

Metis

# Study

Artificial intelligence in the armed forces No. 40 | April 2024

The views expressed in Metis Studies are those of the authors. They do not reflect the opinion of the Bundeswehr, the Federal Ministry of Defence, or the Bundeswehr University Munich. The primary target audience of Metis Studies are practitioners. Metis Studies are based on analyses of scholarly literature, reports, press articles and expert interviews with academics, think tank analysts and policy-makers. References are omitted. Inquiries about sources can be directed at the author(s) via email.

Institute for Strategy & Foresight



# Summary

rtificial intelligence (AI) can accelerate processes and improve decision-making. This study outlines a scenario of military use of AI based on the OODA loop. It then discusses risks of misuse of AI on the basis of the Centaur

and Minotaur models of warfare. Finally, the study advocates moving from talking about to using and especially experimenting with AI in order to be able to deploy useful, accepted and responsible AI solutions in the Bundeswehr in a timely manner.

#### Artificial intelligence in the armed forces

Since the last time uncrewed systems, automation and artificial intelligence (AI) were the subject of a Metis Study, it is not just generative AI that has made headlines for its ability to generate text, image and video. With the Russian invasion of Ukraine, the focus of discussions has come to firmly rest on the military use of AI.<sup>1</sup>

The generative AI products ChatGPT, Dall-E and Sora by the pioneer OpenAI are relatively immature. They were made publicly available, seemingly with no thought given to the associated copyright problems and particularly the consequences for the global information landscape and social decision-making, especially in open, pluralistic democracies.

Military use of AI must of course meet significantly higher standards in terms of robustness, accountability and reliability, especially if it is directly or indirectly linked to the use of weapons and their effects. The automated target recognition system of the Russian loitering munition Lancet, which was recently implemented through a software update and which automates target selection and engagement and renders Lancet invulnerable to jamming as it approaches a target, would never have been considered acceptable for service use by NATO forces. The same is most likely true for similar object recognition software used by the Ukrainian side.

Against the backdrop of these examples and the ongoing Russian invasion of Ukraine, the discussion surrounding military use of Al is often limited to the automation of the final steps of the kill chain, i.e. to autonomy of weapon systems when it comes to selecting and engaging targets without human intervention. This is understandable, as this automation of violence is where great military potential intersects with highly controversial political, legal and ethical aspects.

But when other possible operational contexts of AI in the military are systemised on the basis of the OODA (observe, orient, decide, act) loop, it is clear that even the "act" step of the loop does not always have to involve weapon effects. Autonomous systems for evacuation operations – such as the "Grille" drone, which is currently being tested by the Bundeswehr and can carry up to 700 kg to quickly fly casualties out of combat zones – are just one example of an appropriate use of automation at the end of the decision loop.

See "The security-policy effects of digitisation: Future forms of conflict and conflict management", Metis Study No. 1 (February 2018) and "Uncrewed systems: armaments, control and arms control", Metis Study No. 28 (June 2022). The cover images of Study No. 28 and of this Study No. 40 were both created by Midjourney. The prompt was the same for both. The difference in quality of the resulting images illustrates the rapid development in the field of Al-based image generation. See also Figure 1.



But there are more areas in which AI can play to its strengths as well as optimise and accelerate processes, including more efficient and resilient logistics, in cyber defence, for predictive maintenance, in collecting, merging and automatically analysing large amounts of data to spot developing crises early on, in military intelligence as well as in combat service support and operational command and control to provide a better situation picture for mission planning and operations.

To illustrate the possible future of responsible military use of AI, this study will use the OODA loop as a structure on which to project an ideal scenario of AI being used for an infantry operation of the Rapid Response Forces Division.





# The OODA-Loop





#### **Observe**

Using AI, information provided by drones that conduct reconnaissance along pre-planned flight routes is combined with other informa-

tion from radios, optronics, helmet cameras, official smartphones and other available data sources, including satellite images. At the level of the battle management system, the situation picture for the command and control system is thus enhanced, which facilitates the conduct of operations in the second step. Ultimately, the benefits of AI already result in increased survivability in this first step - but especially for the soldiers on operations, who, in the final step of the loop, are primarily concerned with combat and with their own survival.



