals at the center, focusing on user needs, explainable systems, and meaningful human control [9,24,25,26]. Human-centered AI can augment human capabilities during both cognitive and physical work, such as augmented reality helmets to improve situational awareness or exoskeletons to enhance physical strength [9].

These examples illustrate machines that enhance human capabilities but not necessarily human intelligence. If human-centered AI systems augment human intelligence by integrating human and artificial intelligence, this can be considered hybrid intelligence. We define this overlap between hybrid intelligence and human-centered AI as human-centered hybrid intelligence. Like hybrid intelligence, human-centered hybrid intelligence is primarily involved in cognitive work, such as personalized AI support for firefighters by highlighting how mission characteristics differ from their experience [27].

In Figure 1, we visualize the relationships between all discussed concepts above. Moreover, in Table 1, we summarize the conceptual differences between these concepts. In summary, hybrid intelligence augments intelligence and focuses primarily on cognitive work, while human-machine teaming and human-centered AI can also augment other capabilities and focus on physical work. Furthermore, human-centered AI merely augments humans, while human-machine teaming and hybrid intelligence can augment both humans and machines. Finally, human-machine teaming merely tries to achieve shared team goals, while hybrid intelligence and human-centered AI can also try to achieve only human goals.

2.2. *Current State and Challenges*

Human-machine teaming is increasingly studied across disciplines such as computer science, engineering, and social sciences [28]. It is common to conduct human-machine teaming experiments in simulated and/or controlled environments because studying human-machine teams in the real world can be both time and cost-expensive. Several human-machine teaming testbeds have been used for such experiments, although only a few more frequently [28]. For example, in the Mixed Initiative eXperimental testbed, a

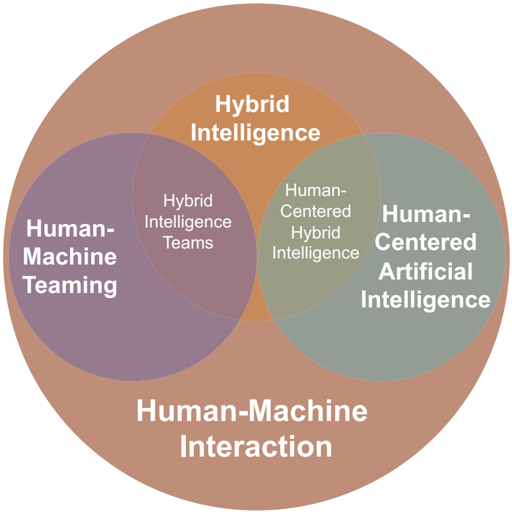

Figure 1. Venn diagram of human-machine interaction, hybrid intelligence, human-centered artificial intelligence, human-machine teaming, hybrid intelligence teams, and human-centered hybrid intelligence.

Table 1. Summary of the conceptual differences between human-machine interaction types.

Concept	Augments	Goals	Focus	Example
Human-Machine Teaming	Human capabilities Machine capabilities	Shared	Cognitive Physical	Robot teleoperation
Hybrid Intelligence	Human intelligence Machine intelligence	Human Machine Shared	Cognitive	Data annotation
Hybrid Intelligence Teaming	Human intelligence Machine intelligence	Shared	Cognitive	Joint decisions
Human-Centered Artificial Intelligence	Human capabilities	Human	Cognitive Physical	Augmented reality
Human-Centered Hybrid Intelligence	Human intelligence	Human	Cognitive	Personalized support

human operator detects targets in collaboration with a RoboLeader that collects information from subordinate robots with limited autonomy [29]. In the Cognitive Engineering Research on Team Tasks testbed, a team of three members, with specific roles such as navigator, photographer, and pilot, must collaborate effectively to achieve the team objective of capturing ground targets [30]. Finally, in the Blocks World 4 Teams testbed, participants must collaboratively deliver blocks in a specific sequence [31].

Some popular research topics facilitated by these testbeds include transparency and explainability [32,33,34], trust calibration and repair [35], and co-learning [36] in human-machine teams. For example, how interdependence relationships or communication modalities affect human-machine teaming [10,28]. In contrast, research and testbeds on multiple humans and machines, machine leadership roles, and communication methods beyond text are still lacking [28].

