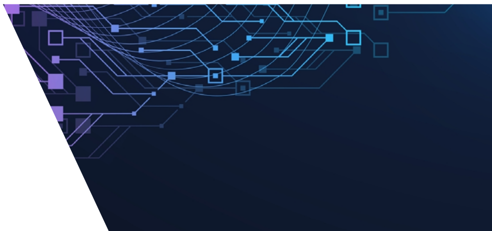
Federal Ministry
Republic of Austria
Defence

verteidigungspolitik.at

Artificial Intelligence in National Defence



Contents

The contents of the individual contributions reflect the personal views and opinions of the respective authors and do not necessarily correspond to positions of the Federal Ministry of Defence, or the institutions for which they work.

Imprint
Republic of Austria
Federal Ministry of Defence

Media owner, publisher and producer: Federal Ministry of Defence, Rossauer Laende 1, 1090 Vienna, Austria

The coordination of this publication was carried out by the Defence Policy and Strategy Division.

Project management and layout: Raphael Spoetta

Coordination of Content: David Song-Pehamberger

Editorial team: Walter Matyas, Miriam Gruber, Laura Reis

Photos and graphics: Courtesy of authors, Shutterstock

Print: Armed Forces Printing Centre, 1030 Vienna, Austria

Published 2025

Please collect used paper for recycling.

ISBN: 978-3-902275-63-9

Shutterstock

Editorial by the Secretary-General

Dear reader,

Arnold H. Kammel

We are in the midst of a global and cross-societal digital transformation, characterised by increasing connectivity and ever-growing computational power. The integration of artificial intelligence (AI) into digital processes of all kinds is now causing this transformation to advance at an increasingly rapid pace and with enormous leaps in efficiency. While this offers great opportunities, it also poses serious risks.

AI offers remarkable opportunities for complex, global challenges for which existing analytical models are insufficient. AI models offer unprecedented possibilities for evaluating and analysing large amounts of data, which can help to achieve new breakthroughs in medicine or in understanding and mitigating the consequences of climate change, for example. Furthermore, the increasing automation of countless processes in industry and administration is an opportunity to better counteract demographic change.

However, AI also exacerbates security risks in our society. A geopolitical race between global

Dr Arnold H. Kammel has been the Secretary-General at the Federal Ministry of Defence since 2022.

players for supremacy in this and other technological domains has already begun and is continuing to gather momentum. Increasingly complex and powerful AI models are thus also intensifying systemic competition and jeopardising international peace.

National Defence too is confronted with the reality of this social transformation and its challenges. To address this, the Federal Ministry of Defence (MoD) of Austria has adopted its first AI Strategy in 2024. This strategy addresses the extensive topic of AI while also covering essential components of digitalisation, including cyber defence and data security.

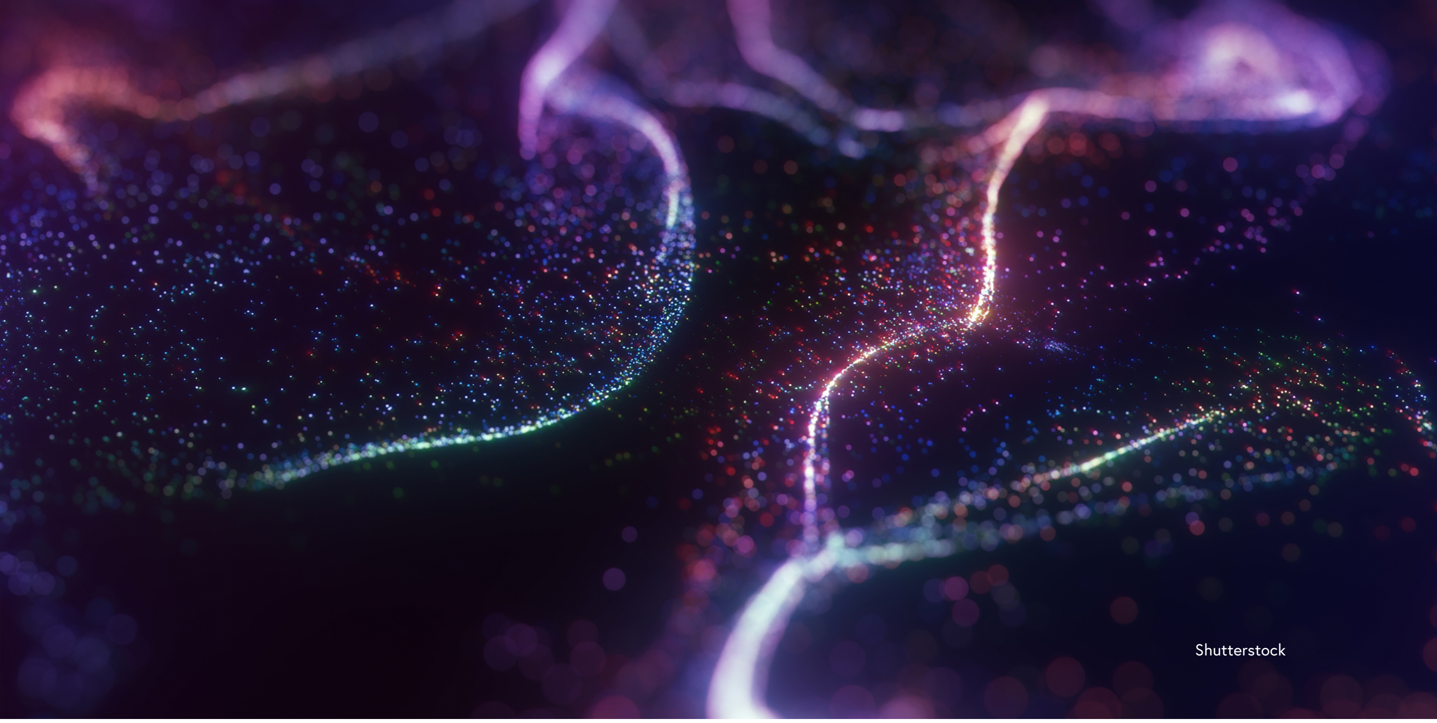
Given that even the best strategy is futile without conscientious application, a holistic implementation is a particular focus of the MoD AI-strategy.

To this end, the strategy includes a ten-year implementation horizon, in which AI is to be applied gradually, sustainably, and in a risk-based manner in the Austrian Armed Forces (capability development) and in the MoD (administration).

The usage of AI is not an end in itself, but a deliberate aspect of the digital transformation of Austrian National Defence. Concerns about data security and privacy, as well as ethical issues raised by certain AI applications, must be carefully evaluated before these applications can be introduced. In addition, risk assessment and strategic foresight with regard to AI and other emerging technologies will be a particular focus in the coming years in order to be prepared for the upcoming technological turning points, including their opportunities and risks.

Because one thing is certain: The digital transformation is happening, and it is up to us whether we want to take advantage of its opportunities and face the risks prepared, or whether we allow ourselves to be overwhelmed by it.

Sincerely,
Arnold H. Kammel


Shutterstock

The AI-supported transformation of the Armed Forces

Rudolf Striedinger

We are living in a time of technological change. The rapid advancement of digitalisation of our society leads to increasing connectivity between humans and machines. With the help of more extensive sensor networks, this connectivity is generating huge amounts of data. The integration of artificial intelligence (AI) is now unlocking the potential of large amounts of data and interconnected Information and Communication Technology (ICT) systems, contributing to the digital transformation of all aspects of society.

This change naturally also has a significant impact on military national defence and thus on the Austrian Armed Forces (AAF), as digitalisation offers both enormous opportunities and significant risks for the armed forces. For this reason, armed forces around the world are currently undergoing a phase of digital transformation. This transformation affects all domains – land, air, sea, space, as well as cyber and information space – and all levels, from the strategic, to the operational, to the tactical and combat level.

Significance for the future battlefield

The digital and AI-supported transformation of the armed forces is a strategic necessity in order to master modern combat. AI-supported networking offers the possibility of connecting previously separated domain structures via so-called Multi-Domain Operations, thereby reducing response times and greatly increasing effectiveness to meet the requirements of the future battlefield.

Not only command structures, but also logistics and maintenance will increasingly benefit from sensor fusion, which increases the efficiency of deployment, while simultaneously extending the service life of equipment and gear. AI-enhanced data analysis enables significant efficiency gains in reconnaissance and situational awareness, as well as in early crisis detection and strategic and tactical planning.

Increasingly sophisticated applications of robotics and autonomous systems also offer a wide range of possible use cases, particularly in areas that pose exceptional dangers to soldiers (e.g. demining) or where human reaction times are insufficient (e.g. missile defence). During ongoing military conflicts, such as Russia's war of aggression against Ukraine, it is becoming clear that drones already play an enormous role on the battlefield in both reconnaissance and combat, which will only increase in the future.

Impact on the capabilities of the Austrian Armed Forces

AI systems, apart from robotics, are predominantly software-based models. Accordingly, the field of cyber security and cyber defence is already benefitting enormously from the use of AI to protect ICT systems. AI will also become an increasingly important tool in defence against AI adversaries, as AI systems are already being used for automated cyber-attacks.

In addition, the use of new digital tools for training, education, and exercise creates clear added value for AAF soldiers. Generative AI also offers a wide range of applications in these and other areas, such as language translation, language and text analysis, and information extraction. Finally, the potential for automating certain administrative processes should not be overlooked. In all these areas, it is crucial to emphasise that AI must always be used responsibly and to support, rather than replace humans.

The AAF 2032+ Development Plan provides a roadmap for the integration of transformative technologies to build and align the AAF for the future. This plan lays out the military-strategic framework that, with the provision of the necessary resources, will enable the AAF to comprehensively protect Austria in the coming decade. Despite the allocation of significant resources for the development of the AAF and its transformation for a future shaped by emerging and disruptive technologies, we – like all armed forces worldwide – face significant challenges in the implementation of AI.

Long-Term disruptive developments

Developments on the battlefield, as well as the latest state of the art in technology and innovation are currently advancing at such a rapid pace that significant technological breakthroughs are happening annually. This trend of rapid development cycles continues to accelerate. However, the size and equipment of armed forces require balanced and long-term planning, based on implementation horizons measured in decades. To keep up with the rapid development of disruptive technologies, planning and procurement cycles must be accelerated and made iterative, while taking into account the established aspects of stable and risk-based military foresight and planning. In short, we must become faster and more flexible without compromising our own security.

However, apart from accelerating processes, other fundamentals must also be taken into account to make the digital transformation effective and sustainable. Networked systems must be kept interoperable, both within the framework of national defence structures and with partners. To this end, international cooperation is a key element for the AAF in order to participate in and benefit from areas such as standardisation, research and development, and joint procurement. For Austria, the framework of the EU's Common Security and Defence Policy (CSDP) is particularly important.

Cross-structural and interlinked systems also raise questions of vulnerability and susceptibility to attacks via cyber space. Therefore, it is necessary to implement the digital transformation of the armed forces based on the principle of security by design. This means that particular attention must be paid to the system architecture and the accreditation

General Mag. Rudolf Striedinger has been the Chief of the General Staff of the Austrian Armed Forces since 2022.

of all ICT systems. No compromises can be made here, even if this leads to delays in operational readiness.

The digitalisation of the armed forces is a process that affects us all, and no one can or should be exempt from it. AI will support this process. It is up to us how responsibly we will deal with it in the future, without restraint, but with the necessary respect for the dangers involved.

Shutterstock

Artificial Intelligence and hybrid threats

The domains cyber and space

Artificial intelligence (AI) has the potential to support conflict resolution—in conventional and asymmetric warfare, as well as in the cyber and space domains.

