

9

Human-Machine Teaming in Artificial Intelligence-Driven Air Power

Future Challenges and Opportunities for the Air Force

Dr. Jean-Marc **Rickli**

Head, Global and Emerging Security Risk, and Federico Mantellassi, Research and Project Officer, Geneva Centre for Security Policy, Switzerland

Abstract

Artificial intelligence is slowly making its way into military operations, with advances in the discipline driving both a qualitative and quantitative increase of autonomy in the battlespace. This entails that warfighters will increasingly co-exist with machines with progressively more advanced autonomous capabilities. As machines make the jump from simple tools to cooperative teammates, human-machine teaming will be at the center of modern warfare. The loyal wingman concept for the air force shows that the quality of the interaction between the human and the machine is as essential to successful human-machine teaming as the technical sophistication of the machine. Understanding how to ensure trust between humans and machines will be critical. AI and machine learning will make trust more necessary and harder to achieve, while convergence with neurotechnologies might further complicate the task, bringing novel challenges.

Introduction

Artificial intelligence (AI) is increasingly making its way into the military domain. While enthusiasts, deniers, and pragmatics might disagree over the extent to which AI will confer battlespace advantages, it is already being used as an analytical enabler, disruptor, and force multiplier (Rickli and Mantellassi, 2023). The conflict in Ukraine has outwardly exemplified this trend, with AI utilized to optimize the targeting cycle by reducing sensor-to-shooter timelines, analyzing intercepted communications, disseminating intelligence and in information warfare (Rickli, Mantellassi, and Juillard, 2022, p.22). Advances in AI are accelerating autonomy and automation in warfare quantitatively but also qualitatively, with respect to the array of tasks performed – entirely or partially – by machines. With the expanding co-existence of humans and machines in the battlespace, successfully understanding how to achieve the best cooperation between humans and machines is key to unlocking the full potential of AI in military operations. This paper will focus on the emerging importance of AI and human-machine teaming for the air force, paying particular attention to the “loyal wingman” concept before zooming in on the issue of trust. It will also discuss how human-machine teaming might converge with other emerging technologies, particularly neurotechnology, to redefine future possibilities.

AI and Human-Machine Teaming

Human-machine teaming (HMT) is the act of combining human judgment with the data processing and response capabilities of computing (Motley, 2022). In the military domain, this signifies integrating humans, AI, and robotics into cooperative, interdependent, and autonomous systems (Hein and Maquaire, 2022). Human-machine interaction is already prevalent in the military and civilian domains, but this traditional model of interaction between machines and humans does not necessarily constitute a “team” (Walliser et al., 2019). It is essential to understand what makes the interaction between the two a *team*, as opposed to one where the machine is merely a tool. In a human-machine team, the machine plays an active role in achieving a goal, it “draws inferences from information, derives new insights from information, learns from past experiences, finds and provides relevant information to test assumptions, helps evaluate the consequences of potential solutions, debates the validity of proposed positions offering evidence and

arguments, proposes solutions and provides predictions to unstructured problems, plus participates in cognitive decision making with human actors" (Bittner et al., 2020, p.3). In HMT, the machine is not just a tool that completes an assigned function after instruction but a teammate able to coordinate with its human counterpart and support decision-making toward achieving a goal (Motley, 2022).

The mere presence of autonomy in a machine is not sufficient to qualify a human-machine interaction as a team if the machine is not working side-by-side and cooperatively with a human (Motley, 2022). HMT is made of three elements: the human, the machine, *and the interaction and nature of the relationship between the machine and the human* (Chahal and Konaev, 2021). AI is accelerating the potential and prevalence of HMT, especially in military operations, with a *qualitative* advance in AI-powered autonomy and a *quantitative* increase in the sheer number of autonomous systems being fielded. Machines will less simply be tools and become integral to operations where they actively participate in decision-making (Walliser et al., 2019). In this emerging context, the need to successfully pair man and machine in ways that extract the highest potential out of both lays the basis for HMT.

Advances in AI lead to a *qualitative* improvement in the variety of functions autonomous systems can perform, placing these systems alongside humans in more critical and cooperative ways. AI-powered autonomy, increasingly driven by machine learning (ML), is leading to an increase in the complexity of tasks that machines can complete independently and more accurately. For example, Primer AI is using different techniques, most notably Natural Language Processing (NLP), to fuse multiple intelligence inputs (audio, visual, text) to provide the Ukrainian military with a real-time, autonomously-generated intelligence picture of the battlespace (Primer, 2022). Similarly, loitering munitions have made their way into the battlespace, relying on AI to identify and engage targets. Weapon systems are increasingly benefiting from dual-use AI technologies that are developed in the private sector. For instance, in May 2022, Deepmind released a new algorithm, "GATO," capable of performing 604 different tasks (Deepmind, 2022).