While these testbeds have facilitated significant progress, human-machine teaming research still faces key challenges. Creating realistic, reusable, and widely adopted testbeds has proven challenging despite the variety of testbeds. Currently, the field lacks (consensus on) effective and comparable research requirements. This often results in a disconnect between human-machine teaming research and its practical applications. Furthermore, it has resulted in numerous methods and testbeds without centralized documentation [37]. For example, most of the identified testbeds in [28] have only been used once or twice and emerged in the past ten years. Moreover, the modalities of these testbeds range from 2D grid worlds to 3D games and augmented to virtual reality. Team characteristics also often vary, especially regarding team composition, task interdependence, leadership structure, and communication structure [28].

2.3. Goals

The lack of widely adopted and highly diverse testbeds suggests that many human-machine teaming researchers try to reinvent the wheel. Instead, we argue that this research community should aim for more reusable, comparable, and modular methods and testbeds. The community already adopted this goal during a recent workshop[2] that we organized at the Lorentz Center (Leiden, The Netherlands) in the summer of 2024. Here, we established and discussed several community goals, such as developing human-machine teaming as a methodology and achieving consensus on the community's positioning. However, the ultimate goal is a common platform with essential tools and resources for human-machine teaming research. Ideally, this platform should contain research guidelines and requirements for experiment design, a library with reusable and comparable team design patterns, modular testbeds, and templates for describing and comparing studies. However, the requirements of such a platform should first be identified and formalized before actually building it.

3. Reusable, comparable, and modular human-machine teaming research

To achieve its goals, we believe the human-machine teaming research community should prioritize reusable, comparable, and modular research. Tasks, environments, and measures could be adapted across different studies and domains instead of reinventing the wheel and starting from scratch. The community should examine what has worked well for others and identify key requirements and customizable characteristics for human-machine teaming research. Moreover, it could benefit from templates for describing, comparing, and designing studies across domains and teams.

Most importantly, however, the community needs more modular methods and testbeds at a common platform to build upon existing research. Ideally, a platform where users can easily browse through modular testbeds that allow customization of tasks, machines, and teams to meet specific research needs. Before the community can achieve these goals, we believe it should first take a step back and study how to conduct reusable, comparable, and modular human-machine teaming research. This way, the requirements for such research can be identified and formalized, contributing to a more structured and systematic research approach. In the next subsections, we outline several concrete directions and outputs to focus on when establishing this approach.

[2] https://lorentzcenter.nl/research-environments-for-human-machine-teaming.html

3.1. Taxonomies and Guidelines

Identifying requirements and guidelines for human-machine teaming research can provide a baseline framework for researchers and support comparable studies. Therefore, we believe these to be a good starting point towards a more structured and systematic research approach. Such requirements and guidelines can also contribute to characterizing realistic and reusable human-machine teaming tasks, environments, and scenarios, providing a stronger connection between research and real-world applications. We believe literature reviews on human-machine teaming research to be crucial for this, such as the review on testbeds in [28] and independent and dependent variables in [10].

One particularly valuable contribution would be a literature review that results in a taxonomy of human-machine teaming tasks. This taxonomy could include task, machine, human, and team-related categories with examples from literature. Task-related categories could include domain, type, and goal; human and machine-related categories could include behavior and communication; and team-related categories could include design, roles, and interdependencies. These categories could even be structured according to the stages of experiment design at which they are defined. For example, determining the task domain, such as firefighting or warehousing, generally precedes decisions about team roles, such as supervisor, performer, or supporter. This taxonomy should help human-machine teaming researchers to identify and determine their task, machine, human, and team-related categories. Moreover, it would allow people to compare human-machine teaming studies more easily. Such a taxonomy with guidelines can be iteratively refined during community workshops. By establishing this baseline framework, human-machine teaming methods and research can become more reusable and comparable.

3.2. Team Design Patterns

Another desired output should be team design patterns that describe generic, reusable, and proven solutions for human-machine teaming [14,38]. These patterns can then be (re)used during experiment design and compared to others during user studies. Several team design patterns already exist with varying levels of abstraction, such as abstract patterns for AI advisors collaborating with humans [39]. However, more concrete patterns also exist, for example, for humans and machines collaborating in morally sensitive situations by allocating moral decisions to humans and non-moral decisions to machines [15,40,41]. Creating more team design patterns based on stakeholder involvement and realistic use cases can strengthen the connection between research and real-world applications, enhancing both translation and generalization [39]. For example, realistic team design patterns for firefighters collaborating with their explore-and-extinguish robots.

