Technology and Arms Control

CNTRMONITOR

2026

Verification in an Era of Rapid Technological Change



Cluster for Natural and
Technical Science
Arms Control Research

cntrarmscontrol.org

Cluster for Natural and
Technical Science
Arms Control Research

Technology and Arms Control

CNTRMONITOR 2026

Verification in an Era of Rapid Technological Change

Editors

Malte Göttsche, Una Jakob and Niklas Schörnig

PRIF – Peace Research Institute Frankfurt

Peace Research Institute Frankfurt
Leibniz-Institut für
Friedens- und Konfliktforschung

Gefördert durch:

Auswärtiges Amt

aufgrund eines Beschlusses
des Deutschen Bundestages

CNTRMONITOR 2026

Editorial 3

Executive Summary 4

Focus 11

Introduction: Verification in an Era of Rapid Technological Change 12
Malte Göttsche, Una Jakob, Niklas Schörnig

Beyond the Black Box: Conditions for Credible Use of AI in Verification 18
Thomas Reinhold, Fabian Unruh

Verification with Quantum Sensing 24
Markus Gräfe, Yan-Jie Schnellbach

From Images to Insights: Data Integration in Satellite Surveillance 33
Sarah McKee, Niklas Schörnig

Novel Particle Probes in Nuclear Verification 40
Luis Pazos Clemens, Yan-Jie Schnellbach, Adhitya Sekhar

Genetic Engineering Attribution: Contribution to Verification and Practical Limitations 48
Kadri Reis

Trust but Verify: Autonomous Spacecraft and the Verification Challenge 55
Sarah McKee

What Cyberspace Regulation can Learn from Traditional Arms Control 60
Anja-Liisa Krichbaum, Helene Pleil, Christian Reuter

Systems for Implementation Review 65
Owen Greene

Societal Verification 71
Robin E. Möser

Trends 79

Artificial Intelligence 80

Artificial Intelligence and Semiconductors—Key Developments Since 2025
Thomas Reinhold

AI in the Lab: From Biological Design to Biological Access Control 86
Kadri Reis

Artificial Intelligence in Chemistry and Chemical Weapons Control: Fragmentation and the "Execution Gap" 91
Philip Gasse

Scaling Disinformation with AI: A Threat to the Chemical Weapons Prohibition? 95
Philip Gasse

Drones 100

Innovation at Drone Speed—How the War against Ukraine is Shaping Future Warfare
Niklas Schörnig

Biology 104

Mirror Life Moves from Warning to Governance Debate
Kadri Reis

Emerging Disruptive Technologies 108

Technology Preferences of Six Military Innovator Countries: An Updated Analysis of Security Policy Technology Strategies
Abdullah Kasim Korkusuz, Thomas Reinhold, Niklas Schörnig, Liska Suckau

Annex 121

Treaties and Instruments Related to Arms Control, Disarmament, and Non-Proliferation of Nuclear Weapons 122

International Treaties and Instruments Related to the Governance of Biological and Chemical Weapons Disarmament, Non-Proliferation, and Dual-Use Risks 127

The Open Skies Treaty 132

International Initiatives and Core Principles for the Safe Use of (Military) AI: New Developments in 2025/2026 133

Glossary 138

Index 140

Authors 143

Editorial

There is no point in sugar-coating it: the international security situation has further deteriorated dramatically over the past year. However, this by no means implies that we should stop thinking about arms control. Quite the contrary: now more than ever is it essential to strengthen the existing regimes – such as the Non-Proliferation Treaty, the Chemical Weapons Convention and the Biological Weapons Convention – and to prepare for a better future for arms control.

We are delighted to present this third issue of the CNTR Monitor, which focuses on verification. Verification provides a technical basis for collective compliance assessments and renders arms control regimes more credible and more effective. In a broader understanding, we see verification as the production of reliable and internationally available knowledge on activities that could destabilize peace and security regionally or globally. In applying this definition, we do not reject the traditional understanding of verification but rather expand it to include significant developments such as the increasing engagement of civil society in knowledge production.

The Monitor shows how verification efforts must be carefully calibrated. Insufficient verification lowers the hurdles for breaches of an agreement, whilst verification that is too extensive can be overly intrusive and threaten national interests. When planning and compiling the 2026 CNTR Monitor, we aimed to present new verification methods that may alleviate some problems current and future verification faces. Some contributions are explorative technological outlines, e.g. how to make broad use of AI, whilst others present concrete techniques, e.g. new detection capabilities for nuclear verification. Some proposals can significantly improve the accuracy of verification, whereas others make verification more cost-effective.