Josef Schröfl

AI in the conduct of hybrid conflict

The term hybrid threats refers to actions by state or non-state actors that seek to weaken or damage a target (e.g. a state or society) through a combination of overt and covert military and non-military means. Hybrid conflicts, which are characterised by a mixture of conventional warfare, irregular and asymmetric tactics, cyber operations, and information warfare, have posed unique challenges to national security policy for the past several years. In this context, AI is proving to be a powerful tool with the potential to support policy-making, strategy development, early warning and intelligence systems, as well as post-conflict mediation efforts.

AI is already utilised in conflict and war scenarios. Examples include the use of so-called “deepfakes” in the course of Russia’s war of aggression against Ukraine. Manipulated videos, such as one showing the Ukrainian president allegedly insulting European allies during a speech, are being used to spread disinformation among the Ukrainian population. The aim is to undermine morale and damage the government’s reputation. Furthermore, a cyber attack targeting the KA-SAT satellite network operated by ViaSat took place a few hours before the start of Russia’s invasion. This network was used by the Ukrainian Armed Forces for their command and control systems (C2), among other things. However, thousands of civilian consumers across the European continent were also affected, including critical infrastructure.

Potential application for conflict resolution

AI-driven early warning and intelligence systems play a crucial role in escalation management. These systems analyse vast amounts of data, including social media posts, satellite images, and communication patterns, to identify anomalies or patterns associated with potential conflicts. By providing timely warnings, AI supports political decision-makers, as well as already deployed troops, in taking preventive measures.

Conflict resolution often involves complex decisions. AI can support decision-makers by analysing historical data, assessing risks, and suggesting optimal courses of action. AI models can recommend strategies for negotiations based on historical facts, while also taking cultural differences and the real-life experiences of those involved into account.

Diplomats who mediate settlements between conflicting parties often face the challenge of finding common ground between them. AI can suggest negotiation strategies, simulate possible outcomes, and identify areas for compromise. By analysing text data from negotiation transcripts and historical treatises, AI can highlight points of convergence and divergence, thereby supporting negotiators in their efforts.

AI can also identify opportunities for dialogue, reconciliation, and confidence-building measures. Natural language processing (NLP) algorithms analyse speeches, interviews, and public statements, as well as social media, to assess the mood of the parties involved in a conflict (society, politics, etc.), identify common values, and recommend confidence-building measures based on this information.

AI models simulate scenarios of conflict and war, enabling political decision-makers to weigh up various possibilities for conflict development and assess the consequences. By adjusting variables, including troop movements, economic sanctions, or cyber attacks, they gain insight into their potential effects. These simulations are used for strategy development as well as crisis management.

Transparency and logical explainability are crucial if AI is ultimately to be used in a trustworthy way. Transparent AI systems are the goal of democracies, whereas in autocracies, AI tends to be used tacitly and without society's knowledge. The European Union has already set an example here with its Artificial Intelligence Act (AI Act).

It is of utmost importance to ensure that AI systems comply with ethical guidelines. Bias, discrimination, and unintended harm must be minimised. However, AI algorithms also learn from historical data, which may contain biases. To mitigate these, AI must be continuously trained and monitored. Effective conflict resolution requires a balance between human expertise and AI capabilities. Although AI can process vast amounts of data, human judgment, empathy, and cultural understanding remain irreplaceable.

AI in the domains cyber and space

Digitalisation has increasingly linked the domains of outer space with the cyber and information space. Satellites, ground stations, and user terminals are increasingly exposed to cyber threats. Understanding the connections between cyber space and outer space is critical to protecting the space assets that society relies on.

Just like any digital device, satellites can be hacked. However, because they are so far removed from most people's everyday lives, their importance and society's reliance on critical space infrastructure can easily be overlooked. A cyber attack on a satellite can simultaneously disrupt financial markets, road traffic, weather forecasts, internet connections, power grids, air traffic control, and military operations.

Colonel Josef Schröfl is Deputy Director at the Hybrid Centre of Excellence in Helsinki, Finland.

Astronomers often have to deal with huge amounts of data from telescopes and satellites. AI helps process this data, cleans up noisy images, and extracts useful information. For example, AI has improved our knowledge of the largest black hole in the centre of the Messier 87 (M87) galaxy, providing a clearer view of its structure. AI can also streamline mission planning by optimising flight paths, resource allocation, and scheduling. It also improves satellite efficiency by automating tasks, including optimal satellite positioning.

In conclusion, AI is revolutionising space exploration, making it faster and more efficient, and enabling discoveries that go beyond what humans alone could have achieved.

As the threat landscape in space evolves, so too should the understanding of cyber risks and mitigation measures to protect space assets and the wide range of services they provide to society. Future challenges in space-based cyber security will include closing the skills and information gap and figuring out how to most effectively conduct and respond to cyber operations in space.

In addition, AI can play an important role in resolving hybrid conflicts. Despite all the obstacles, the responsible use of AI can contribute to improving conflict prevention, decision-making, and peacebuilding.

Shutterstock

AI in the military and on the battlefield

Outlook and future trends

The use of artificial intelligence (AI) in armed forces and on the battlefield will not only grow linearly but exponentially over the next 10 years. This fact is currently still systematically underestimated. The step from “combined arms warfare” to “all domain mosaic warfare” is a big one that will be dominated by AI on all levels.

Joachim Klerx

AI not only increases the efficiency and speed of combat, but also the speed and efficiency of developing new methods of combat. Civilian AI systems are widely regarded as the most important technology of the future, but also as the greatest future threat. Even leading voices in AI development, including Sam Altman (OpenAI) and Elon Musk (xAI, Tesla), warn of the risks of AI and dangerous innovations. The same applies to military AI systems, especially when it comes to how these developments will alter the battlefield.

The near term

The current trend is toward increasing the efficiency, performance, and “intelligence” of AI systems. Current systems with so-called “narrow AI” are optimised for specific tasks. The main characteristic of narrow AI is its specialisation. These systems can only perform the tasks for which they were programmed and are not capable of learning or acting beyond that. Their advantage is that they can process real-time information from various sensors and sources and make decisions rapidly, based on complex data.

In military applications, narrow AI is used for reconnaissance, automated surveillance, target detection, and in the cyber domain. Other examples include robots equipped with AI that assist humans or replace them in dangerous environments. The optimisation of supply routes and inventory is another key area of application, alongside many others.

Newer, general systems are capable of adaptive learning with extensive contextual knowledge and statistical methods. Through data analysis, they can make predictions about enemy movements and derive both tactical and strategic decisions for missions and operations. In the near future, AI systems will increasingly evolve from supporting to taking over human tasks, as they can analyse, predict, plan, and execute with greater efficiency and optimal use of resources. Increasingly autonomous weapon systems will appear on the battlefield, surpassing soldiers with their corresponding weapon systems.

Medium term

The transition from narrow AI to artificial general intelligence (AGI) represents a massive leap forward, the exact path of which has not yet been fully explored. AGI refers to a form of AI that is capable of demonstrating human-like cognitive abilities, i.e. flexibly solving problems, learning, understanding, planning, and adapting to new situations – regardless of the field of application. AGI does not yet exist.

It remains to be seen whether AGI systems will consist of unprecedented neural networks modelled on the human brain or whether they can be constructed in a modular fashion from a combination of limited narrow AI systems. As far as we know at present, this means that military AGI systems will develop situational awareness tailored to their area of responsibility, which will serve to support their respective capabilities in terms of forecasting and decision-making. These capabilities will help to support or replace different levels of leadership in planning.

Regardless of the specific technological solution, general AI will eventually assist military leaders in optimising tactical and strategic planning. This will be achieved through sensor data from so-called "smart devices," i.e. electronic devices equipped with sensors, processors, and often an internet connection. These devices are capable of independently collecting and processing data and communicating with other devices or users.

Parts of US operational planning are already being simulated and optimised using AI. A military AGI would be able to do this in real time. However, this also increases the risk of a "flash war," a special form of mutually escalating AI systems that react catastrophically to the trigger values of the other's systems. This phenomenon first appeared in high-frequency stock market trading and should be considered a risk in the development and usage of military AGI. AGI systems are likely to appear as "digital twins" or "digital companions" on the battlefield and take over planning in command and control systems.

Long term: Superintelligent military AI systems

The move towards military AGI appears difficult, but has predictable consequences. The move from AGI to ASI (Artificial Super Intelligence; a type of AI that is far superior to humans in all cognitive abilities), on the other hand, will probably seem small and insignificant, but will have largely unpredictable consequences. The basis for this could be, for example, a military platform that integrates 'smart devices' on the battlefield into a Joint All-Domain Command and Control (JADC2; networking and integrated command across all domains of operation), such as LatticeOS (Anduril). Since no AI with ASI capabilities exists yet, the military impact is difficult to estimate.

Lethal autonomous weapon systems could become more effective through the combination of smart devices, the Internet of Battlefield Things (IoBT) and ASI. ASI would take the lead and control unmanned drones and smart devices in real time, enabling them to adapt to changing conditions and carry out their missions

Dr Joachim Klerx works at the Austrian Institute of Technology with a focus on research and development of AI to support strategic foresight.

efficiently and accurately. Neuromorphic computers and quantum computers could further reduce the response times of these systems and improve their decision-making capabilities.

In cyber warfare, an ASI could take both defensive and offensive operations to a new level of intensity. It could detect and repel cyber attacks in real time, find and exploit vulnerabilities in enemy systems, and develop and deploy new types of sophisticated cyber weapons. The enemy's AI systems would be a prime target.

A state that does not participate in the competition for the best military AI could fall significantly behind militarily and strategically. Technologically advanced states will be able to process information faster and more accurately by using AI, leading to better strategic decisions and faster response times. Without these technologies, a state could lose its defence capabilities, making it more vulnerable to modern threats such as cyber attacks and hybrid warfare. Overall, a state that lags behind in AI development could lose strategic, economic, and foreign policy importance.

Shutterstock

Military applications of AI

The use of artificial intelligence (AI) in the armed forces has the potential to enhance efficiency and effectiveness in all areas. From strategic and operational planning to the tactical environment and far beyond deployment, AI applications will gradually contribute to the digital transformation of national defence. This transformation is already underway and will take several decades to complete. Herein, a distinction must be made between automation and autonomisation, with the majority of the short- to medium-term benefits of implementing AI models resulting from the automation of existing processes.

David Song-Pehamberger

Command and Control systems

Military command and control (C2) systems have always relied on timely, high-quality data for decision-making and effective communication channels for transmitting orders. AI-supported systems and sensor networks now enable the effective, cross-domain networking of previously separate C2 systems. In real time, data can be transmitted from the tactical level, analysed using AI, and made available for command support at the operational and strategic levels.

The potential speed of these processes far exceeds what is possible for humans, giving decision-makers and analysts a better overview and allowing them to focus on the important decisions. Decisions made by the commander can also be directly incorporated into the targeting cycle. This enables multi-domain operations (MDO) that go far beyond traditional decision-making and communication chains. However, this requires access to large amounts of high-quality data via sensor networks and reconnaissance.

Intelligence and reconnaissance

AI systems already have a major impact on the scope and quality of information provided by intelligence services and military reconnaissance, especially in the field of intelligence, surveillance, target acquisition and reconnaissance (ISTAR). This is also a basic requirement for MDO systems. Due to ever-growing sensor networks and data streams, intelligence services are faced with the challenge of having to search through large amounts of information. AI systems can support automated pattern recognition to dramatically increase both the effectiveness and efficiency of ISTAR. One example is the use of satellite-based sensors for the automated detection of military-relevant troop concentrations or movements.