#### Orient

Thanks to continuous blue force tracking and the ability to report enemy contact with the push of a button, digital command and con-

trol systems have already rendered the cumbersome maintenance of a situation map obsolete. With Al, the image data obtained during the observation step can be processed more quickly, allowing terrain, individual buildings or entire streets to be visualised and rendered as 3D models. Information about covered approach routes, possible breaches or landing zones can quickly be requested using voice control. Sandboxes and cardboard models are no longer necessary. Al can also be used to classify reconnaissance results and prioritise them for targeting. This speeds up the command and control process as well as completion of the kill chain. Soldiers who were previously needed for data processing can now be employed for more high-value military tasks.



# Decide

Human creativity, the ability to abstract and interpret information, but also intuition, charisma, empathy and leadership by example allow people to guide and inspire others, i.e. soldiers, in

a way that machines cannot. AI can support human decision-making, however, by not only highlighting different ways of acting but also by simulating them to their logical conclusion and systematically comparing them. Above all, however, AI can be a supportive asset wherever human decision-making is affected by "typically human" cognitive idiosyncrasies such as bias and "predictable irrationality". Instead of humans' using their own diffuse ideas - or even unrelated considerations, such as certain idiosyncratic preferences of the commander - as the basis for charting courses of action, AI can be used to more systematically and transparently compare a plan against the given constraints and risk assessments and to simulate contingencies to their logical conclusion. The experience gained from using AI-based systems for operational command and control allows more realistic estimates of space and time in operational planning and more precise synchronisation of different assets. None of the people involved has to fear a bad assessment because of content produced by AI.



# Act

Once the human commander has made a decision, command support can use AI to more

quickly write up the relevant order, visualise it and communicate it to subordinate units. Since the receiving side breaks down orders along defined patterns on its command level, this step can also be prepared, supported and accelerated with AI. Returning once more to the level of the individual soldiers on operations: if the individual elements still require reconnaissance results or terrain data, these no longer have to be requested vertically from the superior command but can be obtained horizontally with the aid of AI - "What's the weather forecast for tomorrow and where can we attack with the best possible cover?"



#### **Opportunities and risks**

The main advantage pursued by the military in its use of AI is better and faster detection and decision-making, i. e. faster cycling through the OODA loop – ideally by delegating to the machine those tasks that it is particularly suited for, while having humans take on those tasks that they are uniquely qualified for or which they have to remain in charge of for regulatory reasons. A civilian equivalent would be an oncologist who uses AI to more reliably diagnose skin cancer but then personally makes treatment decisions in consultation with her patient. There are two risks to be considered when it comes to using AI.

The first risk arises from a false understanding of what military AI can do and really contribute to decision-making in interaction with people. The OODA loop should also serve only as an ideal example, one which highlights options but is not intended to further fuel the hype around Al, which misunderstands it as a military universal weapon. Software and AI that is based on machine learning and equipped with a sufficiently large database<sup>2</sup> excel in optimisation tasks and pattern recognition, so they can support humans with a) more reliable predictions, b) filtering out relevant information, and c) visualising complex interactions. Since every human decision is based on a notion of the future – and a desire to shape it – it stands to reason that a more accurate idea of possible futures and a broader portfolio of options for action courtesy of AI will facilitate better decisions.

However, machines *should not* make decisions *for* military commanders, because they cannot grasp the legal and ethical consequences, let alone bear them. In fact, they cannot (yet) do so, technologically speaking, because content generated by Al's inductive reasoning is not to be equated with human cognition and the ability to apply abductive reasoning. Machines do not actually "decide" anything in the proper sense of the word – which is why the term is used in this study only for lack of a better word and should not be taken literally – not least because they cannot independently imagine futures or reflect on and justify their "decision".

Moreover, military dilemmas (in combat) rarely result from black-and-white, clear-cut routine problems, i.e. the optimal conditions for automating entire processes such as those found in warehouses, payment transactions or when sorting photos on a smartphone with facial recognition – conditions that are thin on the ground in war. Thanks to thousands of years of evolution, the human brain is good (and energy-efficient) at imagining futures in complex, unpredictable situations where intricate, elusive and data-poor problems require context-dependent and

2 The collection of sufficient data can be a much greater problem in military contexts than in civilian applications. As a result, AI often has to draw on synthetic data, which comes with its own challenges.

situational decision-making. Humans are thus (at least for now) unique in that our flexible intelligence allows us to overcome such challenges. By offering multiple predictions, information filters and visualisation aids, AI can compensate for some of our known weaknesses. No more, no less.