Beyond the verification focus, as in previous years, the 2026 Monitor also contains chapters on other current technological trends. These include updates on how drones continue to change the battlefield, or broader impacts of AI on international security, such as dual-use risks in chemistry and biology, AI-based disinformation, or the importance of semiconductor supply chains.

We hope you will find this an engaging and informative read.



Prof. Dr. Malte Göttsche



Dr. Una Jakob



Dr. Niklas Schörnig

Introduction: Verification in an Era of Rapid Technological Change

In the current security environment, only few arms-control-related treaties with formal verification regimes remain in place, and geopolitical tensions and power shifts require flexible approaches to old and new security problems. In this world, verification has to be understood in a broader and more inclusive way as it remains ever more crucial to keep producing reliable and internationally available knowledge on activities that could destabilize peace and security regionally or globally. New technologies can help improve these processes: by optimizing the remaining verification mechanisms of existing regimes, by enabling non-governmental actors to participate in OSINT-based knowledge production, and by ensuring that new, non-intrusive verification mechanisms are available when the international situation permits the establishment of new regimes. Artificial intelligence (AI) will play a particular role in this regard, as it has the potential to bring a new level of quality to the verification of arms control treaties in many areas. At the same time, however, the challenges must be considered. For instance, some new technologies may reveal unnecessary confidential or proprietary information and thus cross the line between necessary transparency and undue intrusiveness.

- **Support implementation of new technology in verification.** New technologies can advance current verification mechanisms, especially by increasing accuracy and reducing costs. They should also be considered when negotiating new or updating existing regimes – what was previously difficult to detect may become accurately verifiable.

- **Invest in an ecosystem for the development of verification technologies.** Tailored measures are necessary for a more effective transition of promising emerging technologies from basic research towards implementation. Here, a gap currently exists that should be closed.

- **Use new creative approaches to generate knowledge relevant to security policy.** Opportunities exist beyond the state-led verification methods practiced to date: e.g. less formal processes to gain a better understanding of compliance, or the provision of information by civil society and its integration into compliance assessments.

Beyond the Black Box: Conditions for Credible Use of AI in Verification

Artificial intelligence (AI) is increasingly being integrated into arms control, disarmament, and non-proliferation verification, offering new capabilities for processing large volumes of data, detecting anomalies, and supporting compliance assessments. However, concerns about the "black box" nature of AI have raised questions about the reliability of AI-supported verification. Explainability alone is not the decisive criterion; instead, the key challenge is whether AI systems can be designed and governed in a way that preserves credibility in a limited-trust environment. Reproducibility, integrity, transparency, decentralization, and robust security emerge as the central requirements for the credible use of AI in verification regimes.

- **Focus on credibility and reproducibility of AI.** The black box nature of AI should not prevent its use in arms control verification. For artificial intelligence to become a valuable verification tool, credibility and reproducibility should be treated as the primary design objectives.

- **Agree on standards.** The development of AI-supported verification systems should be guided by agreed standards for data integrity, transparency, reproducibility, and cybersecurity.

- **Enable independent validation.** Participating actors should be able to independently validate AI-generated results through shared datasets, transparent algorithms and procedures, and decentralized technical infrastructures.

Verification with Quantum Sensing

Quantum sensing is an advanced measurement approach that uses quantum mechanics to detect objects over long distances or in noisy environments. It could become relevant to verification where it can strengthen specific tasks, including testing declarations against observations, locating anomalies that warrant follow-up, documenting where and how measurements were taken, and gathering evidence without excessive intrusiveness. Its potential value is practical rather than transformative: more stable field measurements, more selective screening, and more reproducible inspection records. Cold-atom interferometry may support local gravimetry and inertial navigation; NV-center and atomic magnetometers may improve local magnetic-anomaly mapping; and nonlinear interferometers based on induced coherence can probe mid-infrared molecular fingerprints while using visible or near-infra-

red detection. None of these applications has yet been established in treaty inspections. Operational value will depend on a demonstrable advantage over mature instruments, workable measurement times, agreed data rules, and acceptance by the states and organizations involved.

- **Fund mission-driven validation.** Comparative field trials should be carried out, benchmarking performance against established classical instruments at representative sites, and testing false positives, false negatives, calibration stability, ruggedness, operator burden, and data protection requirements.