Apart from military sensors, AI models also support the field of open-source intelligence (OSINT). This refers to the analysis of large, publicly available data sets. OSINT reconnaissance uses data from social media, traditional media networks, various websites, and other data streams to identify relevant patterns and use them to assess the threat situation, for anticipatory planning or for direct command support. Ukraine has extensively used OSINT on the battlefield, inter alia to determine the location of Russian troops or to identify fallen soldiers.

Strategic planning and early warning

AI-supported crisis and early warning systems that are capable of identifying patterns at an early stage and informing commanders about potential threats are particularly important. Strategic planning can benefit from the potential of AI-supported analysis. This also applies to resource planning and integration with other fields, including logistics, procurement and C2 systems. It is important to note that complex AI systems must be transparent and base their conclusions on high quality and relevant data. Furthermore, the final analysis and assessment must always be carried out by experienced experts.

Cyber defence

Cyber defence is an area in which AI systems already play a significant role. Deep learning models can be used, for example, to continuously and automatically monitor IT systems and networks to detect anomalies that indicate malware and cyber attacks. In addition, large amounts of data and software code can be analysed to identify threats. AI systems can also be used to uncover vulnerabilities in one's own systems and to simulate cyber attacks in order to strengthen overall cyber resilience.

AI can make an important contribution to cyber defence. However, it is also increasingly being used offensively to generate malware and for automated cyber attacks. Therefore, the effective use of AI in cyber defence also requires detecting and defending against such AI-based cyber threats.

Robotics and autonomy

Although the field of robotics is not necessarily based on AI, the use of complex AI systems enables vehicles (e.g. drones) to operate with increasing autonomy. Whether on land, at sea, or in the air, transporting a vehicle over rough terrain to a selected destination requires a certain degree of autonomy. The exact degrees of autonomy have not yet been universally defined, but the terms 'human-in-the-loop,' 'human-on-the-loop,' and 'human-in-command' are commonly used, depending on whether the vehicle requires human involvement permanently, only during certain steps, or not at all once commands have been issued. The latter degree of autonomy could be referred to as fully autonomous systems, but there is still no consensus on this, as the term autonomy is highly controversial, especially in weapon systems. Robotics is currently used primarily in remote-controlled or semi-autonomous applications and in non-lethal areas such as logistics, mine clearance and reconnaissance.

Remote-controlled and semi-autonomous drones are already an integral part of armed forces. However, they have the clear disadvantage of only functioning in areas with active and uninterrupted communication. As soon as communication is disrupted (e.g. by jamming or spoofing), such vehicles become useless. That is why we are seeing an increased development of vehicles that can maintain certain functions (e.g. reconnaissance) independently, i.e. autonomously, after communication has been interrupted and until communication is restored. Autonomously operating drones can be used either on single vehicles, in swarms, or to support manned vehicles (human-machine teaming).

Maintenance and logistics

AI systems can support military maintenance and logistics, as well as assist in planning and provisioning. In the area of maintenance, the integration of sensors and the networking of systems can improve the predictability of maintenance requirements, thereby shortening maintenance times and extending the service life of equipment (predictive maintenance). This can in turn be incorporated into intelligent warehouse solutions (smart warehouses), which can secure and minimise the inventory of accessories and spare parts. Barracks and military properties can also be made increasingly efficient, sustainable, and self-sufficient through intelligent solutions (smart camps).

Improved maintenance and provisioning are incorporated into complex logistics planning, which also benefits from the networking of systems and the integration of AI models. To this end, supply chains can also be optimised and made more resilient. This is necessary due to the rapidly changing conditions in military missions and operations, which require adjustments to supply chains. These adjustments in planning can be greatly accelerated through the integration of AI.

Training and exercise

Another area where AI promises major efficiency gains is training and exercises. AI-supported simulations enable more realistic and flexible training scenarios, with the possibility to digitalise much of the required training for soldiers.

Exercises are an important part of preparing armed forces for emergencies. AI enables commanders to improve the effectiveness of simulations. Soldiers can prepare even better for deployment, for example through the integration of augmented and virtual reality. Furthermore, specialist personnel can conduct adaptive and personalised training to stay up to date in their respective fields of expertise.

Conclusion: AI across National Defence

The potential applications of AI-supported systems extend far beyond the areas mentioned here. Human resources, healthcare, disarmament and many other fields will benefit from the integration of AI-supported systems. In addition, AI enables the increasing networking and integration of different domains and levels. The results of this AI-supported transformation will continue to advance rapidly.

David Song-Pehamberger, BA MAIS, works at the Defence Policy and Strategy Division of the Federal Ministry of Defence with a focus on cyber defence, AI, and emerging and disruptive technologies.

Shutterstock

The legal framework for the use of AI in the military sector

Existing law, including International Humanitarian Law (IHL) in the event of armed conflict, is fully applicable to the use of Artificial Intelligence (AI) in the military and can ensure the responsible use of such systems. A distinction must be made between the use of AI in peacetime and in times of armed conflict, as different provisions apply depending on the situation.

Alexandra Duca

Legal requirements in peacetime

In peacetime, the legal status regarding the use of AI is governed by the national law of each state. This includes, in particular, the fundamental rights applicable in the state concerned. Austria has ratified all international human rights conventions developed by the United Nations (UN) to date and ensures the protection of human rights through both constitutional and ordinary law. The European Convention on Human Rights (ECHR) has constitutional status in Austria. Fundamental rights must also be respected and protected when using AI, unless interference is permitted by law. In the military sphere, Austria's Military Powers Act (MPA) in particular must be taken into account.

The MPA refers exclusively to the activities of military bodies in the field of national defence (Art. 79 para. 1 Federal Constitutional Law). It is conceivable that AI could be used in the context of general mission preparation, the immediate preparation of a mission, and a mission itself, including follow-up measures. In guard duty, AI could be used both in the exercise of legal authorities – e.g. surveillance, (identity) checks of persons, entering property – and in the enforcement of those authorities through, inter alia, the exercise of coercive force. The MPA expressly provides for the use of physical force (including computer systems) for this purpose. Additionally, AI could play a role in military intelligence tasks relating to information gathering and processing. In any case, the legally standardised framework must be observed.

Legal requirements in armed conflict

In the event of an armed conflict, IHL applies as lex specialis. This field of law is then also applicable to the use of AI systems, as the applicability of IHL does not depend on the means and methods used, but solely on the factual existence of an armed conflict. The essential provisions of IHL can be found in the four Geneva Conventions of 1949 and in the two additional protocols to the Geneva Conventions of 1977 (API).

IHL aims to protect persons who are either not or no longer directly participating in hostilities and to limit the effects and consequences of a conflict in general. There are four essential principles of IHL: distinction, necessity, proportionality, and humanity. Any attack that takes place in the context of an armed conflict must therefore be measured against these standards. This naturally also applies when AI applications are used.

This means that a clear distinction must always be drawn between civilian and military spheres. Civilians and civilian objects must never be attacked and must always be protected.

Furthermore, even in war, there is no unlimited right to cause damage; every military measure must be justified. Similarly, it is prohibited to cause superfluous injury or unnecessary suffering, or to use weapons, projectiles or materials, or means or methods of warfare that are likely to cause superfluous injury or unnecessary suffering. Finally, foreseeable collateral damage to civilians must not be disproportionate to the expected concrete and direct military advantage.

These assessments must be made by a human being for each individual case, because situations of armed conflict are highly dynamic and require complex forms of human judgement. This is also confirmed by the fact that IHL is directed exclusively at human beings (e.g. Art. 57(2)(a) API: "Whoever plans or decides [...]"). This does not fundamentally preclude the use of AI as an aid for individual assessments, such as for facial recognition or for estimating the expected collateral damage. However, humans always remain the decision-makers; decision-making cannot be transferred to AI applications. This is also an essential point for the attribution of responsibility.

The question of responsibility

In IHL, the responsibility of commanders as set out in Art. 87 API plays a particularly important role. According to this, military commanders are required to prevent violations of IHL with regard to members of the armed forces under their command and other persons within their command, to stop them if necessary, and to report them to the competent authorities. Commanding officers are also to be held accountable if they authorise an attack without sufficiently verifying the data provided by an AI system (e.g. classification of civilians and combatants by AI) and it subsequently transpires that IHL has been violated.

Regardless of the degree of human-machine interaction with an AI system, human operators of AI systems can also be held responsible alongside the commander, who usually authorises attacks or, in the case of a 'human-on-the-loop' decision, decides that an attack will not be aborted. Similarly, the system's programmers and those who provide data for the AI system on the basis of which it subsequently operates cannot fundamentally evade responsibility. There is a complex organisational structure behind the use of AI, whereby it must be possible to determine where things went wrong in case of an IHL violation.

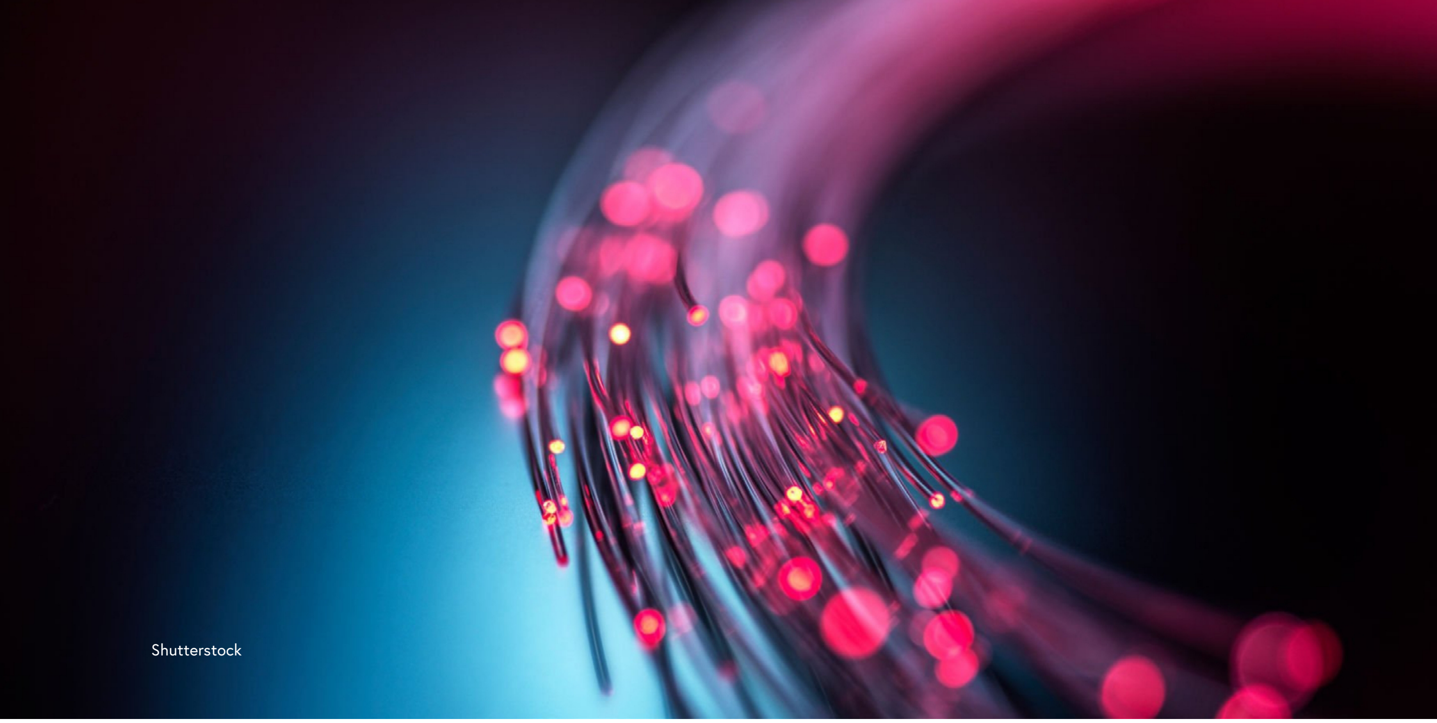
Conclusion

The first step in the legal assessment of the use of AI in the military sphere is to classify the situation from a legal perspective. This classification of the situation will determine the applicable legal provisions thereto.