The more functions an algorithm can perform, the more functionally equivalent it gets to a human being, increasing the potential for HMT.

The more functions an algorithm can perform, the more functionally equivalent it gets to a human being, increasing the potential for HMT. These developments also drive a *quantitative* increase in the number of autonomous systems or systems with some level of autonomy (Boulanin and Verbruggen, 2017). The advent of swarm technology – already demonstrated in experiments – will accelerate this trend (Xin et al., 2022; Mehta, 2017). Many of these technological advancements are pioneered in the private sector and are gains that are more in reach for smaller powers, non-state actors, and individuals, unlike with legacy systems (Ashby et al., 2020; Rickli, 2020a). With this quantitative and qualitative increase in robotics and autonomy, the tempo of war will accelerate (Rickli, 2019). In algorithmic wars, the only way for human beings to retain a meaningful role in the decision-making process will be through integrating and synchronizing human and machine inputs (Walsh, 2021). HMT is imperative for air forces to create symbiotic relationships and interactions between its human and machine elements.

While the successes and high potential for AI are reshaping HMT, the inherent limitations and relative brittleness of AI means that machines must continue to co-exist with humans (Rickli, 2020b). It is unrealistic to assume that all aspects of warfare will be automated in the near future and that AI will replace all – or even most – of the tasks of the warfighter. AI remains “narrow” for the time being, able to outperform humans only in activities that can be easily codified or where there are clearer rules and metrics (UK Ministry of Defense, 2018). Additionally, these narrow AI applications are prone to adversarial attacks and data biases which can lead to spectacular failures (Scott-Hayward, 2022). While AI's superior analysis, data processing, and statistical correlation capabilities vastly outperform humans, the latter will maintain the cognitive advantage on the battlespace, understanding context, relying on intuition, breaking rules when necessary, and adapting in novel ways (Losey, 2022). The prioritization of HMT is precisely driven by the need to achieve the best outcomes from combining machine algorithms with humans, where each can play to its strengths (Jatho and Kroll, 2022). When the right tasks are assigned to the right element of the team (human or machine), and the interaction between the two is of high quality, human-machine teams vastly outperform humans and machines acting on their own (Jatho and Kroll, 2022).

The Loyal Wingman Concept and Trust

Air forces have been particularly attuned to the requirements for HMT (Briant, 2021). The adoption of AI and interaction between man and machine increasingly lie at the heart of air power (Briant, 2021). Several high-profile programs around the world are experimenting with the "Loyal Wingman" concept, which proposes autonomous, unmanned aircraft work alongside manned aircraft collaboratively. The need for HMT in the air force stems from multiple sources. First, integrating AI, robotics, and computers into operations accelerates data production and collection, running the risk of "information overload" on warfighters (Johnson et al., 2014). Pilots operating complex systems require high levels of concentration and multi-tasking, with limited scope to analyze yet more data in real time. AI assistance can alleviate some of this burden. As part of its "Vanguard programs," the U.S. Air Force is developing "Skyborg" – an "autonomous aircraft teaming architecture" which will enable unmanned aircraft such as the Kratos Valkyrie to fly in a team with manned counterparts (U.S. Air Force, n.d). These drone wingmen will enhance the performance of pilots by offloading some data analysis tasks, mapping targets, and air defense systems, and suggesting flight corridors for pilots (Losey, 2022). In time, onboard AI systems will be able to suggest appropriate courses of action to pilots. Second, U.S. near-peer adversaries such as China have invested massively in developing anti-access/area denial (A2/AD) capabilities, making the operational environment extremely contested and lethal (Grynkewich, 2017). To increase survivability in these highly contested environments – and given the advent of swarming technology – loyal wingmen could be used *en masse* to penetrate and saturate adversarial defenses, act as decoys, or deliver kinetic effects (Perret, 2021).