- **Prioritize bounded pilot cases.** Such pilot cases should include local gravimetry and magnetometry for anomaly mapping, quantum-assisted inertial navigation for GPS-denied inspections, and nonlinear-interferometer mid-infrared screening with visible or near-infrared readout for follow-up chemical analysis.

- **Define the verification question before deployment.** For any new instrument, the performance threshold, evidentiary role, managed-access safeguards, and approval route should be agreed.

From Images to Insights: Data Integration in Satellite Surveillance

Soon after the launch of Sputnik I in 1957, gathering information from space for national security purposes became a major driver of exploration of the "final frontier." Both nuclear superpowers, as well as an increasing number of other states, have invested heavily in satellite-based reconnaissance. However, the most recent developments in the fields of artificial intelligence (AI) and the fusion of data from various space-based observation technologies have the potential to take space surveillance to a new level, including, for example, the space-based detection of stealth assets. In addition, data fusion will enhance societal verification based on commercial satellites and private access to satellite images. Although these advances create new opportunities for verification, they also entail emerging risks. To fully assess the risks and opportunities associated with the use of satellite imagery, it is necessary to understand the current technological limits of the different technologies in use.

- **Invest in AI-enabled intelligence.** Strategic advantage increasingly depends on integrating and analyzing multisensor data rather than collecting imagery alone. Policymakers should prioritize AI, data fusion, and analytical capacity.

- **Maintain rigorous technology assessments.** National spectrum authorities, defense research organizations and the European Defence Agency (EDA) should ensure systematic and independent assessment of novel sensing technologies and emerging uses of the radio spectrum to determine their technical credibility, rate of development, and potential security implications, including the circumstances under which they could become strategically significant.

- **Review remote sensing policy.** Germany should reassess whether its National Data Security Policy for Space-Based Earth Remote Sensing Systems remains fit for purpose in an increasingly international commercial Earth observation market, and whether the current framework continues to effectively address national security objectives while preserving the scientific, commercial, and humanitarian benefits of commercial Earth observation.

Novel Particle Probes in Nuclear Verification

Nuclear energy is experiencing a global resurgence, including the development of novel advanced reactor designs and deployment concepts. Accommodating these innovations would require major modifications to existing nuclear safeguards as well as additional personnel. In light of this, researchers and academic institutions are exploring new nuclear verification techniques, with new non-destructive assay (NDA) approaches looking particularly promising for addressing future safeguards scenarios. These efforts build on advances in industrial technologies and ongoing research and development (R&D) in signal detection and data processing. Analyzing the evolution of frontier verification techniques enables us to identify common technological drivers and the approaches fueling their development.

- **Include nuclear verification in security-focused programs.** Governments should ensure that nuclear verification is included in their security-focused R&D funding programs and schemes, while simultaneously maintaining an emphasis on the civil security aspect, would foster trust through balance and transparency and promote collaborative innovation.

- **Recognize the benefit and impact of adapting basic physics research to verification applications.** Adapting innovations in detection technologies emerging from fundamental physics research has enabled novel tools forming the building blocks of new nuclear verification techniques – this spin-off benefit and impact needs to be more clearly recognized by

academic institutions such as universities involved in fundamental research to support making such research attractive to scientists.

- **Support interdisciplinary research environments.** Sustaining innovation for verification purposes by providing continual support for interdisciplinary maturation environments is likely to be the most effective pathway for leveraging existing R&D efforts; however, while interdisciplinary funding and exchange is available, it often competes, to its detriment, in funding schemes with single-discipline focused research or trends.

Genetic Engineering Attribution: Contribution to Verification and Practical Limitations

Genetic engineering attribution (GEA) remains relevant to ongoing biosecurity and verification discussions. Recent work suggests that engineered biological material may contain detectable traces of design choices and laboratory practices. However, much of this progress relies on curated datasets, such as plasmid repositories, which differ substantially from real-world incidents. Current methods may help with screening and narrowing down possibilities but using them for verification will require better data, more realistic testing, and careful interpretation.

- **Improve data standards for engineered biological sequences.** Biosecurity and attribution initiatives should support more consistent documentation of construct type, vector system, backbone, engineering method, and relevant depositor context in biological repositories and sequence databases.

- **Match real-world conditions in evaluations.** Research programs supporting GEA should prioritize testing under conditions closer to real-world incidents, including partial, degraded, mixed, or poorly documented samples, and distinguish engineering detection from source attribution.