In times of armed conflict, IHL, as a highly flexible area of law, prescribes a series of rules that often have to be weighed up and assessed on a case-by-case basis.

These assessments must always be made by a human being, although AI applications may be used to assist in this process. However, the use of AI must always remain a conscious decision made by a human being and must not be an autonomous act in itself, in order to ensure, among other things, that responsibility can be attributed in the event of legal violations.

Mag. Alexandra Duca, LL.B works at the International Law Division of the Federal Ministry of Defence with a focus on, inter alia, international legal assessment of new technologies, including AI.


Shutterstock

Artificial Intelligence as a subject of arms control

The struggle to regulate a military technology of the future

The arms control of artificial intelligence (AI) for military use has gained momentum, prompting the Federal Ministry of Defence (MoD) of Austria to take a clear stance.

Michael Retter

Defence ministries and armed forces everywhere expect that the increasing integration of AI will lead to faster and better decisions and a certain degree of automation. The enormous opportunities resulting from the application of AI for military missions are explained in detail elsewhere in this publication. However, in addition to the numerous opportunities, this development also gives rise to potential risks, especially when the benefits of AI relate directly or indirectly to the use of force, including the production and use of weapons. This quickly raises legal, ethical, humanitarian, and security policy questions, such as whether AI systems can independently trigger wars, whether AI-supported weapon systems lead to more civilian casualties, or whether AI applications make it easier for terrorists to gain access to weapons. It is imperative that security policy responses are provided in this instance.

This is where arms control comes into play. Its aim is to contribute to maintaining intergovernmental stability through targeted arms regulations, to prevent humanitarian suffering, and to ensure that actors with sensitive security interests do not gain access to weapons. The results of such arms control processes are usually internationally negotiated treaties.

What often makes the whole endeavour lengthy, arduous, and in some areas even futile, is the fact that states must cooperate on arms control to achieve results. Against the backdrop of the current geopolitical situation, which is characterised by conflict, polarisation, and competition for future technologies, this is no easy task. However, current arms control processes show that this is absolutely necessary.

Between “responsible use” ...

As international regulation of civilian AI applications either explicitly or implicitly excludes the field of ‘security and defence,’ the arms control process has evolved distinctly from it.

The first initiative in this regard was the conference on ‘Responsible AI in the Military Domain’ (‘REAIM Conference’) hosted by the Netherlands in February 2023. As a result of this event, many of the countries present adopted the so-called ‘Call-to-Action.’ In addition, the Netherlands announced the establishment of the ‘Global Commission on Responsible AI in the Military Domain,’ which will present reports and recommendations prepared by experts and thus make a substantive contribution to the international discussions. To the surprise of some participants, the US also presented its own initiative at the REAIM conference, the US Declaration on ‘Responsible Military Use of AI and Autonomy.’

Both the Call-to-Action and the US Declaration are politically binding for the states; they are, however, not international treaties. Both agreements revolve around the central concept of the “responsible use” of AI in the military. This is to be achieved through the establishment of certain, internationally binding norms, including principles, standards, and practical measures. States are thus called upon to take these norms into account when integrating AI into their defence ministries and armed forces.

The merits of both initiatives include the fact that they focus, for the first time, on the potential risks of AI used for military purposes, identify meaningful norms to reduce these risks, and provide a platform for dialogue and exchange. However, the limited participation remains a drawback. Only around 60 countries worldwide support the Call and the Declaration.

... and the regulation of Autonomous Weapon Systems

In an article dealing with the topic of AI in military applications, one issue cannot be overlooked: Autonomous Weapon Systems (AWS). These are systems that, once activated, can select and engage targets without further human intervention.

Experts see AI as the technology that could trigger a quantum leap in the field of AWS. In this area, arms control has adopted a forward-looking approach, with states discussing these systems in the context of the United Nations' Convention on Certain Conventional Weapons (CCW) since 2014, long before the hype surrounding machine learning and generative AI took hold. The main question here is whether the use of AWS in armed conflicts could jeopardise international humanitarian law or ethical principles.

Although the CCW addressed possible arms control regulations for AWS at an early stage, results are still pending. This is due to the complexity of the issue and the strong interests of the states involved. However, there is currently hope that a breakthrough could be achieved by 2026. Since all militarily relevant states are members of the CCW and would therefore be bound by a possible agreement, this would be particularly effective.

In order to advance the arms control process, the United Nations General Assembly adopted a resolution on AWS for the first time in 2023. Austrian diplomatic efforts played a key role in this.. There are currently two United Nations processes addressing the AWS subject: the CCW and the General Assembly.

Position of the MoD

As outlined above, arms control of AI for military use is gaining momentum. The year 2023 marked a turning point in this regard with the emergence of new processes and the differentiation of existing ones.

It is therefore becoming increasingly important for the MoD to take a clear position in this field.

The Ministry's guiding principle is to seize opportunities while minimising risks.

An important part of this ambition is the responsible use of all AI systems by the MoD or the Austrian Armed Forces, in accordance with legal provisions, political guidelines, and ethical principles. Potential risks associated with AI must be identified at an early stage, taken into account in planning, and minimised in a targeted manner. This applies in particular to the use of force. In the spirit of responsible use, this must always take place under human control.

Michael Retter, BA MA, works at the Military Policy Division of the Federal Ministry of Defence with a focus on arms control and new technologies.

Shutterstock

Artificial Intelligence and automation in government

An ethical approach

The potential of automating sovereign acts can only be fully understood if, at the same time, the question of the limits of automation is kept in mind.

Max Gottschlich

The term ‘artificial intelligence’ (AI) is misleading. A machine is not a conscious self that feels, perceives, imagines, judges, and concludes. It is pointless to talk about action, guilt, accountability, and responsibility in relation to machines. Talk of ‘autonomous’ machines is also misleading. Machines are not autonomous in the sense of self-determination, but rather function according to their programming in such a way that they interact with objects in a manner that is useful to us, without the need for human control. AI is a system that relates data in a manner regulated by algorithms. Signals are linked according to a calculated probability – without awareness of their meaning. Nevertheless, the appearance of understanding and logical thinking arises. We have thus created an assistant that organises masses of data through 'automated thinking'. It would be more appropriate to speak of algorithmic systems or assistance systems.

Two types of machines can be distinguished: one that performs its function only under human control, and one that operates without constant intervention – i.e. automaton. The execution of the commands objectified in the program generates the appearance of self-activity. Within the concept of automatons, a further distinction must be made between one, in which automated thinking and the outsourced interaction with the world proceed linearly and rigidly along programmed paths, and one, in which automated thinking in hardware and software enables mobility and interoperability. This gives the automaton the potential to accompany interactions with the world as an assistance system from a utility perspective. An assistance system is a tool that not only serves to relieve physical strain, but has become so powerful that it can support the execution of interactions with the world – in understanding, judging, and reasoning, as well as in deciding and acting.

Within the assistance system, there exists a relevant difference between the terms 'smart' and 'intelligent'. 'Smart' refers to finding and using appropriate means for intended purposes in order to generate and test solutions to problems. AI is also an 'intelligent' assistant that becomes a seemingly independent agent. There are two types of this 'autonomous' assistant: artificial narrow intelligence, which is limited to a specific area, and artificial general intelligence, in which AI connects different areas. This is made possible by an artificial neural network. A processing layer receives activating signals through an input layer. The result of the processing is displayed by the output layer.

Unlike a circuit, the neural network allows for plasticity of functionality in use. On the one hand, according to the rules of the algorithm, weightings of certain connections should emerge in the neural network, enabling the 'recognition' of complex patterns; on the other hand, these weightings must remain plastic in order to be adaptable. This enables machine 'learning'. The appearance of reflexivity depends on this plasticity. Through the use of the algorithm, connections are established in the substrate of the neural structure, which are reduced to the basis of recalibrating change. The usefulness of the process in the plasticity of the data relational process is therefore based on the purposefulness of neural mediation present in nature. The 'intelligence' of AI is based on the principle of imitating nature through technology. Modern bionics takes up the idea, dating back to antiquity, that art perfects nature by imitating it. This concept is a guiding principle in the development of AI.

Benefits

In the name of digitalisation, vast amounts of data are generated in relation to nature and the social world, enabling the real world to be translated into increasingly comprehensive, accurate, and adaptable models. In a model, reality is transformed into a system of clearly definable, concrete relations between phenomena. Once incorporated into the fixed forms of a model, behaviour becomes predictable and thus controllable. The more comprehensive and accurate the models, the greater the need for automatic assistants, without which the use of data masses in the management of technical civilisation would be impossible. Such 'harvesting machines' use algorithmic rules to process data according to utility aspects. This makes it possible to maximise the efficiency of processes.

Algorithmic decision-making systems have been used by banks and insurance companies since the 1980s to process data relating to loans and insurance policies for assessment purposes according to mathematical models. Since the 2000s, these systems include machine learning. This has enabled large amounts of data to be processed in a wide range of areas. The more powerful and complex these decision-making systems are, the more difficult it becomes to trace the paths and results. This has led to the development of supplementary systems that explain the decision-making systems – explainability technologies.

Government administrations deal extensively with rule-based processes, which is why the use of AI is an obvious choice. They primarily use Artificial narrow intelligence in the foreground (chatbots, application submission, data collection, etc.) and in the background (case processing, document organisation and classification, workflow management, predictive modelling of expenditures, etc.). Applications in the legal field make it possible to compare facts with the large number of previously documented cases and decisions. Considering increasingly powerful systems, questions of limits and objective repercussions of this technology arise.

Problems

AI unfolds its functionality on the basis of certain prerequisites that influence the results. This applies to the selection of data, the form of processing in the algorithm, and the mode of 'training' through corrections. Since the public sector is concerned with organising resources for the common good and the self-preservation of the state as a world of freedom, such systems must be used to support administrations under the premise of maximum transparency. Both the administration and politicians must be able to relate to the results of these assistance systems, which can only fulfil their purpose, if blind trust in the results is not required. The problem of the growing dependence of state actors on big tech and its particular interests already touches on a more fundamental level. This is due to the fact that it undermines the trust placed in the state as an institutionalised instrument for the common good.

The following question leads to a deeper problem: To what extent is automation of administrative processes and the shaping of the political community desirable? First, it should be borne in mind that automation also leads to new dependencies and a loss of skills. There can be no upskilling without de-skilling. This must be counteracted in a targeted manner. Automation also has an impact on the self-image of law and the state. The more sovereign action is outsourced to assistance systems, the more the state alienates itself from its purpose. The state is not a machine, not a technocracy in which citizens are managed like data records, but an organic unity of institutions in which citizens must encounter freedom in practice if they are to recognise the state. If, for example, the administration of justice were to be automated and defendants received their sentences from machines, i.e. from objects, this would violate the right of the person to be recognised as a non-object, as a presence of freedom. The limitation of assistance systems is that no algorithm can replace considerations that arise from the knowledge and will of the common good – the good and just for the political community. A judicial decision is not an automatable subsumption of a 'case' under a rule, because justice demands that the respective situation be taken into account in the interests of fairness.