A myriad of challenges come into play to complicate the goal of achieving effective HMT for loyal wingmen, not all of which are always of a technical nature. Multiple factors impact and influence the quality of interaction in HMT, some of which are illustrated in Figure 9.1. The focus of much research into HMT is not wholly on technical capabilities and characteristics but on the nature and quality of the relationship between the human and the machine and their interaction. Effective HMT is only partially dependent on the sophistication of AI but heavily contingent on the quality of the interaction. A parallel with human-human teams is often made: The effectiveness of teams is not simply the aggregate of separate member abilities and inputs but actually depends on the successful integration and coordination of individual efforts through team processes and teamwork

(Funke et al., 2022). In HMT, the human has to understand 1) its role, 2) the AI system, 3) how to interact with the AI system/teammate, and 4) how to interact with the other human teammates (Puscas, 2022).



Figure 9.1: Influence Factors on the Quality of Human-Machine Interaction

Trust is a vital element in HMT but a highly complex concept with many variables which affect it, such as demographic, geographic location, context, and multiple other factors (Chahal and Konaev, 2021). Davis, Mayer, and Schoorman define trust as “the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, *irrespective of the ability to monitor or control that other party*” (1995). The jury is still out on whether the variables and dynamics which govern human-human trust are the same as those which influence human-machine trust (Celaya and Yeung, 2019). This article will consider trust as an “individual’s confidence in the reliability of the technology’s conclusions and its ability to accomplish defined goals.” (Chahal and Konaev, 2021). As the UK Ministry of Defense notes, AI systems will not only be

limited by what can be done but also “what actors trust their machines to do” (UK Ministry of Defense, 2018). In fact, without trust, AI will not reach its full potential because its use may be limited to low-risk scenarios rather than where it can offer a real advantage (Motley, 2022). The challenge for air forces relates to how trust in AI can be enhanced to augment the human-machine relationship qualitatively. Training and the characteristics of user interfaces are central considerations in improving trust in AI and improving HMT.

Trust in AI systems can be impacted – among other things – by mechanical understanding, system predictability, familiarity, and context (UK Ministry of Defense, 2018, p.48). These elements can be – at least partially – addressed through appropriate training to instill confidence and understanding in AI systems, as well as through user-friendly interfaces that enable control over machines through heightened familiarity and a sense of predictability (Puscas, 2022). However, increased AI-driven autonomy and the expanding role of ML means that appropriate training and user interfaces will simultaneously become more complex and challenging to achieve but also more necessary – but all of these can have an adverse effect on trust (Puscas, 2022). The “black-box” problem of AI impedes understandability, explainability, and predictability, thus reducing trust, which only grows with the degree of autonomy delegated to machines (Michel, 2020). As ML algorithms learn over time, the pilot/human teammate must know how the machine changes, what it is learning, and how that will affect its outcomes. The complexity of automation therefore makes continuous training primordial but also more complex. Similarly, the usability of interfaces becomes more essential but simultaneously also more complex. As the machine teammate offers more autonomy and its algorithmic processes are complexified, its outputs are ever more challenging to explain and communicate (Puscas 2022). This complexity, in turn, may contribute to reducing trust in the system.



As ML algorithms learn over time, the pilot/human teammate must know how the machine changes, what it is learning, and how that will affect its outcomes. The complexity of automation therefore makes continuous training primordial but also more complex.

Lack of trust affects HMT, reducing the efficiency and potential of man-machine teams and thereby lowering the probability of their operational use. However,

excessive trust in a machine can also negatively impact HMT. (Scharre, 2018; Puscas, 2022). Indeed, increased autonomy in systems can lead to the “automation conundrum,” whereby the loss of user alertness is directly proportional to the system’s enhanced automation and perceived reliability (Puscas, 2022). In this sense, the complexity of a machine can both reduce trust or increase it to an excessive level. Therefore, military operators need to have a healthy dose of skepticism regarding the autonomous systems they operate or oversee, which entails the capacity to accurately assess the limitations of the system (Scharre, 2018, p.144). This again stresses the importance of appropriate training to gain understanding of the system, lest it becomes “fully autonomous by neglect” (Puscas, 2022). In 2003, a series of incidents involving the American “Patriot System,” a *human-in-the-loop* air defense system, resulted in fratricides attributed to excess trust and inappropriate training, which made the system *de facto* fully autonomous (Scharre, 2018).

Convergence with Neurotechnology

The rise of any individual emerging technology does not happen in a vacuum, independently from other technological innovations. Hence, the convergence between sets of emerging technologies requires careful attention to anticipate future challenges arising from the unpredictable interactions between these technologies (Rickli and Mantellassi, 2022). The fact that human interactions for now always happen through an interface (such as screens) implies some limitations in their efficiency. Operational demonstrations employing loyal wingmen have featured operators directing machine teammates with handheld tablets (Trevithick, 2021). One way to do away with the need for interfaces is to directly connect the machine with the human brain using a set of technologies that allow bi-directional interaction between the brain and machines. The ongoing convergence between the fields of AI and neurotechnology – the field of science which seeks to connect technical components to a nervous system – may allow optimal human-machine interaction in the future (Rickli, 2020c).