- **Combine GEA with other methods.** Investigators and verification practitioners using GEA should interpret its results alongside epidemiological, environmental, and investigative evidence rather than as standalone proof.

Trust but Verify: Autonomous Spacecraft and the Verification Challenge

Satellite autonomy (the ability to operate and navigate with limited dependence on ground control) is simultaneously beneficial and risky. On one hand, it improves resilience, redefines in-orbit verification, and operational independence. On the other hand, it introduces new governance and security challenges. As fewer decisions can be directly attributable to human operators, it becomes harder to determine whether behavior is accidental, emergent, negligent, or deliberate. In contested orbital regions, autonomous rendezvous and proximity operations (RPO) could heighten mistrust and contribute to escalation between states. A key concern is the risk of miscalculation under conditions of uncertainty, particularly where strategic interpretation is already difficult. Existing governance frameworks and norms have not kept pace with technical capability, limiting their ability to effectively manage emerging autonomous systems.

- **Develop shared verification protocols and minimize ambiguity.** Verification-related maneuvers, trajectories, and approach velocities should be transparent and observable to minimize ambiguity and reduce perceptions of threat.

- **Establish open-ended working group for inspector satellites.** The UN First Committee could task an open-ended working group, in consultation with technical verification bodies such as the International Atomic Energy Agency (IAEA), with developing proposals for a fleet of internationally sanctioned inspector satellites.

- **Specify information barriers.** Space-faring states should establish agreed limits on verification sensors and inspection activities (information barriers) to ensure that orbital verification can be conducted without enabling intrusive surveillance.

What Cyberspace Regulation can Learn from Traditional Arms Control

Cyber operations are increasingly being used for military purposes, particularly in the context of hybrid warfare. However, no binding international regulatory framework for cyberspace currently exists. Drawing on the Chemical Weapons Convention (CWC), often described as one of the most successful arms control regimes, the regulation of cyberspace should focus less on specific technological applications and more on harmful behavior and contexts of use.

- **Focus on behavior and intent.** Cyber arms control should focus on regulating harmful behavior, intended effects, and context of use, rather than specific technological applications.

- **Establish an attribution mechanism for cyber attacks.** An independent but globally recognized attribution mechanism similar to stateless attribution practices proposed in the literature could become a practical alternative to traditional verification approaches in cyberspace.

- **Include soft law and norm building.** Since soft law instruments and norm building are likely to be prerequisites for future binding cyber arms control agreements, they should be included in approaches to regulate cyberspace.

Systems for Implementation Review

Systems for implementation review (SIRs) are international or regional institutions and mechanisms through which parties to a multilateral agreement share information on implementation, compare their implementation activities, review one another's performance, learn from implementation experience, address noncompliance or poor compliance, and, where useful, adjust their commitments. Experience has shown that the development and operation of SIRs can make important contributions to treaty or regime effectiveness, particularly for complex multilateral agreements. Systems for implementation review improve implementation and compliance by helping to mobilize timely international attention and expert responses to emerging problems or opportunities in national implementation of international commitments. They have proved particularly effective in the context of environmental agreements. Although SIRs operate across several multilateral arms control and disarmament regimes, they remain relatively unrecognized and are often inadequately developed. Greater focus on developing and using SIRs could help to improve the dynamism and effectiveness of arms control efforts.

- **Include SIRs in efforts to enhance multilateral arms control and disarmament agreements and soft law.** The German Foreign Ministry and its European Union and international partners should devote greater attention to the development and more purposeful use of SIRs in their efforts to

  - improve the dynamism and effectiveness of existing multilateral arms control and disarmament agreements, including the CWC, BWC, NPT, and humanitarian arms control agreements; and
  - develop international codes of conduct and good-practice guidelines as part of their soft law approaches to promoting international cooperation on safety and security issues in outer space, as well as to limiting the risks posed by the use of cyber weapons and misuse of emerging technologies.

Societal Verification

With the widespread use and increasing accessibility of satellite imagery, societal verification, which assigns analysts outside the government–intelligence nexus a role in producing verification-relevant knowledge, has experienced renewed interest. The monitoring of North Korean nuclear facilities and the tracking of Iranian sites by non-state actors, including nongovernmental organizations (NGOs), think tanks, and independent researchers, illustrates how analysis produced by NGOs feeds into public debate and media coverage. Interviews with open-source intelligence (OSINT) analysts from NGOs and think tanks as well as government officials indicate that societal verification initiatives have two distinct effects: (1) informing the public about pressing security issues, thereby also releasing new information for government actors to use; and (2) stimulating and shaping domestic debates about security policies. Improved dialogue between non-state and governmental experts would help policymakers make better use of information produced by think tanks and NGOs. It could also help alleviate public fears and reduce reliance on worst-case assumptions by addressing the existing gap in reporting on threats and security issues that the intelligence sector and the government only reluctantly fill.