Therefore, there is a trap lurking here: Precisely because the state administration wants to perfect itself through the use of AI, this can become an existential threat. For the state owes its existence to the consciousness of its citizens, who recognise it as the place where they enjoy their freedom. If the population lives in the awareness that it is regulated by a technocracy, it can only see the state as an external force in which it no longer recognises itself. Only those who keep this fundamental problem in mind can avoid the danger that the means, as in Goethe's Sorcerer's Apprentice, will take on a life of their own and defeat the purpose.

Military aspects

The benefit of assistance systems for the military lies in their ability to take into account as many relevant factors as possible by quickly processing large amounts of data from different sources. This is intended to generate a reliable basis for the decisions that need to be made on an ongoing basis in the field, which in turn is intended to secure one's own advantage in the complex dynamics of interaction with the enemy. This is a matter of time, precision, and adaptability of the armed forces at all levels. Analysis tools are intended to model the network of events and make them manageable, right down to the anticipation of future scenarios. In addition, the hope is to reduce dependence on 'human factors' in assessing the situation.

However, now that 'systemic competitors' are also using assistance systems, a dynamic of automation is emerging in military armament, in which – as in AI-supported stock trading – it is also a race against time. The fear of the enemy gaining an advantage is driving the urge to integrate AI into military technology. The arguments in favour of military robotics for reconnaissance and precision strikes are obvious. Minimising the risk to one's own soldiers and reducing collateral damage through machine precision are also cited as ethically relevant arguments for their use.

The problems are obvious: On a technical level, the main issues are the susceptibility of algorithms to errors in distinguishing between friend and foe, and a growing and sensitive dependence of military functionality on the functionality of systems, manufacturers, and necessary resources.

It is inherent in the logic of assistance systems that, due to the way they process data, they implicitly suggest decisions based on utility considerations. The more powerful the systems, the more dependent humans become, and the more binding these recommendations for action appear to be. Automation then indirectly encompasses the decisions themselves.

The growing automation of warfare generates an accelerating and disinhibiting, dehumanising dynamic. All the inhibiting factors in warfare that Clausewitz pointed out are linked to a questioning reflection on the deadly interplay of events. Such inhibition does not apply to machines. This underscores the urgency of the task facing the international community to bring about legal restrictions on the military use of AI in accordance with international humanitarian law.

DDr Max Gottschlich is senior lecturer at the Institute of Practical Philosophy and Ethics of the Catholic Private University Linz.

Shutterstock

AI in geopolitical competition

AI is an emerging and disruptive technology that touches all aspects of society. Control over this area of technology therefore holds enormous potential for economic and military power projection. The resulting geopolitical competition pits two sides against each other: democracies and autocracies.

Daniel Hikes-Wurm

The geopolitical rivalry between the democratic model of the "West" and the power-centred and autocratic model of China, supported by like-minded states such as Russia and Iran, also contains a technological component. This concerns both control over the technology itself and the rules governing its use. The cross-sectoral nature of AI and its inherent dual-use character also increasingly blur the traditional boundaries between state and society.

Control over technology

For years, both China and the US have been competing over dominance in AI, to exert power in both the economy and in national security. Both sides have achieved technological advances that they want to deny the other.

China is a leader in the development of AI patents, both in terms of quantity and quality. Around half of the world's leading AI researchers come from China. The Chinese government is keen to keep this expertise within the country through strict monitoring of research and development and intellectual property. China's liberal approach to data security and privacy also gives Chinese companies more leeway than Western firms when training AI models, especially in areas considered particularly sensitive in the US and Europe, such as surveillance of their populations.

However, China continues to lag behind in hardware. Western-oriented companies work together in global supply chains, including NVidia (USA) for design, ASML (Netherlands) for production equipment, and TSMC (Taiwan) for manufacturing. Due to US-imposed export controls, China is largely excluded from these supply chains. In the field of economically successful applications such as large language models, companies from the US, including OpenAI, Meta, and Google, continue to dominate globally. European AI companies, which are often successful in niche areas, are still missing at the global forefront. With the AI Act, the world's first comprehensive legislation on AI models, the EU is now attempting to establish its own position of power.

Control over norms and standards

The rapid advancement of this cross-societal technology raises far-reaching ethical questions, especially in areas where existing standards and regulations are insufficient. This is another field in which there is an ongoing geopolitical power struggle over what should be allowed and what should be restricted. On the one hand, there is the democratic, value-based system of the West, and on the other, the autocratic and power-centred system of China.

In the West – i.e. the US, Europe, but also like-minded countries such as South Korea, Japan, and Australia – rights to freedom of expression, privacy, and human rights are considered paramount and inviolable. Securing these rights requires strict regulations on the use of personal data and a risk-based approach to AI models, as contained in the EU's AI Act. At the same time, a certain degree of openness and transparency of AI models is required to ensure the protection of democratic values. All countries and companies must adhere to these values, although there are naturally certain differences in interpretation within this Western faction (e.g., between the US and the EU).

This is in stark contrast to the Chinese model, which is designed to strictly follow government requirements. Technological sovereignty is a top priority for China. This goes hand in hand with a very narrow understanding of sovereignty and simultaneous scepticism toward supranational obligations, which China sees as strongly influenced by the West. Data protection is important in China, but only insofar as data is not allowed to leave the country. Domestically, personal data may be used much more extensively than in the US or the EU. This implies a certain degree of reticence surrounding AI, as each country may set and change its rules at short notice. In addition, there is no democratic requirement for AI systems to be transparent to the population – instead, greater accountability to the state itself is demanded.

Society's subtle militarisation

Due to the increasing interconnectedness of society and the cross-border nature of the internet, conflicts are becoming an increasingly common part of everyday life in civil society. This not only affects reporting, which allows conflicts to be broadcast in real time around the world, but also how civil society and the private sector interact with conflicts. Private companies are already an important part of warfare. Prominent examples of this are the involvement of Microsoft and Starlink in Russia's war of aggression against Ukraine, as well as the widespread use of drones from the civilian market by the Chinese company DJI.

Furthermore, cooperation with private-sector technology companies from Silicon Valley is already an integral part of the digital transformation of the armed forces in the US. In the EU, too, the social taboo surrounding civil-military cooperation has now been broken. The billion-euro EU research funding programme Horizon Europe has recently been opened up to dual-use applications. This means that national security and defence are no longer excluded.

The merging of state security with civil society is even more explicit in China, within the framework of its civil-military fusion concept. Here, it is a legal requirement that companies and research institutions of all kinds must share their innovations with the state if they are classified as relevant to national security. This can happen at any time and without prior notice. This also applies to all data stored within China, which must be made available to the state.

This gradual blurring of civil and military boundaries around the world goes far beyond state-controlled aspects. Anyone with internet access can intervene in conflicts. For example, thousands of private individuals from all over Europe have supported Ukraine by fundraising to purchase ammunition and weapons or by participating in extensive cyber campaigns against Russia. This raises the question of whether such private individuals who voluntarily participate in conflicts via the internet can still be considered civilians under existing international humanitarian law. The same applies, of course, to companies that are playing an increasingly active role in conflicts. For geopolitical actors, all this also means that control over strategic companies and influence over civil society are becoming increasingly important.

The geopolitical contest continues

The advantages and disadvantages of strict regulations must be weighed up. This applies to Europe above all else. The EU still sees itself as a standard-setting superpower, as regulations adopted in Brussels (e.g. GDPR) are observed worldwide, but this 'Brussels effect' can be undermined if regulations cause economic damage. It remains to be seen whether this will be the case with the AI Act and such strict regulations.

Although China's approach may appear more efficient at first glance, it has some disadvantages that cannot be dismissed. One reason why Chinese language models lag behind their US counterparts in terms of performance is that Chinese models must adhere strictly to government guidelines and are therefore not allowed to reproduce whatever the data would provide for.

An increasing separation of the two spheres of influence is likely to occur, not only because of differing value systems, but also because of mutual efforts to decouple economies. Whether it is 'decoupling' in the US, 'de-risking' in Europe or 'securitisation' in China, all geopolitical actors currently

Colonel Daniel Hikes-Wurm works at the Directorate-General for Defence Policy in the Federal Ministry of Defence with a focus on hybrid threats and new technologies.

want to protect their economies and societies from excessive influence from others.

The EU, which is seeking to diversify its supply chains in the name of strategic autonomy, is also including trade with the US in its evaluations. Nonetheless, transatlantic cooperation remains necessary to enforce the democratic, Western regulatory approach. The EU-US Trade and Technology Council (TTC) serves this purpose. The US envisages a special security policy role for the TTC, while Brussels still sees it primarily as an instrument of economic cooperation. However, closer security cooperation is essential if the EU wants to have a say in US sanctions and export controls.

China, meanwhile, has been using sanctions and other economic and social pressure for years to achieve national goals. Examples include the rare earth sanctions against Japan (2010), the consumer boycott against South Korea (2016), the import ban on Lithuanian products after Taiwan opened a representative office there (2021), and the sanctions against Australian export industries after the Australian Prime Minister called for an independent investigation into the COVID-19 outbreak in Wuhan (2020). New technologies, including AI and other strategic technologies, may serve as additional leverage for China in the future.

Finally, there remains the question of the West's legitimacy in pushing its views. While the Western approach is taken for granted, most non-Western states are not convinced. There continues to be a struggle for influence over the Global South, many of which question the legitimacy and transferability of the Western approach. China offers many developing countries opportunities that Europe and the US can only counteract to a limited extent. One example of this is the expansion of 5G in Africa, which has been carried out almost exclusively by Chinese companies such as Huawei.

While Western governments warned against cooperation with China, they offered few alternatives. Chinese companies provided comprehensive solutions, including the construction and operation of telecommunications infrastructure, at affordable prices. If comprehensive technological AI solutions are offered at affordable prices, including integrated solutions for better surveillance and control of societies by state authorities, then Western warnings and efforts to achieve a genuinely 'human-centred' approach to AI are likely to be unattractive to many outside the Western hemisphere.

Shutterstock

AI’s role in current conflicts

At the beginning of the 21st century, the conflict patterns we had previously been accustomed to changed fundamentally – from asymmetric warfare to hybrid warfare to high-level conventional warfare. Additionally, we are faced with technological developments whose effects are increasingly far-reaching and whose full extent cannot yet be assessed. Digitalisation, including the development of powerful information technologies and artificial intelligence (AI), has created the conditions for a new revolution in warfare.

Markus Reisner

A decisive step forward in the field of national defence has been achieved through the automation and autonomisation of military reconnaissance, weapons, command and control (C2), and target acquisition systems. The military forces of various countries recognised the opportunities this presented early on and attempted to further develop procedures and tactics. Tests including the 2018 “AlphaDogfight Trials” (virtual air combat between humans and machines conducted by DARPA), the British “Storm-Cloud” trials in 2021 (cross-domain deployment in a transparent battlefield), and real-world deployments such as the US “Maven” project in Afghanistan in 2017 (automatic tracking of human targets) served as both blueprints and previews.

Impact on the current battlefield

Contemporary wars, such as Russia's war of aggression against Ukraine, but also the war in Gaza, are increasingly turning fiction into reality. Two key applications for AI can already be identified: on the one hand, the use of reconnaissance and weapon systems in the various domains of warfare and, on the other hand, support for accelerating military decision-making in C2 and target acquisition processes. The use of unmanned, semi-autonomous, robot-like reconnaissance and weapon systems opens up previously unimagined cross-domain possibilities for the armed forces involved, and the acceleration of processes influences the speed of their own weapons.