The quality of the relationship between humans and machines is essential for HMT and is influenced by the characteristics of interfaces. Engaging with an external interface (screen, tablet, computer) to supervise or otherwise engage with a machine or semi-autonomous agent can be highly demanding cognitively, resulting in a loss of alertness and complacency (Puscas, 2022). Advances in neurotechnologies may resolve the need for pilots to interact with an external

link, such as a screen or display, to view, communicate, and transmit information to and from the machine. Brain-computer interfaces (BCIs) would seamlessly integrate the control of loyal wingmen into the cognitive processes of pilots in ways that reduce cognitive overload, accelerate the Observe, Orient, Decide, and Act (OODA) loop, illustrated below in Figure 9.2, and remove the complex task of designing interfaces. BCIs could “facilitate multi-tasking at the speed of thought” and allow pilots to “interface with smart decision aids” and multi-vehicle partnering (Bartels et al., 2020). DARPA, which has efforts underway to operationalize the use of BCI, has simulated dogfights in which algorithms get “inside” adversarial OODA loops to defeat human opponents through faster decision-making (Tegler, 2020).

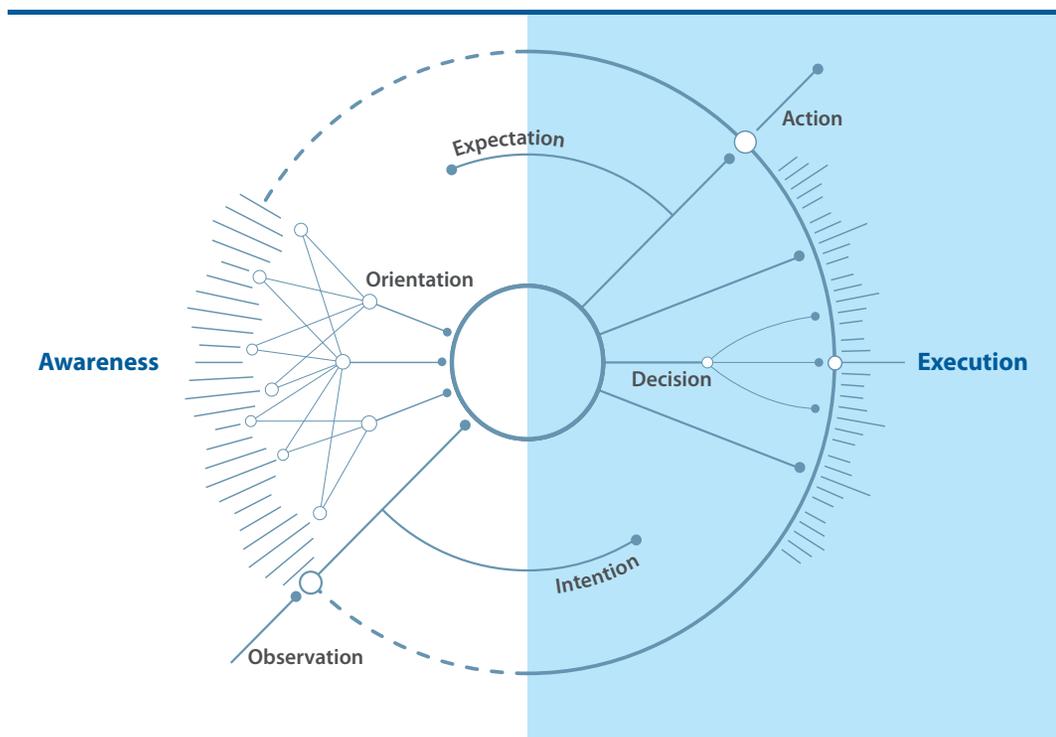


Figure 9.2: A Dashboard-Style Representation of the OODA Loop

The convergence of technologies merges the qualities of each technology, accelerates their development, and leads to novel innovations previously not possible. However, convergence also enables the transfer of risks associated with individual technologies to the other, sometimes creating entirely novel and

unpredictable challenges. Hence, while neurotechnologies such as BCI could, in time, make possible optimal HMT, they also introduce new and more complex risks. For neurotechnologies, these include ethical dilemmas of data privacy and cognitive and mental integrity, novel avenues for brain manipulation and waging cognitive warfare, as well as unprecedented capacity for surveillance, to name a few (Rickli and Ienca, 2021).

