Metaverse Military TMAF26

Are You REDy? Using Red AI Tactics And The Defence Metaverse To Augment Future Force Design

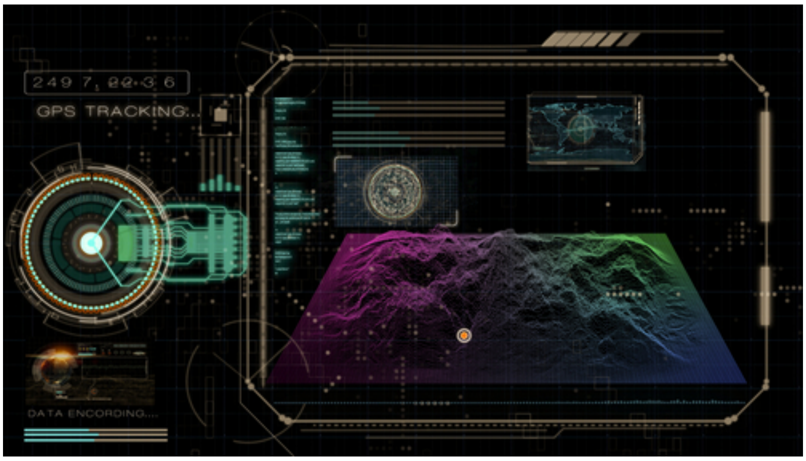
Written by Heiko Borchert on July 16, 2026

share share share

Abstract: Force designers face demanding methodological challenges when considering how best to adapt concepts, successfully adopt new technologies, and overhaul existing structures and processes amid a fundamentally changing geostrategic environment. The Defence Metaverse, an ultra-realistic virtual twin of the battlefield, and powerful new red AI tactics offer sovereign European solutions to tackle these challenges by stress-testing the assumptions underpinning concepts, procurement decisions, and training against adaptive non-conventional peer adversaries.

Problem statement: When do force designers, willing to adapt existing concepts and adopt new technologies, know that novel ways of delivering fighting power will make a difference?

So what?: The German Bundeswehr should invite allied partners to join the Defence Metaverse to validate future force concepts in a multinational digital environment. By exposing allied concepts to an AI-enhanced red peer, allied forces will understand what works, where their concepts fall short of expectations, and how best to close existing shortfalls.



Source: shutterstock.com/Tridsanu Thopet

Unleashing Red to Properly Prepare Blue

Capability development needs to balance operational (and future) needs, conceptual readiness, and technological maturity, but the forces shaping these three dimensions hardly evolve in tandem. During the missile war between Iran and Israel in 2025, for example, "Tehran changed tactics and found gaps in Israel's armour through trial and error. Tehran began to launch more advanced and longer-range missiles from a wider range of locations inside Iran (….). The regime also altered the timing and pattern of attacks (…) As the war went on, Iran fired fewer missiles, but its success rate rose (…)."[1] This example illustrates that 'probing the adversary' can reveal conceptual and technological gaps that enemy forces can exploit to their advantage.

Commenting on the development of the U.S. Air Force concept of collaborative combat aircraft (CCA), then-Assistant Secretary of the Air Force for Acquisition, Technology and Logistics Andrew Hunter argued in 2024, that the service "has to develop concepts of operations for them, or exactly how they would be controlled and operate in battle" before "CCAs can be integrated into the squadrons. "As the experimental operations unit sets those tactics," it is important for them to have actual CCAs to experiment with.[2] However, building before conceptualising future products can undercut the feedback loop between conceptual readiness and technological maturity, as the latter is likely to override the former.

Both examples suggest that capability developers and force designers would benefit greatly from tools that help them explore how alternative – and hitherto unseen – red force behaviour affects blue force options. Although there are traditional methods of simulation and military operations research, problems stem from a lack of adaptive red-force methods and from analytical bias in how blue-force designers model red-force behaviour. This not only shapes adversarial performance, but also blue force preparedness.[3], [4] Consequently, surprise does not really emanate from adversarial action but from blue negligence in shaping red forces. Moreover, this challenge becomes ever more problematic the more blue and red forces operate at comparable levels of technological sophistication or conceptual maturity.

What Is the Defence Metaverse?