We believe the human-machine teaming research community would also greatly benefit from a library with all created team design patterns. Such a team design pattern library can facilitate reusing and comparing human-machine teaming methods. Ideally, a library that is divided between more abstract and concrete patterns and where more abstract and generalizable patterns can be created from more concrete ones. More concrete patterns could also be categorized into common and popular research topics, such as transparency and explainability, trust calibration and repair, and co-learning in human-machine teams. These more concrete patterns can describe generic reusable behaviors of humans and machines for supporting effective and resilient teamwork in these scenarios. For example, what information explore-and-extinguish robots should explain, and at which moments, to firefighter supervisors during semi-autonomous victim search tasks.

3.3. Modular Human-Machine Teaming Testbeds

Ultimately, these taxonomies and team design patterns should contribute to modular human-machine teaming testbeds accessible on a common platform. They can inform the community about which aspects of human-machine teaming testbeds can be standardized but configurable to allow comparisons, such as interdependence and role distribution. The community recently started moving in the right direction by sharing customizable testbeds. For example, a customizable testbed in a 2D grid world where participants collaborate with an autonomous, rule-based, explainable artificial moral agent during a firefighting task [42]. In addition to these fixed characteristics, the testbed is customizable with respect to explanation type (technical, ethical, or none) and artificial moral agency (low or high). Other important elements are currently fixed, such as explanation modality (hybrid) and role distribution (human supervision). Customizing these would further strengthen this testbed and facilitate reuse and comparisons with other studies. For example, allowing textual explanations or machine supervision. Ultimately, this can accelerate progress in the field of moral decision-making in human-machine teams.

Another example is a modular testbed in a 3D environment where participants collaborate with an autonomous, rule-based robot that violates their trust during a warehousing task [43]. This testbed allows researchers to alter robot reliability (100% or 70%), physical form (human- or machine-like), and trust repair strategy (apology, denial, promise, explanation, or none). Other important elements are currently fixed, such as trust violation type (competence-based) and communication modality (audio). Customizing these would further strengthen this testbed and enhance reusability and comparability. For example, adding integrity-based trust violations or textual communication.

This testbed is similar to studies in different environments on the influence of trust repair strategies [44] and interdependence [35] on human trust development. However, the former concerns a 3D first-person shooter game and the latter a 2D search and rescue task. The results of these studies would have been more comparable if performed in the same testbed. Moreover, building upon this customizable testbed prevents researchers from starting from scratch and reinventing the wheel. We believe this would accelerate progress in the field of trust calibration and repair in human-machine teams.

These recent customizable testbeds show that the community is moving in the right direction. However, these examples only allow for the adjustment of some machine-related settings. Instead, we consider modular testbeds to allow the adjustment of entire task, team, or machine-related components, such as modifying machines regarding autonomy, behavior, and communication. A starting point could be to create modular testbeds for the most common human-machine teaming topics, such as trust calibration and repair, and transparency and explainability in human-machine teams. Creating these testbeds based on stakeholder involvement, realistic use cases, and real-world applications could greatly benefit both research and practice. We believe hackathons at conferences could be great opportunities to work towards such modular testbeds.

3.4. Templates for Describing and Comparing Studies

Finally, the community could benefit from templates to describe, compare, and design studies across domains and teams. Such templates could be used on a common platform for contributors to facilitate sharing research and for users to foster creativity by browsing through and comparing prior research. They should prioritize essential information over supplementary details, such as task descriptions and variables over autonomy levels and hypotheses. So, key characteristics of human-machine teaming studies should be identified first. We believe the aforementioned taxonomies and guidelines can contribute to this identification, while team design patterns and testbeds can be template categories.

An example could be a template with dropdown options and tags for contributors to facilitate uploading their human-machine teaming research. This template could include the categories of team design pattern, task, and research environment. The team design pattern category could specify the number of humans and machines and their roles and relationships. The task category could specify the domain, task type, and interdependencies. Finally, the research environment category could specify the modality, measures, variables, machine embodiment, machine behavior, communication modality, and link to the testbed. Platform users could then enter search queries to find their studies of interest. For example, a user interested in VR studies on trust repair in triads with two humans and one machine. Such templates can facilitate reusable, comparable, and replicable human-machine teaming research. We believe they can be constructed based on literature reviews and domain expertise, and iteratively refined during workshops based on feedback from both researchers and practitioners.