In Ukraine, the simultaneous deployment of tens of thousands of drones on both sides is creating a transparent battlefield. It is no longer possible to deploy forces undetected. Every manoeuvre is halted by a hail of kamikaze drones and rapid-fire artillery. At the same time, these drones deliver data that is analysed by AI-based software. Targets are thus quickly identified and engaged immediately. One example of this is the Ukrainian GIS-ARTA software, which collects data on Russian targets for its own artillery batteries.

Other examples show the areas of application that AI is already capable of. In the spring of 2024, a video emerged in Ukraine showing a Russian attack on a Ukrainian infantry base being repelled by ground robots using 'first-person view' (FPV) drones. In addition, the US Army is now supporting the Ukrainian armed forces with algorithms to predict when Ukrainian howitzers will need new barrels. At the same time, Russia and Ukraine are developing software to enable drones to navigate to a target and autonomously home in on it, even if jammers interrupt the connection between pilots and drones. Finally, in connection with the Gaza war, it became known in April 2024 that the Israeli Defence Forces (IDF) were using an AI tool known as "Lavender" to identify suspected Hamas terrorists among thousands of Palestinians.

New framework

These AI-supported developments for military operations have changed the boundaries of space and time. The parameters previously known for military operational thinking, namely force, space, time, and information, are beginning to change. This opens up new opportunities for the armed forces, but also requires them to adapt to changed conditions. Human control over manned and unmanned weapon systems and military software in C2 information systems is exercised via network structures in cyber space or in the electromagnetic field. In the event that an adversary succeeds in controlling and penetrating one's networks, optional attack or defence strategies must be in place.

Due to limited communication, these strategies can only be based on a higher degree of autonomy of software and hardware. The development of semi-autonomous AI programs is therefore being actively pursued, particularly in the cyber domain. One example is the development of a program called "Monster Mind" for the US National Security Agency (NSA). The aim of this program is to detect and neutralise potential cyber attacks on the US at an early stage. Due to the high speed at which such operations are carried out, the goal is to use the program in a completely autonomous mode. There are also considerations to use AI in decision-making in the nuclear field. Such considerations have existed for a long time: during the Cold War, the Soviet Union worked on a so-called "dead hand" concept as part of its "perimeter" system. This was intended to enable a nuclear counterstrike in any case.

Future potential

In the longer term, it can be assumed that, at the end of a corresponding development process, fully autonomous reconnaissance and weapon systems as well as C2 information and target selection systems with minimal AI will be able to independently resolve situations of medium complexity. This includes, for example, unarmed reconnaissance, armed patrols, limited attacks in a defined and designated area, and initial target selection through data analysis.

In the future, semi-autonomous robots equipped with AI will be able to use sensors to independently gather information about their environment. This information will be processed by high-performance processors and form the basis for a decision, which will then be implemented using built-in components (such as movement mechanisms or weapons). With growing experience, the robot will become increasingly capable of optimising itself and

Colonel Mag. Dr Markus Reisner, PhD, is the head of the Institute for Officer Training at the Theresan Military Academy. He focuses on unmanned weapons systems.

becoming more effective. Humans are reduced to a supervisory role.

Current events in war zones make it clear that unmanned reconnaissance and weapon systems as well as AI-supported C2 and target selection systems have become standard in modern warfare. The Pentagon's annual report on Chinese military power recently noted that the People's Liberation Army PLA has begun discussing "multi-domain precision warfare." It concerns the use of "big data" and AI to quickly identify key vulnerabilities in US military systems such as satellites or computer networks. These could then be attacked.

It can therefore be assumed that the transformation of warfare initiated by this will accelerate even further in the future. However, it is only a matter of time before the first drone controlled by terrorist groups or state actors operating in the background targets a football stadium or critical infrastructure with malicious software.

Drones have already proven their worth as effective weapon carriers, whether transporting air-to-ground weapons or carrying explosives. Drones could also be used to deploy chemical or biological weapons. Should such a deployment take place in an AI-controlled swarm, or should malware spread rapidly in cyber space, the consequences of such attacks could be catastrophic.

Shutterstock

The digital transformation of Armed Forces

An international comparison

The global digital transformation is leading to new organisational structures for the implementation of modern technologies, with artificial intelligence (AI) playing an important role in enabling new military capabilities for seamless operational superiority across all domains. This article uses publicly available sources to examine how countries are integrating digital technologies into their armed forces to modernise their defence capabilities and operational efficiency.

Michael Suker

United States

Within the United States Department of Defence (DoD), the Chief Digital and Artificial Intelligence Office (CDAO), established in 2021, and the US Cyber Command (USCYBERCOM) are responsible for digital transformation. The CDAO emerged from the Joint Artificial Intelligence Center (JAIC), which was established in 2018, and plays a central role in developing and implementing strategies aimed at integrating technological innovations and improving efficiency and decision-making within the DoD. Its focus is on data management, AI, automation, and the development of robust and secure networks. This also includes responsibility for implementing initiatives such as Joint All-Domain Command and Control (JADC2), which aim to integrate and optimise command and control (C2) capabilities across all military domains.

In contrast to the DoD AI strategy from 2018 and the data strategy from 2020, the Data, Analytics, and Artificial Intelligence Adoption Strategy, published in 2023, is based on an AI hierarchy of needs. This new focus emphasises the importance of high-quality data and aims to ensure fast and flexible adaptation processes, data analysis, and responsible implementation of AI. Currently, these organisations have different but complementary roles within the US military. While USCYBERCOM is responsible for cyber defence and operations, the CDAO focuses on the strategic integration of digital technologies with an emphasis on AI to improve the efficiency and effectiveness of military and administrative processes. With its personnel, the US also provides the essential infrastructure for NATO's Federated Mission Networking (FMN), which was established based on the experience of the war in Afghanistan to improve cooperation in Joint Operations within NATO.

The US also plays a key role in the Defence Innovation Accelerator for the North Atlantic (DIANA), which was established in 2021 to promote civilian and military innovation and transatlantic cooperation on critical technologies. DIANA supports companies through a network of over 200 accelerators and test centres in currently 28 of the 32 NATO member states.

Germany

The Bundeswehr Centre for Digitalisation and Cyber and Information Domain Capability Development (BwCDig) plays a key role in the digitalisation process of the Bundeswehr (Germany's Armed Forces). As a centre of excellence and driver of digitalisation, the agency is subordinate to the Cyber and Information Domain Command and promises increased efficiency and shorter innovation cycles thanks to the bundling of digitalisation tasks. The BwCDig is responsible for coordinating and developing capabilities in military intelligence, electronic warfare, operational communications, geo-information, and information security.

The main tasks of the BwCDig, which has around 800 employees, include the dimension-specific development and integration of new software products and the adaptation of existing commercial and military software solutions to meet the Bundeswehr's requirements. It is also responsible for harmonising IT systems and improving interoperability within NATO through the FMN. One of its flagship projects is the introduction of a comprehensive digital information and data network, including command and weapon systems, through an integrated battle management system.

France

France published its national strategy for "Artificial Intelligence in Support of Defence" back in 2019, which pursues a long-term and holistic approach to the gradual introduction of AI in all areas of the armed forces. The internal Directorate-General for Digitalisation and Information and Communication Technology (Direction générale du numérique et des systèmes d'information et de communication, DGNUM) is responsible for the strategic direction and implementation of digital transformation across the entire structure of the French Ministry of Defence.

The Defence Digital Agency (Agence du Numérique de Défense, AND) is the central body for digital transformation within DGNUM. It was established in 2021 and has around 400 employees. The AND plays a key role in implementing new technologies and coordinating all major digital projects, and is responsible for modernising the armed forces' digital infrastructure. Its tasks include pooling existing capacities and coordinating resources for the implementation of digital transformation, as well as advising other departments within the Ministry of Defence.

Switzerland

Building on its "Vision 2030," the Swiss Armed Forces' Digital Transformation Strategy aims at strengthening its operational readiness of digital technologies through AI application. Its core element is the standardisation of data, to facilitate integration within the armed forces and into national and international networks.

One priority is the development of the "Sensor-Message-Com-

Michael Suker, BSc MSc is the head of the Cyber Documentation and Research Centre of the National Defence Acydemy. He focuses on identification of fake news and automated information gathering.

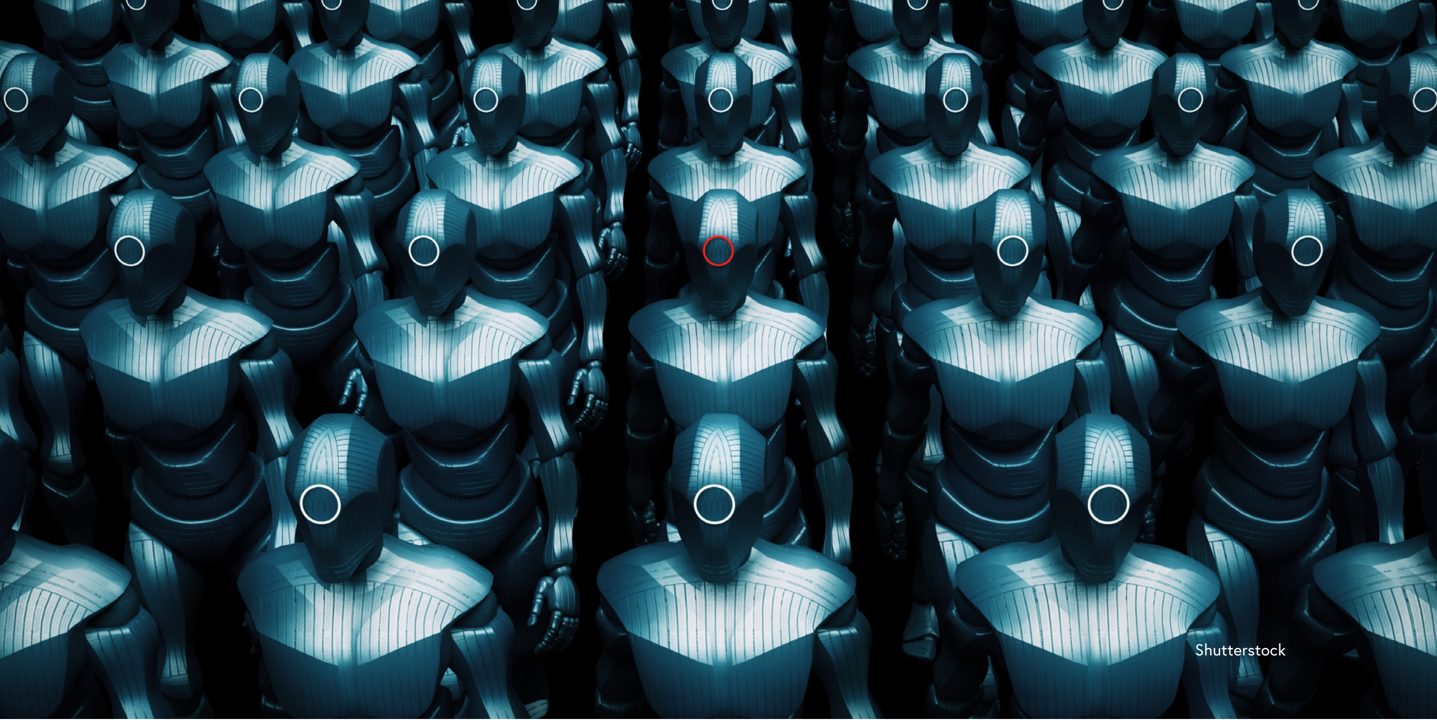
mand-Effect System," which supports and promotes integration in accordance with international standards. The overarching goal of this system is to seamlessly connect various elements and lay the foundations for effective and efficient operational command. To implement the digital transformation, the "Long-Term Development of the Defence and Army Group" Project was established and the Command Introduction Unit was set up as a special implementation organisation, which will be gradually integrated into the Cyber Command. One of the implementation organisation's projects is the administration of a digitisation platform that will serve as the basis for software products.