GhostPlay, a capability and technology development project of the German Bundeswehr's Centre for Digitalisation and Technology Research (dtec.bw), was launched in 2021 to address this challenge.[5] The goal is to demonstrate that AI can self-learn optimal tactical behaviour under uncertainty. To this purpose, GhostPlay has developed the Defence Metaverse. The Defence Metaverse implements the concept of model-based digital engineering, understood as an integrated digital approach "that uses authoritative sources of system data and models as a continuum across disciplines to support lifecycle activities from concept through disposal."[6] Akin to a virtual twin of the battlefield, the Defence Metaverse creates an ultra-realistic, high-fidelity environment that models sensor and effector characteristics, weather, vegetation, terrain, and urban and rural infrastructure. What sets GhostPlay's Defence Metaverse apart from other solutions is its counterplay logic for developing self-learning multi-agent red teams.

Modelling the Operational Environment in the Defence Metaverse; Source: Heiko Borchert and Christian Brandlhuber.

To understand why this approach is different, it is important to note that the Defence Metaverse has been built to advance tactical assertiveness. Thus, AI tactics must meet the demands of the operational environment for which they have been conceived. To develop the required level of detail, automated pipelines were created that use different data (see figure above). The pipelines start with using georeferenced Digital Terrain Elevation Data (DTED).[7] This initial data set was augmented with information on buildings. Both have been fused with Light Detection and Ranging (LIDAR) data to create terrain and building data profiles with a granularity of up to 10 centimetres. In addition, synthetic vegetation has been generated to replicate the vegetation in the mission area. The simulation also provides different weather conditions, such as day/night light, rain, fog, or snow. If needed, trafficability analyses and tactical assessments of geo-positions for red and blue forces can also be conducted.

Unlike traditional approaches that script red forces according to analytical assumptions, GhostPlay's approach "enables AI to develop successful tactics in interaction with a smart adversary." This means that the system is "forward planning a series of play moves to determine the optimal decision sequence." Whenever red or blue has developed successful policies, the other side learns how to counteract. This counterplay logic helps develop decision rules by "training them against more complex and challenging adversaries" to define non-exploitable policies.[8]

This approach deviates from traditional analytical simulation and offers novel ways to develop scenarios for force planning that rest on interaction rather than scripting (see box below). This difference is essential as it generates substantially more and materially different scenarios than traditional simulations can produce.[9] Consequently, the scenarios are not just slight variations of a predefined script but may include completely new and comparably complex tactical patterns, such as red forces playing 'free jazz' to create surprising and unexpected behaviour that challenges blue-force thinking. In addition, unexpected outcomes counter the risk of predictable red forces that play according to blue assumptions, significantly reducing the risk of analytical bias.[10]

Three Core Features Separating the Defence Metaverse from Traditional Analytical Simulation:
– The Defence Metaverse seamlessly integrates strategic, operational, and tactical levels of analysis into one digital environment to realistically illustrate effects and their impact on sensor-to-shooter webs. This is important because traditional simulation models focus on a single layer and model their assumptions accordingly. As a result, simulation results depend on these pre-specified assumptions and thus only confirm what is already known. By contrast, the Defence Metaverse – thanks to its permeability – overcomes this lock-in effect and yields much more realistic simulation results.
– State-of-the-art simulation methods have limits. Monte Carlo simulations, for example, are used to conduct tactical analyses by introducing random environmental effects, sometimes also of enemy behaviour. In general, however, these simulations lack red force policies that would generate a red force exploiting a situation in a structured, coordinated way, thereby creating entirely new tactical red force behaviour to represent true adversarial intent. That said, intelligent red forces can systematically exploit tactical benefits or flaws in specific defence systems or capabilities.[11] In addition, scripted scenarios cannot offer large, representative sets, as efforts to do so quickly become disproportionate. Therefore, an optimal simulation environment needs to be able to perform both tasks. The Defence Metaverse fulfils this requirement by presenting the results of large-scale runs as statistical aggregates of individual scenarios.
– Contextual reasoning means AI-enabled red forces consider the context in which they operate, factor in the consequences of their own actions, anticipate adversarial behaviour, and act accordingly. This leads to AI red forces that can coordinate actions of large groups of challengers across multiple domains, thereby stimulating learning by interacting with adversarial forces that do not behave as scripted software agents.

What Separates the Defence Metaverse from Traditional Analytical Simulation; Updated from Heiko Borchert, Torben Schütz, and Christian Brandlhuber, "Leveraging the Defence Metaverse. Unlocking the Power of Artificial Intelligence for Force Development", The Air Power Journal (Fall 2023), 26–27, https://theairpowerjournal.com/leveraging-defense-metaverse-artificial-intelligence-for-force-development/.