4. Conclusion

We identified three significant challenges for the human-machine teaming research community: a lack of requirements for effective research, numerous methods and testbeds without centralized documentation, and a disconnect between research and real-world applications. To address these challenges, we proposed that the community establish a more structured and systematic research approach. We outlined four key research directions and actionable outputs of such an approach: taxonomies and guidelines to streamline research, team design patterns to describe reusable solutions, modular testbeds to facilitate comparability and reuse, and study templates to foster creativity and encourage sharing. We believe these elements can help formulate requirements for effective human-machine teaming research and foster the development of fewer but more modular, well-documented methods and testbeds. Achieving these two goals can contribute to more ecologically valid human-machine teaming research and, thus, a stronger connection between research and real-world applications. Ultimately, all these efforts can accelerate progress in the field and lay the foundation for a common platform with essential tools for human-machine teaming research.

Acknowledgements This work is part of the research lab AI*MAN of Delft University of Technology. It is the output of the Lorentz Center workshop *Research Environments for Human-Machine Teaming*.

References

[1] Akata Z, Balliet D, De Rijke M, Dignum F, Dignum V, Eiben G, et al. A research agenda for hybrid intelligence: augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. Computer. 2020 Jul;53(8):18-28. Available from: https://doi.org/10.1109/MC.2020.2996587.

[2] Johnson M, Bradshaw JM, Feltovich PJ, Jonker CM, Van Riemsdijk MB, Sierhuis M. Coactive design: Designing support for interdependence in joint activity. Journal of Human Robot Interaction. 2014 Mar;3(1):43-69. Available from: https://doi.org/10.5898/JHRI.3.1.Johnson.

[3] Salas E, Sims DE, Burke CS. Is there a “big five” in teamwork? Small group research. 2005 Oct; 36(5):555-599. Available from: https://doi.org/10.1177/1046496405277134.

[4] Johnson M, Vera A. No AI is an island: the case for teaming intelligence. AI Magazine. 2019 Mar;40(1):16-28. Available from: https://doi.org/10.1609/aimag.v40i1.2842.

[5] Miller T. Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence. 2019 Feb;267:1-38. Available from: https://doi.org/10.1016/j.artint.2018.07.007.

[6] Guidotti R, Monreale A, Ruggieri S, Turini F, Gianotti F, Pedreschi D. A survey of methods for explaining black box models. ACM computing surveys (CSUR). 2018 Aug; 51(5):1-42. Available from: https://doi.org/10.1145/3236009.

[7] Lipton ZC. The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. Queue. 2018 Jun; 16(3):31-57. Available from: https://doi.org/10.1145/3236386.3241340.

[8] Luo R, Du N, Yang XJ. Evaluating effects of enhanced autonomy transparency on trust, dependence, and human-autonomy team performance over time. International Journal of Human–Computer Interaction. 2022 Jul; 38(18-20):1962-1971. Available from: https://doi.org/10.1080/10447318.2022.2097602.

[9] Shneiderman B, Human-centered artificial intelligence: Three fresh ideas. AIS Transactions on Human-Computer Interaction. 2020 Sep; 12(3):109-124. Available from: https://doi.org/10.17705/1thci.00131.

[10] O’Neill T, McNeese N, Barron A, Schelbe B. Human–autonomy teaming: A review and analysis of the empirical literature. Human factors. 2022 Oct; 64(5):904-938. Available from: https://doi.org/10.1177/0018720820960865.

[11] Walliser J, Mead PR, Shaw TH. The perception of teamwork with an autonomous agent enhances affect and performance outcomes. In: Proceedings of the human factors and ergonomics society annual meeting. 2017 Sep; 61(1):231-235. Available from: https://doi.org/10.1177/1541931213601541.

[12] Wynne KT, Lyons JB. An integrative model of autonomous agent teammate-likeness. Theoretical Issues in Ergonomics Science. 2018 Feb; 19(3):353-374. Available from: https://doi.org/10.1080/1463922X.2016.1260181.

[13] Lyons JB, Sycara K, Lewis M, Capiola A. Human–autonomy teaming: Definitions, debates, and directions. Frontiers in Psychology. 2021 May; 12. Available from: https://doi.org/10.3389/fpsyg.2021.589585.

[14] van Diggelen J, Johnson M. Team design patterns. In: Proceedings of the 7th International Conference on Human-Agent Interaction. 2019 Sep:118-126. Available from: https://doi.org/10.1145/3349537.3351892.