Summary and outlook

Advancing automation and the use of technical assistance systems combined with high speed, precision, and lethality characterise modern armed conflicts. The potential of AI has led to a paradigm shift in C2 systems, as it significantly increases the speed and efficiency of military operations. When implementing any digital transformation, new technologies must be actively evaluated and implemented in all identified business processes on an ongoing basis. A comparison of international approaches shows that institutions created specifically for this purpose evaluate the development and use of advanced technologies and coordinate their implementation in the armed forces. Close cooperation between (national and international) military and civilian actors, as well as innovation ecosystems that connect military agencies, research institutions, and defence companies to develop new technologies and strategies and increase military efficiency, are also being strongly promoted in the various approaches of other European armed forces, such as the Estonian Armed Forces.

The military concept of "Mosaic Warfare" involves the cross-dimensional use of semi or fully automated, modular, and small – in particular unmanned – systems in combination with other highly automated systems.

Concepts of combat with digital capabilities (unmanned aerial vehicles, Internet-of-Battlefield-Things, smart devices, AI targeting systems, AI-supported tactical planning, etc.) are clearly superior to the traditional approaches of Air-Land Battle and multi-domain operations within the framework of Mosaic Warfare. This requires a fundamental paradigm shift based on modular, functional, and networked platforms.


Shutterstock

Autonomous Systems Technologies and AI

A European perspective

Autonomous Systems Technology is widely regarded as a field in Emerging and Disruptive Technologies (EDT). It is becoming an increasingly important area that has a significant impact on future capabilities in numerous domains. The European Commission's 'Action Plan for Synergies between Civil Industry, Space Infrastructure and the Defence Industry' classifies it as critical technology. Both the EU and NATO consider autonomous systems technology to be a strategic "enabler".

Gerlof de Wilde

The field of autonomous systems technology is developing rapidly. In the defence sector, the European Union is financing numerous research, development, and investment projects through the European Defence Fund (EDF). This was previously done through predecessor programs such as the Preparatory Action on Defence Research (PADR) and the European Defence Industrial Development Program (EDIDP).

Similar projects are being carried out within the intergovernmental framework of the Permanent Structured Cooperation (PESCO). In addition, there are various initiatives and projects within the European Defence Agency (EDA) aimed at integrating autonomous systems into unmanned vehicles for combat missions (CAT-B projects). Furthermore, an action plan for autonomous systems for defence is being developed. The defence industry is also calling for a coordinated EU action plan for the development of autonomous land systems.

Market forecasts predict enormous growth in the field of autonomous systems and the underlying technologies. In the defence sector, the total market volume is expected to more than double from US$41 billion in 2022 to up to US$90 billion in 2030.

Unmanned systems have already been used in numerous scenarios to increase mission endurance, enhance safety and reliability, and reduce risks and costs. Boring, dirty, and dangerous tasks are among the priority areas for such systems. Granting these systems a certain degree of autonomy broadens their range of applications and would maximise their benefits. Thanks to artificial intelligence (AI), autonomous systems are capable of outperforming humans in various areas, such as processing large amounts of data, solving complex problems, and making quick decisions. In general, autonomy is needed or particularly valued when:

- The cadence of decision-making exceeds the limitations of communication channels (e.g. due to delays, limited bandwidth, or communication windows),
- Time-critical decisions must be made by the system or on board the vehicle (e.g. control, health, life-support measures, etc.),
- Decisions can be improved by using large amounts of on-board data compared to transmitted data (e.g. adaptive science),
- Local decisions improve robustness and reduce the complexity of the system architecture,
- Autonomous decisions reduce the cost of a system or improve its performance.

In summary, autonomous systems can contribute to improving activities in all phases of the so-called OODA (observe, orient, decide, and act) loop.

Autonomous systems or autonomous system technologies can be defined as systems that can perform specific tasks in a defined context within a specified period. This enables people in predefined roles to intervene in specific ways or perform control functions. At the same time, the system can independently take over perception, planning, and action under predetermined circumstances. Autonomous system technologies are technologies that also serve as "enablers" for autonomous systems.

An autonomous system therefore requires a certain degree of "intelligence". It needs a model of the world, must have a perception capacity, and be able to work within the framework of a defined task, for example in the form of an auxiliary or target function. Autonomous systems can function as independent, individual units or in the form of a "swarm" (multiple units). An autonomous robotic system (ARS) is an autonomous system applied to a specific hardware platform, such as an unmanned ground, air, or sea vehicle. An ARS faces the additional challenge of navigating the world in which it operates.

Areas for improving capabilities in European defence are prioritised by the EDA Common Defence Policy (CDP). These include ground combat capabilities, improved logistics and medical support capabilities, maritime manoeuvrability, underwater control – which contributes to resilience at sea – and air superiority and mobility. The CDP also contributes to strengthening reactive cyber defence operations, enables the development of space-based capabilities such as information and communication services, integrates military air capabilities in a changing aviation sector, and implements cross-domain capabilities that contribute to the EU's level of ambition.

More specifically, the EDA CDP is intended to support the following aspects: surveillance, detection and identification, combat support, counter-mine and counter-drone measures, search and rescue services, monitoring of chemical, biological, radiological, and nuclear (CBRN) agents as well as decontamination, disposal of explosive artillery ammunition, logistics (convoys) and (medical) care, and target deception or attraction.

In addition, autonomous systems can support combat capabilities by providing direct or indirect fire support, increasing force protection, and supporting intelligence, surveillance, target acquisition, and reconnaissance (ISTAR) capabilities.

Gerlof de Wilde works at the Directorate-General for Defence Industry and Space (DG DEFIS) of the European Commission.

The use of autonomous systems in the military is challenging. Reasons for this include unstructured and difficult environments, legal and ethical issues related to the use of force, a lack of defence-specific data for machine learning, and possible interference by adversaries. Meaningful human control is a legally and ethically relevant element in the use of autonomous systems in all areas. This is particularly salient in the defence sector, where combat and the use of force are part of the system's tasks. In such cases, autonomy receives special attention.

Meaningful human control includes at least the following three elements:

1. Humans make informed, conscious decisions about the use of weapons.
2. Humans are sufficiently informed to ensure that, taking into account the available information about the target, the weapon and the context in which it is used, the use of force complies with the rules of international law.
3. The weapon in question has been designed for a realistic operational scenario and tested in such a scenario. The people involved have received adequate training to ensure that they use the weapon responsibly.

In summary, autonomous system technologies will have a significant and disruptive impact on future defence capabilities. AI technologies will enable decision-making in autonomous systems. For the defence sector, this amounts to meaningful human control and the requirements of international law guiding future development.

Shutterstock

The Austrian Defence AI Strategy

Arnulf Kopeinig

The AI strategy of the Federal Ministry of Defence (MoD) of Austria aims to systematically integrate artificial intelligence (AI) into military and administrative processes. This integration is intended to increase the efficiency and effectiveness of processes and to secure the competitiveness and innovative strength of the Austrian Armed Forces (AAF) in the long term. In the context of ongoing digital transformation, AI is seen as a key technology for meeting the challenges of an increasingly complex security situation and ensuring impact, efficiency gains, and demographic balance.

Strategic objective

The digital transformation requires a strategic realignment of the MoD in its approach to AI. This technology is central for managing complex, data-intensive tasks more effectively and optimising both military and non-military processes. The ever-growing volume of data and the need to make quick and informed decisions make the use of AI indispensable. The comprehensive digitalisation of the armed forces and administrative processes is a key prerequisite for increasing the effectiveness and efficiency of task performance.

The MoD's strategy aims to master the challenges of technological change and digital transformation through the targeted use of AI. A structured and technology-agnostic approach will ensure that the dynamic developments in AI are continuously monitored and adequately integrated into the MoD's processes. This requires close integration of technology developments and organisational adjustments. To this end, a central office will be established within the ministry to implement this strategy.

Guidelines and areas of action

The strategy defines several areas of action to guide the integration of AI.

Governance and organisation: Centralised control and clear responsibilities should ensure that AI applications are properly planned, developed, tested, and operated. A central responsible body will be set up to coordinate and monitor all AI-related measures. Governance establishes the principles according to which AI applications are developed and operated. This also includes ensuring the verifiability of AI systems and their use throughout their entire life cycle.

Performance Areas: The use of AI should help reduce workloads and optimise results. This particularly applies to complex processes such as command and control (C2), strategic planning, logistics, military medicine and cyber operations. AI should also help mitigate demographic challenges. By integrating market-oriented applications, the benefits of civilian technologies can also be utilised for military purposes.

Agility: Organisational flexibility and adaptability are crucial for keeping pace with rapid technology cycles. Structures and processes must be designed in such a way that they can quickly respond to technological changes and are adapted to developments in the maturity of AI. This requires a proactive and initiative-driven approach in order to implement necessary changes in a timely manner.

Personnel: It is necessary to build up and maintain AI expertise in a sustainable manner. This includes recruiting and training experts and promoting diversification among specialist staff. It is essential to position the MoD as an appealing employer, to attract and retain qualified technical specialists. Flexible working models should be promoted to increase the motivation and performance of employees.

Training and Literacy: Ministry staff need to develop a basic understanding of AI to optimise its use. This includes introducing comprehensive training programmes to teach the essential principles and applications of AI, as well as continuously updating the skills of all employees. Modern learning platforms and simulation tools should be used to create practical and interactive learning environments. A particular focus lies on AI Literacy, i.e. the ability to understand and critically question how AI works and how it can be applied.

New types of professions: The integration of AI and new technologies will lead to the emergence of new positions. These new roles will require specialised skills and knowledge that go beyond the traditional training profile.

Trust: Trust in AI systems will be promoted through a combination of cyber security and ethical standards. It is important that the systems are reliable, predictable, and secure. Active communication and transparency in the development and application of AI should strengthen trust in these technologies.

Norms and Standards: AI will be used in accordance with national and international legal provisions, which can be adapted to the latest or upcoming developments in AI. Ethical standards are also of central importance, particularly with regard to human control and the trustworthy use of AI. The accreditation process for complex AI systems will be adapted accordingly.

Defence Policy: The MoD is committed to the responsible use of AI within the framework of relevant norms and standards and to ensuring the protection of soldiers in the field. AI must contribute to improving the capabilities of the AAF and strengthening military national defence. Meaningful human control remains essential in this context.

Foundations for AI integration

AI models are only integrated after a comprehensive assessment of the potential risks and after undergoing rigorous verification, validation, and certification and sufficient auditing by the MoD.

Market orientation is a central aspect of the MoD's AI Strategy. AI projects and marketable solutions can affect all areas of MoD activities. However, it is generally not possible to purchase, install and immediately use a marketable AI application for the purposes of the MoD. Project and innovation management are therefore particularly important, alongside integration and adaptation. AI applications are predominantly dual-use applications. As such, they will be used in both security-related environments and for civilian purposes. In the short term, manifold diverse smaller dual-use applications are to be expected.

The integration of AI requires not only measures for the existing ICT environment, but also an adaptation of the processes and regulations involved, as well as accompanying initiative in terms of personnel and training. Internal guidelines for procurement processes need to be evaluated with regard to accelerating AI projects and adapted if necessary.