Force Design and the Defence Metaverse

GhostPlay's focus on counterplay and non-conventional AI red forces responds to changes in today's threat environment. With the growing threat of military confrontation with peers operating at comparable levels of conceptual and technological sophistication, the idea of distributing force is gaining traction. One prominent example of this is Mosaic Warfare, a conceptual idea developed by the U.S. Defence Advanced Research Projects Agency (DARPA) in 2018.

Mosaic Warfare rests on the assumption that force packages are to be decomposed into the smallest possible units and reconfigured to meet mission goals and address the challenges posed by adversarial forces. Most importantly, Mosaic Warfare suggests deviating from hierarchical command-and-control (C2) structures that run through centralised command posts and common operational pictures, in favour of decentralised, largely non-hierarchical decision-making. However, the more one decomposes, the more one needs to focus on methods and procedures that enable optimal coordination of decentralised force packages across domains and echelons.[12]

Thinking about alternative force designs thus turns into analysing and developing optimal cooperation policies under uncertainty. This implies, first, the ability of systems to comprehensively assess the situation they operate in to anticipate adversarial behaviour. Second, systems need to learn "how to orchestrate and organise (…) to achieve objectives across time-extended scenarios and in response to enemy action." Finally, these policies must incentivise systems to learn on their own "when and how to cooperate to solve complex tasks with partners."[13]

As a result, GhostPlay enables emergent[14] "defence responses against any incoming threat" with a setup that does not require a C2 component at all, although it may be equipped with one for user interaction. Rather, individual elements of an aggressor swarm of unmanned aerial systems (UAS) or a federated Ground-Based Air Defence (GBAD) solution act as members of a multi-agent system, whereby each agent learns policies that "motivate other agents in the same team to cooperate."[15]

AI Tactics for Red UAS Swarms vs GBAD

WILD HORNETS, an advanced concept study conducted in cooperation with the German Army Concepts and Capability Development Centre (ACCDC), provided fertile ground for validating and further refining the ideas presented above. With the help of the Defence Metaverse, the study assessed future capability requirements to use air-launched effects (ALE) against GBAD.[16] As Justin Bronk and Jack Watling[17] argue, developing a UAS-based precision strike complex is demanding, given the trade-offs among many design parameters, such as airframe, navigation, sensors, effectors, datalinks, and more. That's why WILD HORNETS examined different design parameters to determine which requirements would deliver optimal tactical value. In addition, the study team was also tasked with developing new tactics, techniques, and procedures (TTPs) for ALE deployment with the purpose of assessing how AI-enhanced TTPs would change future UAS battlefield behaviour.

The ACCDC provided four vignettes for evaluation. In the basic configuration, an Army aviation unit consisting of four attack and three transport helicopters flew within 20km of the front edge of the battlespace at altitudes below 30m to identify and neutralise an adversarial high-value target, protected by several Short-Range Air Defence (SHORAD) systems. German Gepard Artillery Anti-Aircraft systems, which use four different types of munitions,[18] have been used as models for the SHORAD system. Attacking red swarms consisted of Switchblade 600 loitering munitions and SPIKE non-line of sight (NLOS) missiles. ALE-NLOS missile communication was simulated in the end game, and signal-based, probabilistic, and auto-encoder models were used to simulate different sensors.[19] To identify the added value of AI tactics, the analytical baseline (see box below) was established by running scenarios ranging from helicopter-guided swarms to swarm members guiding their companions, using different attack trajectories against the SHORAD, and operating SHORAD systems under different daylight conditions and at randomly chosen positions.

Based on 10,000 out-of-sample evaluation scenarios, findings suggest:
– Only in one out of 10 cases will a state-of-the-art ALE using traditional trajectories succeed against the SHORAD constellation shown in the simulation.
– A SHORAD system using active sensors will be even more powerful as these sensors improve the lead time to detect and engage incoming ALE. Tracks can be automatically initialised if the ALE swarm operates in the absence of ground clutter. Electro-optical (EO) sensors compensate for occasional track losses.
– The performance of a SHORAD system that is manually operated (human in the loop) and uses passive sensors (and laser) decreases significantly when facing a swarm of three to four ALE approaching from different directions. This is mainly due to the zoom latency to detect, classify, and identify small flying objects, such as small Unmanned Aerial Systems (sUAS).
– Things change dramatically to the benefit of the defender if the SHORAD system is AI augmented. In this case, it is almost impossible for ALE swarms – using conventional trajectories and flying at around 120-180 km/h – to overwhelm the SHORAD constellation. In around 60% of scenarios, ALE swarms with up to 40 members were destroyed.