Conclusion

Advances in AI lead to a qualitative and quantitative increase in AI-driven machines with autonomous capabilities in military operations. The relative brittleness of AI today, however, implies that warfighters will co-exist with machines on the battlespace, increasingly as teammates. Air forces must invest in attritable, unmanned capabilities such as loyal wingmen to overcome challenges with information burden, cognitive overload, and low survivability in highly contested airspaces. The challenge of HMT lies at the center of future air power capabilities. Optimizing the relationship between man and machine, not only the algorithms that fly these loyal wingmen, is critical. Ensuring pilots have trust in the functioning and performance of AI-driven autonomous systems and a robust understanding of their inherent limitations is crucial. In addressing these issues, the importance of training and user interfaces stands out. Understanding how autonomy impacts trust and affects the relationship between pilots and their machine teammates will be decisive for advancing HMT. Neurotechnologies, specifically BCIs, is an emerging area that is likely to converge with loyal wingmen concepts and allow for more optimal connections between man and machine. Anticipating the novel challenges which come from the convergence between BCIs, loyal wingmen, and AI through foresight will be vital in steering the development of future air power.

References

- Airforce Research Lab (2022). Skyborg.
- Ashby M., Bourdeaux B., Curriden C., Grossman D., Kilma K., Lohn A., Morgan F.E. (2020) Military Applications of Artificial Intelligence: Ethical Concerns in an Uncertain World. RAND Corporation.
- Bartels E.M., Binnendijk A., Marler T. (2020) Brain-Computer Interfaces: US Military Applications and Implications, an Initial Assessment. RAND Corporation.

- Bittner E., Briggs R.O., Elkins A., Maier R., Merz A.B., Oeste-Reis S., Randrup N., Seeber I., Sollner M., Schwabe G., de Veerde G.-J., de Veerde T. (2020) "Machines as teammates: A research agenda on AI in team collaboration" *Information and Management*, 57(2). Available at: <https://doi.org/10.1016/j.im.2019.103174>
- Boulanin V., Verbruggen M. (2017) Mapping the Development of Autonomy in Weapon Systems. Stockholm, International Peace Research Institute.
- Briant R. (2021) La synergie homme-machine et l'avenir des opérations aériennes. Institut français des relations internationales, n106.
- Celaya A., Yeung N. (2019) "Confidence and trust in human-machine teaming." Homeland Defense & Security Information Analysis Center, 6(3).
- Chahal H., Konaev M.(2021) "Building trust in human-machine teams." Brookings Institute, February 18.
- Deepmind (2022) A Generalist Agent. Available at: <https://www.deepmind.com/publications/a-generalist-agent>
- Funke G., Greenlee E.T., Lyons J.B., Matthews G., Tolston M.T., (2022). "Editorial: Teamwork in Human-Machine teaming." *Frontiers in Psychology*. doi: 10.3389/fpsyg.2022.999000
- Grynkewich A. (2017) "The future of air superiority, part III: Deveating A2/AD." *War on the Rocks*. January 13. Available at: <https://warontherocks.com/2017/01/the-future-of-air-superiority-part-iii-defeating-a2ad/>
- Hein S., Maquaire E. (2022) "Human-Robot teaming Research Initiative for a combat aerial network (Hurricane)." Assets Plus. Available at: <https://assets-plus.eu/wp-content/uploads/2022/05/HURRICANE.pdf>
- Jatho E., Kroll J.A. (2022). "Artificial intelligence: Too Fragile to Fight?" U.S Naval Institute. Available at: <https://www.usni.org/magazines/proceedings/2022/february/artificial-intelligence-too-fragile-fight>
- Johnson, S.T., Porche I.R., Tierney, S., Saltzman E., Wilson B. (2014). Big Data: Challenges and Opportunities. In Data Flood: Helping the Navy Address the Rising Tide of Sensor Information. RAND Corporation.
- Losey S. (2022). "Air Force must build trust to add drone wingman, report says." *Defense News*, October 13. Available at: <https://www.defensenews.com/air/2022/10/13/us-air-force-must-build-trust-to-add-drone-wingmen-report-says/>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). "An Integrative Model of Organizational Trust." *Academy of Management Review*, 20(3), 709-734. doi:10.2307/258792
- Mehta, A. (2017). "Pentagon launches 103 Unit Drone Swarm." *Defense News*, January 10. Available at: <https://www.defensenews.com/air/2017/01/10/pentagon-launches-103-unit-drone-swarm/>
- Michel A.H., (2020). The Blackbox, Unlocked: Predictability and Understandability in Military AI.' United Nations Institute for Disarmament Research, Geneva, Switzerland.
- Motley J.O. (2022) "The testing and explainability challenges facing human-machine teaming." Brookings Institute, March 31.
- Perret B. (2021), "Loyal wingmen could be the last aircraft standing in a future conflict." Australian Strategic Policy Institute. November 22. Available at: <https://www.aspistrategist.org.au/loyal-wingmen-could-be-the-last-aircraft-standing-in-a-future-conflict/>