The ecosystem of defence, business, and research sectors forms the basis for the successful implementation of the MoD AI Strategy. Cooperation between these areas is crucial for developing innovative solutions and ensuring competitiveness. National partners are preferred for cooperation with institutes from the scientific research and university sectors, research companies, industrial development and manufacturing, and start-ups. Research projects carried out by these bodies are supported by technical expertise from the MoD. This builds up internal expertise and creates a network within the AI ecosystem that enables a rapid approach for in-house projects.

The processes for research, development, and provision at the MoD must be continuously optimised. In addition, active participation in relevant research projects at the EU level and with international partners remains necessary. This will enable synergy effects between national and international work and research on military AI applications. This applies both to the development of defence systems and to the promotion of dual-use products. Due to restrictions in time and resources, independent research by the MoD into the development of applications in the AI technology field is of secondary importance.

The integration of AI products, whether from the market or from research collaborations, should be implemented as quickly and flexibly as possible in every project. To this end, solution providers should be consulted in a flexible manner. Strategic partnerships, both with Austrian and globally active IT companies, should not be ruled out for these purposes.

Key implementation steps

Technology Monitoring and Evaluation: It is necessary to continuously monitor and evaluate technological developments in the field of AI, to identify their potential for military and administrative applications. Systematic monitoring allows trends to be identified at an early stage and for appropriate measures to be taken.

Process Redesign: Existing processes must be redesigned taking the technological possibilities offered by AI into account. This requires close cooperation between the various departments of the MoD. The aim is to adapt processes with the support of AI in order to increase efficiency and improve the quality of results.

Centralised Control and Decentralised Implementation: It is necessary to find a balance between centralised control and decentralised development to ensure the efficient implementation of AI. Clear responsibilities and decision-making powers are crucial in this regard. Decentralised implementation allows for flexible adaptation to specific requirements and circumstances.

Personnel Training: Personnel must receive ongoing training for the use of AI. It is important that employees acquire and develop the necessary skills to use the new technologies effectively. Training programmes and courses should be offered regularly to keep knowledge up to date.

Legal Framework and Ethical Use: Both the legal framework and ethical standards are taken into account in all phases of the development, implementation, and application of AI. This includes the development, procurement, introduction, and use of the systems. Ethical assessments are incorporated into risk management and help to ensure that AI systems are used safely and responsibly.

Cooperation and Networks: National and international cooperation should promote knowledge transfer and utilise synergy effects. Innovative solutions should be developed and implemented through cross-sectoral cooperation with research and industry. The MoD will actively participate in networks and committees to promote the exchange of knowledge and experience.

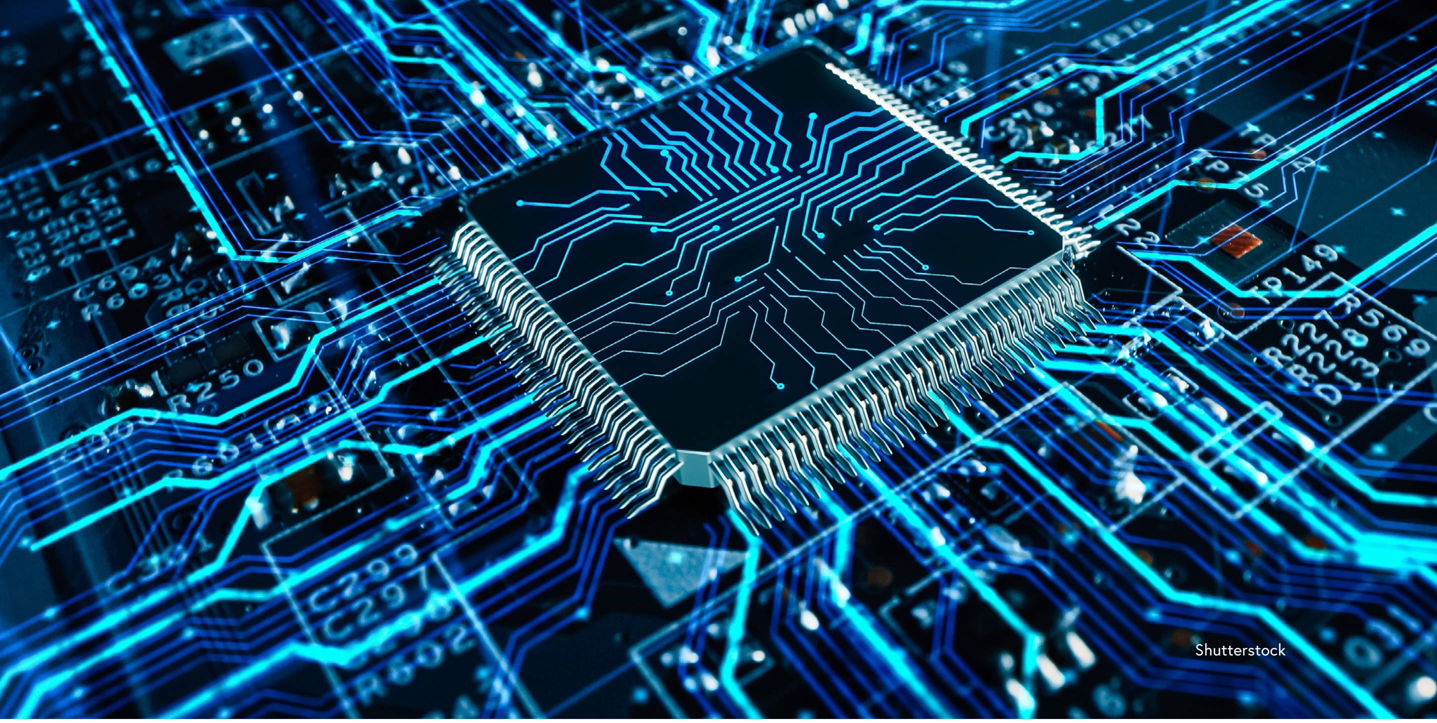
Conclusion

The MoD AI Strategy offers a flexible, comprehensive, and structured approach to the use of AI for military and non-military applications. The systematic integration of AI is intended to increase efficiency and competitiveness. The strategy requires continuous adaptation and further development.

Only through close integration of technology development, organisational adjustments, and qualified employees can the MoD exploit the full potential of AI.

Organisational and human-oriented aspects play a central role in ensuring the successful implementation and sustainable use of AI. Cooperation with industry and research sectors is essential for developing innovative solutions to future challenges. Education and training of personnel and the creation of new types of positions are crucial elements in fully exploiting the potential of AI and shaping the transformation successfully.

Brigadier General Arnulf Kopeinig is the head of the Information and Communication Technology Planning Division of the Federal Ministry of Defence. He is responsible for drafting the Austrian AI Defence Strategy.


Shutterstock

Artificial Intelligence in cyber defence

Even before ChatGPT hit the headlines, artificial intelligence (AI) was opening up fascinating possibilities for many areas of application. As expected, this includes the defence industry, particularly in the field of cyber security and defence, where machine learning and AI are used in various contexts. This brings numerous opportunities, but also challenges and risks.

Victoria Toriser
Florian Silnusek

Opportunities and areas of application

AI is used in many ways at the Federal Ministry of Defence (MoD) of Austria, especially to support the human factor. Since 2023, AI systems have been in productive use in specific areas of cyber security. In the Austrian Armed Forces (AAF), pilot installations help to gain experience and knowledge. AI supports incident and malware analysis by detecting anomalies in the behaviour of networks and devices that indicate attacks. The MoD also uses adapted deep learning models to classify files with a focus on embedded malware or other attack vectors.

In the field of research and development, the EU-level project "Artificial Intelligence Deployable Agent" aims to develop a common European framework to support users and decision-makers in various scenarios. This project, which launched in 2024 and includes Austrian participation, integrates AI-based cyber defence agents that perform autonomous and semi-autonomous actions and cover the entire lifecycle of a cyber incident. The project addresses two key challenges facing end users in the defence sector: on the one hand, a growing attack surface due to increasing digitalisation and, on the other hand, the use of AI-based solutions by attackers.

AI is already widely and diversely used in the field of cyber security, and its importance will continue to grow. It offers numerous advantages:

- In incident and malware analysis, AI helps to detect anomalies in network and device behaviour that indicate an attack. Adapted deep learning models can be used, for example, to classify files with a focus on embedded malware or other attack vectors.
- AI improves threat detection and response by quickly analysing large amounts of data. This enables timely responses to threats, faster detection of cyber-attacks, and thus minimises damage.
- AI helps manage the flood of data, contributing to improved situational awareness and decision-making. Security personnel are quickly provided with utilisable insights that enable better decisions.

AI also plays a special role in Cyber Range technologies. A Cyber Range is a specialised training and simulation environment designed to train and test cyber security skills. It provides a secure, controlled environment where participants can simulate and practice real-world cyber attacks and defence scenarios. This helps security teams improve their skills without putting actual systems at risk.

AI can be used in Cyber Ranges to create realistic and dynamic threat simulations, develop personalised training programs, provide real-time analysis and feedback, and identify and analyse behaviour patterns. It enables automated assessments and reporting, integrates up-to-date threat information, and provides support through virtual assistants or mentors. These applications improve the effectiveness and efficiency of training environments and better prepare participants for real-world threats.

Challenges

In addition to numerous success stories, AI has attained an important place in the public consciousness. However, sometimes AI does not yet meet expectations and, in some cases; limitations of the system are revealed. For example, AI systems have a reputation for being error-prone because the systems are not yet robust and reliable enough to function in every situation. In the field of cyber defence, this means that anomalies or indications of attacks are incorrectly flagged ("false positives"). Furthermore, the lack of transparency and interpretability of many AI systems raises concerns among experts. AI systems are often opaque "black boxes," making it difficult to understand how they make decisions and learn from them.

Security experts also express concerns about systematic errors or biases in AI models. The quality of the training data and the algorithms used are crucial here. If the quality of the training data is poor, this can result in bias, which in turn can have a negative impact on the performance of an AI system.

AI solutions are most effective when they are well integrated into the existing security architecture. Poor or incomplete integration into the existing security infrastructure or incompatibility with other systems limits the usefulness of the AI system.

In light of these challenges, the MoD is working intensively on a national project in the field of cyber security to address the risks and threats posed by the use of AI by external (i.e. other governmental and non-governmental) actors to the AAF's mission-critical information and communication technology systems. The results of the research project are intended to strengthen the AAF's ability to act in the field of AI by providing as comprehensive a picture as possible of the current risk and threat situation. The results are aimed at both technically skilled specialists and decision-makers who, based on the findings, will be able to better assess the current situation and plan and implement further necessary steps in dealing with AI.

Consequence analysis and necessary steps

AI security must go beyond the boundaries of software and hardware implementation and be understood within the triad of technology, application, and people in a cycle. The safe and responsible operation of AI is a key factor for its successful use in various domains. The primary goal is to protect against risks and threat scenarios. A significant part of the discussion focuses on two concepts related to the inclusion of human interaction in automated processes and systems. The “human in the loop” approach, which involves direct human interaction and control over AI systems, ensures that human judgment remains an integral part of the decision-making process.

In the “human on the loop” approach, humans monitor AI operations and intervene when necessary, taking on a supervisory role rather than direct control.

Despite advances in AI, it is crucial that the ultimate decision-making power remains with human decision-makers. AI serves as a support system that provides valuable data and analysis for information purposes, but should not make autonomous decisions.

DI Florian Silnusek ist the head of the Cyber Security Technology Division of the Military Cyber Centre.

Mag. Victoria Toriser is the head of the Cyber Basics and Innovation of the Military Cyber Centre.

Federal Ministry
Republic of Austria
Defence