Source: Author.

Demonstrating the ability to self-learn optimal attack patterns was a key goal of GhostPlay. To start with, the swarm's flight characteristics were gradually improved. Swarms that exploit terrain features and vary flight trajectories and tempo challenge SHORAD systems, as these flight patterns are difficult to detect. This is an important insight, as most current Counter-UAS solutions assume scenarios with stationary flying objects stacked in lines, rather than flying demanding patterns.

As algorithmic behaviour matured, the red swarm's ability to coordinate collaboratively developed. This requires ALE swarm members to develop ad hoc strategies that define mission goals and sub-goals, and to delegate their execution to team members. Today, swarms predominantly fly in patterns, where swarm members fly according to predefined waypoints or mimic each other's behaviour.[20] WILD HORNETS, by contrast, demonstrates the logic of distributed control, which implies that swarm members pursue their own goals and cooperate when giving up individual objectives will help them achieve a higher-value common goal. Most importantly, this tactic – demonstrated in the Defence Metaverse two years prior (!) to the first battlefield observation in Ukraine[21] – increases swarm efficiency and effectiveness. Against a modern, automated Gepard SHORAD system, a low-and-slow red swarm succeeded in only 4 of 10 missions. However, when the swarm coordinated collaboratively, it achieved mission success in 8 out of 10 cases. At the same time, the size of the swarm shrank drastically. When facing a manually operated Gepard, ALE swarms capable of terrain following flight manoeuvres needed around 7 members to outpace the SHORAD – but with collaborative coordination, only 4 members were needed, thus significantly reducing red's logistical footprint.

WILD HORNETS also demonstrates that ALE swarms exhibit cooperation patterns observed among animals. Sarah M. E. Fortune, Xi Cheng, Keith Holmes, and Andrew W. Trites showed that killer whales (Orcas) and Pacific white-sided dolphins opportunistically engage in "cooperative foraging" vis-à-vis adult Chinook salmon. According to them, "whales may eavesdrop on dolphin echolocation to scan broader areas to locate large Chinook salmon," while prey too big for dolphins to swallow was broken apart for sharing by Orcas. In military parlance, dolphins and Orcas seem to have formed a most effective sensor-to-killer web that rests on each partner's comparative advantage.[22]

The same behaviour was observed in the Defence Metaverse when operating heterogeneous swarms of ALE and SPIKE NLOS missiles against Russian 76N6 Clam Shell low-altitude acquisition radar systems. This specific radar sits atop a mast of up to around 40 meters. At this height, attacking the radar from below to benefit from low-flying cover is unfeasible, as attacking ALE swarms are unable to build up sufficient kinetic energy to ignite the charge. Attacking from above would make ALE members easy prey for the air defence system. In this situation, ALE members learned that their attrition rate would be too high to succeed. Thus, they called in more powerful SPIKE NLOS missiles to engage the target and clear the way for ALEs to accomplish their mission. While it is too early to tell if this opportunistic collaboration would succeed in combat, it is most interesting to note that AI has self-learned a cooperative behaviour that mimics cooperation among animals, without human operators telling AI to replicate this specific attack pattern.

Fail in the Defence Metaverse to Succeed on the Battlefield

Wars between Russia and Ukraine or the U.S., Israel and Iran drive home one important message: Effectiveness takes centre stage when red is playing the attrition game. Attritional warfare, the saying goes, is a numbers game with rapid (and sustainable) replenishment of weapons as a key parameter. Nevertheless, this logic changes if one of the parties ups the ante on effectiveness. This is where AI-based collaborative coordination comes in. Simulation findings for WILD HORNETS clearly show that attacking swarms can drastically increase their effectiveness while at the same time shrinking in size. That is, AI tactics serve as a means to generate superior effects with fewer assets.[23]

This finding constitutes an important contrarian signal to adjust current discussions about future force size and design. It is an observation that underlines the value of the Defence Metaverse for force planners, as it enables them to test their assumptions against non-cooperative red AI peers. It is an observation that emphasises effectiveness, not only efficiency, which will gain in prominence given growing concerns about defence value for money amid rising public debt in Europe. In this environment, defence planners must make sure they have vigorously tested their hypotheses and do not fall victim to the promises of defence contractors. Real-life tests and experiments are one very important tool to make sure the armed forces deliver. However, organising real-life tests costs money, requires sophisticated testing sites, and falls out of favour with citizens in densely populated areas. Using the Defence Metaverse as a digital test lab, by contrast, offers all the advantages at a fraction of these costs.