The second risk lies in the fact that, even with an appropriate understanding of the strengths and limitations of AI (whatever its current level of development), the ratio of the overall mix of human and machine can still be off. As a result, military potential may go to waste and human judgement and control may be lost, which in turn entails legal, ethical and security risks.<sup>3</sup>

#### **Centaurs vs. Minotaurs**

Metaphorically speaking, the military would ideally achieve a system of AI-augmented decision-making and a human-machine team which amount to a centaur, the figure of Greek mythology with a horse's body and a human head - human reasoning, judgement and responsibility combined with a machine's many senses (sensors) and extraordinary muscle power (effectors). One example would be a loitering munition that, as part of a sensor network, helps to reconnoitre a military target and then, after critical examination and a decision by a human who is familiar with the operational context and the current situation, engages it in accordance with international law of armed conflict and the rules of engagement, with the human bearing all ethical and legal responsibility. The human in this scenario improves their military effectiveness and reduces their own physical risk; the machine, in this case an expendable drone, can be "sacrificed".

In keeping with the theme of Greek mythology, the Minotaur, a human body with a bull's head, would represent the outcome of a negative development: the machine is in charge, while the human succumbs to automation bias and cedes control, merely simulating human oversight and no longer making decisions on the basis of precise situational knowledge and human judgement, yet still formally responsible for whatever happens. An example would be a battle management system that automatically generates targets and presents them, in quick succession, to a human to approve for engagement, putting the human under pressure to the point that they will eventually trust the machine blindly and give the "green light" every minute or every second without fully grasping the context of the operation and the precise circumstances surrounding the "why, what, who, where, when, and how?" of the resulting effects.<sup>4</sup>

<sup>3</sup> The details were discussed in previous Metis studies. See also: "What degree of human involvement should there be in the use of force?", Metis Interview No. 4 (April 2021).

<sup>4</sup> One example from a civilian context: large, automated logistics centres where people have zero cognitive involvement in decision-making and, like robots, merely fetch, bring and pack as instructed by the algorithm.





If own forces were to be sent on dangerous missions on the basis of machine-generated orders, the Centaur situation would be turned on its head entirely: the machine, whose existence is not particularly valuable, would endanger valuable human life.

> A human-machine relationship modelled after the Minotaur would lose what experts call meaningful human control and which should rightly be preserved. This control paradigm should not be confused with mandatory

remote control. It merely means that, when using automation, human operators in the human-machine team must be able to a) understand how it works and predict its effects (situational awareness), b) step in to administer the process again at any time, if necessary, and, as a result, c) take legal and ethical responsibility for the consequences.



Centaur		Minotaur	
Human	Machine	Human	Machine
Attributes: (moral, legal, and political) judgement, contextual knowl- edge, flexible intelligence to overcome complex and unforeseen challenges	Attributes: speed, potential for optimisation, expendability	Attribut: versatile interface to the physical world	Attributes: speed, potential for optimisa- tion, decision superiorityt
Tasks: understand, predict, assess, mon- itor, control, take responsibility	Tasks: collect, fuse, process and present data	Tasks: take responsibility but with- out real human control	Tasks: collect, fuse, process and priori- tise data, predetermine decisions (automation bias)
The human controls the machine		The machine directs the human	
Table 1 Centaur vs. Minotaur warfare. Author's own work, based on: Paul Scharre 2016: Centaur Warfighting: The False Choice of Humans vs. Automation, in: Temple International & Comparative Law Journal 30, 151–165 & Sparrow, Robert J./Henschke, Adam 2023: Minotaurs, Not Centaurs: The Future of Manned-Unmanned Teaming, in: The US Army War College Quarterly: Parameters 53: 1, 115–130.			



## The benefits of national regulation

The Bundeswehr urgently needs a policy document to regulate the responsible handling of AI and automation, especially with regard to the use of autonomy in the critical functions of target selection and engagement.<sup>5</sup> Such a guideline for the Bundeswehr, the content of which should remain at a fairly abstract level – establishing guiding principles without delving into specifics of individual case studies – as it stipulates meaningful human control over Bundeswehr weapon systems, would have several advantages:

- Political rapprochement and military interoperability with alliance partners, many of whom already have such doctrines, would be facilitated.
- Cooperation with industrial partners would be improved as it would provide them with planning certainty and politically defined "guardrails" for research and development.
- Communication with the wider public could become more objective as the Bundeswehr could openly disclose its military requirements and opportunities while signalling that it recognises and avoids risks.