[15] Verhagen RS, Neerincx MA, Tielman ML. Meaningful human control and variable autonomy in human-robot teams for firefighting. Frontiers in Robotics and AI. 2024 Feb;11. Available from: https://doi.org/10.3389/frobt.2024.1323980.

[16] Dell’Anna D, Murukannaiah P, Dudzik B, Grossi D, Jonker C, Oertel C, Yolum P. Toward a quality model for hybrid intelligence teams. In: Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems. 2024 May:434-443. Available from: https://dl.acm.org/doi/10.5555/3635637.3662893

[17] Tiddi I, de Boer V, Schlobach S, Meyer-Vitali A. Knowledge engineering for hybrid intelligence. In: Proceedings of the 12th Knowledge Capture Conference. 2023 Dec:75-82. Available from: https://doi.org/10.1145/3587259.3627541.

[18] Desolda G, Esposito A, Lanzilotti R, Piccinno A, Costabile MF. From human-centered to symbiotic artificial intelligence: a focus on medical applications. Multimedia Tools and Applications. 2024 Nov:1-42. Available from: https://doi.org/10.1007/s11042-024-20414-5.

[19] van der Meer M, Liscio E, Jonker CM, Plaat A, Vossen P, Murukannaiah PK. Hyena: A hybrid method for extracting arguments from opinions. In: HHAI2022: Augmenting Human Intellect. 2022:17-31. Available from: https://doi.org/10.3233/FAIA220187

[20] Liscio E, van der Meer MT, Siebert LC, Mouter N, Jonker CM, Murukannaiah PK. Axies: Identifying and evaluating context-specific values. In: Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems. 2021 May:799-808. Available from: http://www.ifaamas.org/Proceedings/aamas2021/pdfs/p799.pdf.

[21] Molenaar I. The concept of hybrid human-AI regulation: Exemplifying how to support young learners' self-regulated learning. Computers and Education: Artificial Intelligence. 2022;3. Available from: https://doi.org/10.1016/j.caeai.2022.100070.

[22] Zschech P, Walk J, Heinrich K, Vössing M, Kühl N. A picture is worth a collaboration: Accumulating design knowledge for computer-vision-based hybrid intelligence systems. arXiv preprint arXiv:2104.11600. 2021 Apr. Available from: https://doi.org/10.48550/arXiv.2104.11600.

[23] Zhang Q, Lee ML, Carter S. You complete me: Human-ai teams and complementary expertise. In: Proceedings of the 2022 CHI conference on human factors in computing systems. 2022 Apr;114:1-28. Available from: https://doi.org/10.1145/3491102.3517791.

[24] Li FF. How to Make AI That's Good for People. International New York Times. 2018 Mar. Available from: https://link.gale.com/apps/doc/A553326828/AONE?u=anon~6d863912&sid=googleScholar&xid=dfa7ad19.

[25] Robert L, Bansal G, Lütge C. ICIS 2019 SIGHCI workshop panel report: Human computer interaction challenges and opportunities for fair, trustworthy and ethical artificial intelligence. AIS Transactions on Human-Computer Interaction. 2020 Jun;12(2):96-108. Available from: https://doi.org/10.17705/1thci.00130.

[26] Shneiderman B. Human-centered artificial intelligence: Reliable, safe & trustworthy. International Journal of Human–Computer Interaction. 2020 Mar;36(6):495-504. Available from: https://doi.org/10.1080/10447318.2020.1741118.

[27] van Leeuwen B, Gasaway R, Spaling G, Netage BV. Adopting ai to support situational awareness in emergency response: a reflection by professionals. In: Proceedings of the First International Conference on Hybrid Human-Artificial Intelligence. 2022. Available from: https://www.hhai-conference.org/wp-content/uploads/2022/08/hhai-2021_paper_96.pdf

[28] Chung H, Holder T, Shah J, Yang XJ. Developing a Team Classification Scheme for Human-Agent Teaming. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 2024 Aug; 68(1). Available from: https://doi.org/10.1177/10711813241260387.

[29] Chen JY, Barnes MJ, Qu Z. RoboLeader: A surrogate for enhancing the human control of a team of robots. Army Research Laboratory (ARL), Aberdeen Proving Grounds, MD. 2010 Feb.