The European Central Bank is stress-testing the resilience of the European financial system against hypothetical financial and economic shocks. The Defence Metaverse and red AI tactics provide a novel military equivalent to stress-test national and European defence assumptions, to develop concepts and capabilities for the new age of protracted warfare characterised by adversaries deliberately depleting European military capabilities. Consequently, the German Bundeswehr should invite allied partners to test and validate future force concepts in the Defence Metaverse. By exposing allied concepts to an AI-enhanced red peer, allied forces will understand what works, where their concepts fall short of expectations, and how best to close existing shortfalls.

This research paper is funded by dtec.bw – Digitalization and Technology Research. Center of the Bundeswehr. Dtec.bw is funded by the European Union – NextGenerationEU.

---

[1] Zvi Smith and Benoit Faucon, “Through Trial and Error, Iran Found Gaps in Israel’s Storied Air Defence,” Wall Street Journal, July 15, 2025, https://www.wsj.com/world/middle-east/iran-israel-

air-defense-362826e3.

[2] Stephen Losey, “Air Force Buying More Drone Wingmen to Develop Operational Tactics,” Defense News, November 15, 2024, https://www.defensenews.com/air/2024/11/15/air-force-buying-more-drone-wingmen-to-develop-operational-tactics/.

[3] Yvonne Hofstetter, “From Scripted to Adaptive: Integrating Self-Learning Red Forces into Existing Military Simulators,” paper submitted to ITEC 2026, London, April 14–16, 2026.

[4] Millenium Challenge 2002 remains one of the best cases illustrating this problem. This USD 250 million exercise, which took more than two years to prepare, involved over 13,500 service members from seventeen simulation locations. By using, inter alia, asymmetric tactics, Red Force Commander LTG Paul Van Riper landed major victories during the opening hours of the exercise. This prompted the Joint Forces Command to reset the exercise and confine the red force to strict constraints to ensure blue force success. See Micah Zenko, Red Team: How to Succeed by Thinking Like the Enemy (New York: Basic Books, 2015).

[5] Sponsored by the Centre for Digitalization and Technology Research, GhostPlay was funded by the NextGenerationEU programme. See “GhostPlay,” accessed May 2026, https://www.ghostplay.ai/.

[6] Department of Defense, Digital Engineering Strategy (Washington, DC: Office of the Deputy Assistant Secretary of Defense for Systems Engineering, 2018), 3, https://ac.cto.mil/wp-content/uploads/2019/06/2018-Digital-Engineering-Strategy_Approved_PrintVersion.pdf.

[7] The scenario under analysis takes place in Franconia, Bavaria, a region typical for Central Europe. While WILD HORNETS benefited from Bavarian open-source terrain data, this is unlikely to be the case elsewhere. Therefore, a development pipeline using photogrammetric procedures and Neural Radiance Fields was established to reconstruct and model three-dimensional infrastructure data based on remote sensing. See Heiko Borchert and Christian Brandlhuber, His Hands Can’t Hit What His Eyes Can’t See (Hamburg: Defence AI Observatory, 2024), 14, https://defenseai.eu/daio_study2426_his_hands_cant_hit_what_his_eyes_cant_see_borchert_brandlhuber.

[8] Borchert and Brandlhuber, His Hands Can’t Hit What His Eyes Can’t See, 13.

[9] WILD HORNETS ran approximately 1.8 million scenarios over about 3.5 weeks using a mixed CPU/GPU cluster consisting of 200 nodes.

[10] Hofstetter, “From Scripted to Adaptive.”

[11] This resembles Anthropic’s Claude system, which recently identified previously unknown cybersecurity flaws in IT systems. See Nicolas Carlini et al., “Assessing Claude Mythos Preview’s Cybersecurity Capabilities,” Anthropic Red Teaming, April 7, 2026, https://red.anthropic.com/2026/mythos-preview/. During WILD HORNETS, AI red forces similarly discovered potential vulnerabilities in actual air defence systems. See Borchert and Brandlhuber, His Hands Can’t Hit What His Eyes Can’t See, 27.