- Primer (2022) AI Tool for Monitoring Fast-Evolving Information.
Available at: <https://primer.ai/news/ai-tool-for-monitoring-fast-evolving-information/>
- Puscas I., (2022) Human-Machine Interfaces in Autonomous Weapon Systems Considerations for Human-Control. United Nations Institute for Disarmament Research. Geneva. Switzerland, pp.9-10
- Rickli J-M. (2020a). "Surrogate Warfare and the Transformation of Warfare in the 2020s," Observer Research Foundation, Mumbai, 30 December.
- Rickli J-M. (2020b). "Containing Emerging Technologies' Impact on International Security," in Jonsson, Oscar (ed.). *Modern Warfare: New Technologies and Enduring Concepts*. Stockholm, Transatlantic Leadership Forum, pp. 76-84.
- Rickli J-M. (2020c). "Neurotechnologies and Future Warfare," Nanyang Technological University, Singapore, 7 December.
- Rickli J-M. (2019). The Destabilizing Prospects of Artificial Intelligence for Nuclear Strategy, Deterrence and Stability, in Boulanin, Vincent (ed.). *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk: European Perspectives*. Stockholm: Stockholm International Peace Research Institute, Volume I, pp. 91-98.
- Rickli J-M. and Ienca M. (2021). "The Security and Military Implications of Neurotechnology and Artificial Intelligence" in Friedrich and al. (eds). *Clinical Neurotechnology meets Artificial Intelligence*. Berlin: Springer, pp. 197-214.
- Rickli J-M. and Mantellassi F. (2023) Artificial Intelligence in Warfare: Military Uses of AI and their International Security Implications. In Michael R., and Bitzinger R.A. (eds) *The AI Wave in Defense Innovation: Assessing Military Artificial Intelligence Strategies, Capabilities, and Trajectories*. Routledge, forthcoming.
- Rickli J-M. and Mantellassi F. (2022). *Our Digital Future: The Security Implications of Metaverses*. Geneva, GCSP Strategic Security Analysis.
- Rickli J-M, Mantellassi F. and Julliard V. (2022) Implications for the Future of Warfare. In Greminger, T., and Vestner, T. (eds) *The Russia-Ukraine War's Implications for Global Security: A First Multi-issue Analysis*. Geneva Centre for Security Policy.
- United Kingdom Ministry of Defense (2018). *Joint Concept note on Human-Machine Teaming*.
- Scharre, P. (2018) *Army of None: Autonomous Weapons and the Future of War*. New York. W.W Norton & Company.
- Scott-Hayward S. (2022). "Securing AI-based Security Systems." Geneva Centre for Security Policy. Strategic Security Analysis, Issue 25.
- Tegler E. (2020). "AI Just Won A Series Of Simulated Dogfights Against A Human F-16 Pilot, 5-0. What Does That Mean?" *Forbes*. August 20.
- Trevithick, J. (2021). "Here's how Fighter Pilots Could Control "loyal Wingmen" via a Tablet on Their Thigh." *The Warzone*. September 7.
- Walliser, J. C., de Visser, E. J., Wiese, E., & Shaw, T. H. (2019). Team Structure and Team Building Improve Human-Machine Teaming With Autonomous Agents. *Journal of Cognitive Engineering and Decision Making*, 13(4), 258-278.
- Walsh, B. (2021). "An insider's view of algorithmic warfare." *Axios Future*. November 18.
Available at: <https://www.axios.com/2021/11/17/robert-work-artificial-intelligence-warfare>
- Xin, Z. et al. (2022) "Swarm of Micro Flying Robots in the Wild." *Science Robotics*, vol 7, issue 66.