[30] Demir M, McNeese NJ, Johnson C, Gorman JC, Grimm D, Cooke NJ. Effective team interaction for adaptive training and situation awareness in human-autonomy teaming. In: 2019 IEEE conference on cognitive and computational aspects of situation management (CogSIMA). 2019 Apr:122-126. Available from: https://doi.org/10.1109/COGSIMA.2019.8724202.

[31] Johnson M, Jonker C, van Riemsdijk B, Feltovich PJ, Bradshaw JM. Joint activity testbed: Blocks world for teams (BW4T). In: Engineering Societies in the Agents World X: 10th International Workshop. 2009 Nov;10:254-256. Available from: https://doi.org/10.1007/978-3-642-10203-5_26.

[32] Mercado JE, Rupp MA, Chen JY, Barnes MJ, Barber D, Procci K. Intelligent agent transparency in human–agent teaming for multi-UxV management. Human Factors. 2016 Feb;58(3):401-415. Available from: https://doi.org/10.1177/0018720815621206.

[33] Wright JL, Chen JY, Barnes MJ, Hancock PA. Agent reasoning transparency's effect on operator workload. In: Proceedings of the human factors and ergonomics society annual meeting. 2016 Sep;60(1):249-253. Available from: https://doi.org/10.1177/1541931213601057.

[34] Verhagen RS, Neerincx MA, Tielman ML. The influence of interdependence and a transparent or explainable communication style on human-robot teamwork. Frontiers in Robotics and AI. 2022 Sep;9. Available from: https://doi.org/10.3389/frobt.2022.993997.

[35] Verhagen RS, Marcu A, Neerincx MA, Tielman ML. The Influence of Interdependence on Trust Calibration in Human-Machine Teams. In: HHAI 2024: Hybrid Human AI Systems for the Social Good. 2024:300-314. Available from: https://doi.org/10.3233/FAIA240203.

[36] van Zoelen EM, van den Bosch K, Neerincx M. Becoming team members: Identifying interaction patterns of mutual adaptation for human-robot co-learning. Frontiers in Robotics and AI. 2021 Jul;8. Avail-

able from: https://doi.org/10.3389/frobt.2021.692811.
[37] Wong JH, Chiou EK, Gutzwiller RS, Cook MK, Fallon CK. Human-Artificial Intelligence Teaming for the US Navy: Developing a Holistic Research Roadmap. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 2024 Sep;68(1):380-385. Available from: https://doi.org/10.1177/10711813241260352.
[38] van Diggelen J, Neerincx M, Peeters M, Schraagen JM. Developing effective and resilient human-agent teamwork using team design patterns. IEEE intelligent systems. 2018;34(2):15-24. Available from: https://doi.org/10.1109/MIS.2018.2886671.
[39] van Zoelen E, Mioch T, Tajaddini M, Fleiner C, Tsaneva S, Camin P, Neerincx MA, et al. Developing team design patterns for hybrid intelligence systems. In: HHAI 2023: Augmenting Human Intellect. 2023:3-16. Available from: https://doi.org/10.3233/FAIA230071.
[40] van der Waa J, Verdult S, van den Bosch K, van Diggelen J, Haije T, van der Stigchel B, Cocu I, et al. Moral decision making in human-agent teams: Human control and the role of explanations. Frontiers in Robotics and AI. 2021 May;8. Available from: https://doi.org/10.3389/frobt.2021.640647.
[41] van der Waa J, van Diggelen J, Cavalcante Siebert L, Neerincx M, Jonker C. Allocation of moral decision-making in human-agent teams: a pattern approach. In: Engineering Psychology and Cognitive Ergonomics. Cognition and Design: 17th International Conference, EPCE 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark. 2020 Jul;22(2):203-220. Available from: https://doi.org/10.1007/978-3-030-49183-3_16.
[42] Verhagen, RS. Explainable AI for Meaningful Human Control. GitHub repository. 2024. Available from: https://github.com/rsverhagen94/XAI4MHC.
[43] Esterwood C, Robert LP. The warehouse robot interaction sim: An open-source HRI research platform. In: Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction. 2023 Mar:268-271. Available from: https://doi.org/10.1145/3568294.3580086.
[44] Kox ES, Kerstholt JH, Hueting TF, de Vries PW. Trust repair in human-agent teams: the effectiveness of explanations and expressing regret. Autonomous Agents and Multi-Agent Systems. 2021 Jun;35(2):30. Available from: https://doi.org/10.1007/s10458-021-09515-9.