[12] Bryan Clark, Dan Patt, and Harrison Schramm, Mosaic Warfare: Exploiting Artificial Intelligence and Autonomous Systems to Implement Decision-Centric Operations (Washington, DC: Center for Strategic and Budgetary Assessments, 2020), https://csbaonline.org/research/publications/mosaic-warfare-exploiting-artificial-intelligence-and-autonomous-systems-to-implement-decision-centric-operations; Bryan Clark, Dan Patt, and Timothy W. Walton, Implementing Decision-Centric Warfare: Elevating Command and Control to Gain an Optionality Advantage (Washington, DC: Hudson Institute, 2021), https://www.hudson.org/national-security-defense/implementing-decision-centric-warfare-elevating-command-and-control-to-gain-an-

optionality-advantage; Tim Grayson, “Mosaic Warfare,” DARPA presentation, July 27, 2018, https://www.almendron.com/tribuna/wp-content/uploads/2019/07/sto-mosaic-distro-a.pdf; John Haystead, “DARPA’s Mosaic Warfare: Moving to Address the Ever-More-Rapidly-Paced Advances/Changes in Fielded Threat Capabilities,” Journal of Electronic Defense (February 2020): 20–25.

[13] Heiko Borchert, Christian Brandlhuber, Armin Brandstetter, and Gary S. Schaal, Free Jazz on the Battlefield: How GhostPlay’s AI Approach Enhances Air Defense (Hamburg: Defense AI Observatory, 2022), 11, https://defenseai.eu/daio_study2203.

[14] In this context, emergence emphasizes the ability of a complex system to adapt to unforeseen events, where formation rather than formulation matters. See Heiko Borchert, “The Very Long Game of Defence AI Adoption: Introduction,” in The Very Long Game: 25 Case Studies on the Global State of Defence AI, ed. Heiko Borchert, Torben Schütz, and Joseph Verbovszky (Cham: Springer, 2024), 1–38, https://link.springer.com/chapter/10.1007/978-3-031-58649-1_1.

[15] Borchert, Brandlhuber, Brandstetter, and Schaal, Free Jazz on the Battlefield, 17.

[16] This section summarizes the main findings of Borchert and Brandlhuber, His Hands Can’t Hit What His Eyes Can’t See, 17–34.

[17] Justin Bronk and Jack Watling, Mass Precision Strike: Designing UAV Complexes for Land Forces (London: RUSI, 2024), https://www.rusi.org/explore-our-research/publications/occasional-papers/mass-precision-strike-designing-uav-complexes-land-forces.

[18] High-explosive incendiary (HEI), semi-armour-piercing high-explosive incendiary (SAPHEI), frangible armour-piercing discarding sabot (FAPDS), and advanced hit efficiency and destruction (AHEAD) munitions.

[19] The simulation was conducted at unclassified levels using only open-source information.

[20] Stacie Pettyjohn, Evolution Not Revolution: Drone Warfare in Russia’s 2022 Invasion of Ukraine (Washington, DC: Center for a New American Security, 2024), 39–40, https://www.cnas.org/publications/reports/evolution-not-revolution.

[21] Field observations of Russia’s latest-generation V2U UAS suggest that “these drones operate as distributed, partially swarm-capable systems in which each unit processes information locally while remaining aware of nearby drones.” See Kateryna Bondar, How Russia Is Building a Sovereign Drone Ecosystem for AI-Driven Autonomy (Washington, DC: CSIS, 2026), 27, https://www.csis.org/analysis/how-russia-building-sovereign-drone-ecosystem-ai-driven-autonomy.

[22] Sarah M. E. Fortune, Xi Cheng, Keith Holmes, and Andrew W. Trites, “Cooperative Foraging Between Dolphins and Fish-Eating Killer Whales,” Scientific Reports 15 (December 11, 2025): 42897, https://www.nature.com/articles/s41598-025-22718-4.

[23] This argument does not dispute the relevance of mass. Rather, it advances a more nuanced proposition: AI tactics can compensate for lack of mass or multiply the effects mass can deliver.

← Book Review: The Cyber and Information Space Matrix

MORE POSTS

Are You REDy? Using Red AI Tactics And The Defence Metaverse To Augment Future Force Design July 16, 2026

Book Review: The Cyber and Information Space Matrix July 13, 2026

Academic Education For Officers' Training In A New Age Of Warfare July 9, 2026

Competing Narratives In Media Framing Of U.S. Arms Sales To Taiwan July 6, 2026

Dr Heiko Borchert is Co-Director of the Defense AI Observatory, Helmut-Schmidt-University/University of the Federal Armed Forces Hamburg, runs Borchert Consulting & Research AG, a strategic affairs consulting boutique, and is partner with the Economic Security Forum. He has been working on arms cooperation, corporate geoeconomics, defence, and military innovation for close to three decades. The views contained in this article are the author's alone.

The views contained in this article are the author's alone and do not represent the views of any institution affiliated with the author.