If the policy document is designed to be future-proof, there is no cause for concern that available options could be limited through premature national regulation. Such future-proofing is achieved by focusing on abstract principles for military human-machine interaction based on political guidelines, the applicable legal framework and guiding values. The policy document could also be updated whenever necessary, much like similar documents of Alliance partners such as the US, the UK and the Netherlands. The fact that EU and NATO partner nations have had such rules in place for years also shows not only that the political, legal and ethical risks have largely been studied and determined but also that, at least in the West, the established ideas of what the necessary "guardrails" for automation and human control should be do not differ all too widely.

While the Bundeswehr has some catching up to do when it comes to AI in the wake of the *Zeitenwende*, it could roll out useful, responsible technology more quickly if innovation and procurement cycles were sped up, which would require the armed forces to be provided with the necessary IT infrastructure as well as more funding, agility and opportunities for experimentation.

5 See "Uncrewed systems: armaments, control and arms control", Metis Study No. 28 (June 2022).



Fig. 1 Comparison of two images generated by the Midjourney Al image generator in response to the same prompt. Left: generated for Metis Study No. 28 in August 2022; right: generated for Metis Study No. 40 in May 2024. See also footnote 1



#### Useful, accepted and responsible AI in armed forces

Despite the hype, much of the military use of AI, even in other armed forces, remains at the theoretical level or in the early phases of testing. Systematic testing is the only way to find out what is really *useful* for military purposes, *accepted* by the force and politically, legally and ethically *responsible*, i. e. ensures predictability, administrability and accountability. After all there is no *one-size-fits-all* blueprint for meaningful human control when AI is used in the armed forces – control must be differentiated, structured according to the operational context and then trained with appropriate tactics, techniques and procedures.

Training and red teaming will not work in the abstract. Finally, despite the great potential of AI, we must not lose sight of a) the fact that the responsible handling of AI requires carefully considered interface ergonomics as well as new training and exercise concepts in order to prevent automation bias when dealing with probabilistic output of AI, and b) that AI also introduces specific new sources of error and Achilles heels that can be targeted by the enemy. The old adage of data science remains true: garbage in, garbage out. Training data sets for AI may also be distorted by bias. Adversaries can also potentially interfere with any step of the OODA loop, from compromising the "observe" step at the very beginning with information and cyber operations to, in the "act" step, misleading expensive automated targeting systems with potentially the simplest of means.<sup>6</sup> The trust of the armed forces in new systems is one last aspect to consider when it comes to the use of Al in armed forces, and it is unfortunately often underexamined. Over the past ten years, social scientists have repeatedly surveyed members of the US armed forces with the same surprising result: soldiers are *more* sceptical about increasing autonomy in weapon systems than the broader public. Practical tests are therefore crucial in this regard too. Once they are used in training scenarios, Al solutions that have prominently been discussed on paper and in PowerPoint presentations may turn out to be unsuitable in practice because they are not accepted by the troops. By the same token, other useful applications can sometimes only be achieved through creative experimentation in the context of an exercise.

The question of what is useful, accepted and responsible in terms of military use of AI cannot be definitively answered at a desk, in a working group or during the umpteenth panel discussion. When it comes to AI, the primary goal of the Bundeswehr should therefore be to translate words into actions.

<sup>6</sup> See "Uncrewed systems: armaments, control and arms control", Metis Study No. 28 (June 2022) and "Trends and developments in hybrid threats", Metis Study No. 35 (June 2023).



# IMPRINT

# Publisher

Metis Institute for Strategy and Foresight University of the Bundeswehr Munich Web: metis. unibw.de Bluesky: @metis.unibw.de

#### Author

PD Dr. Frank Sauer

Julian Werner

metis@unibw.de

### **Creative Director**

Christoph Ph. Nick, M.A.

zum-staunen.de

# **Image credits**

#### Cover:

"European soldiers in dense forest with tanks and drones" C. Nick, motif created with the help of Midjourney.

Inside pages: All motifs created with the help of Midjourney. | C. Nick

#### Translation

Federal Office of Languages

# **Original title**

Künstliche Intelligenz in Streitkräften

#### ISSN-2627-0609

This work is licensed under the **Creative Commons Attribution 4.0** International License.