- **Dialogue between non-state actors and governments.** Non-state actors using OSINT and governments should intensify their dialogue to broaden debates and involve more stakeholders in the production of verification-relevant knowledge.

- **The verification-relevant knowledge produced by non-state actors should inform government actors in their decision-making.** Analyses produced by non-state actors provide additional perspectives and context for the public about policy impact and threat potential, at times taking a more critical stance than government releases.

- **Engagement between state and non-state actors should aim at increasing the credibility of future verification approaches and arms control policies.** This could also enhance public support for political measures.

Artificial Intelligence and Semiconductors – Key Developments Since 2025

Semiconductors are no longer merely an economic concern; they have increasingly become a matter of national and international security. Competition over artificial intelligence (AI) capabilities is closely tied to control over semiconductor supply chains, critical raw materials, and the infrastructures required to support advanced computing. Strengthening resilience across semiconductor ecosystems—including manufacturing capacity, energy provision, and international partnerships—has therefore become a key strategic objective for states and firms alike.

- German foreign policy should treat semiconductor supply chains and AI compute infrastructure as core issues of economic security and critical infrastructure protection.
- Germany and the EU should strengthen and diversify international partnerships across the semiconductor value chain, from critical minerals and manufacturing equipment to advanced chips and memory.
- International coordination should improve resilience against geopolitical disruptions by strengthening cooperation on supply security, export controls, and access to critical semiconductor and AI infrastructure.

AI in the Lab: From Biological Design to Biological Access Control

As artificial intelligence (AI) tools become increasingly important in the life sciences, the policy debate is shifting from general dual-use concerns to more practical questions of control, access, and implementation. This includes access to sensitive pathogen data, managed access to higher-risk biological AI tools, stronger nucleic acid screening, and deployment controls at frontier AI firms. The central policy question is how to govern access, oversight, and accountability before riskier capabilities become routine.

- The EU should support stronger and more consistent requirements for nucleic acid screening and customer verification across member states.
- National and EU frameworks should support managed access to higher-risk biological AI tools and sensitive pathogen datasets.
- International coordination should promote compatible standards for nucleic acid screening, biological AI governance, reporting, and evaluation.

Artificial Intelligence in Chemistry and Chemical Weapons Control: Fragmentation and the "Execution Gap"

Although artificial intelligence (AI) is advancing rapidly in chemistry, its application in research and its use in chemical weapons control remain fragmented and underregulated. While many models are being developed, they are rarely integrated into stable workflows. Meanwhile, the widespread availability of open-source and poorly secured tools is facilitating the discovery of novel synthetic routes and toxic molecules. Opportunities for verification, forensic analysis, and mission planning remain largely untapped. To preserve the strength of the Chemical Weapons Convention (CWC), closer alignment between technical capabilities and the policy framework is crucial.

- States and the Organisation for the Prohibition of Chemical Weapons (OPCW) should systematically monitor the application of AI in chemistry to identify and mitigate, at an early stage, the risks posed by fragmented and unregulated models.
- The OPCW and relevant national authorities should develop targeted capabilities for the operational use of AI in verification, analysis, and mission planning to narrow the current "execution gap."
- Developers, research institutions, and policymakers should work together to establish and enforce binding safeguards and access restrictions for particularly sensitive applications, such as toxic molecules, retrosynthetic planners and autonomous laboratories.

Scaling Disinformation with AI: A Threat for the Chemical Weapons Prohibition?

Agent-based, large language model-driven artificial intelligence (AI) can make disinformation more scalable, cheaper, and harder to detect. This undermines trust in institutions and democratic processes. This development has a particularly significant impact on chemical weapons control. The Organisation for the Prohibition of Chemical Weapons (OPCW) has long been the target of coordinated campaigns seeking to weaken accountability, institutional legitimacy, and governance structures. Traditional methods of detecting disinformation campaigns using bots are ineffective against new technical developments. Therefore, what is needed are resilient institutions capable of detecting attacks early, communicating transparently, and developing counter-

strategies cooperatively, rather than relying on technical censorship measures.

- The OPCW should establish a monitoring and analysis framework to systematically track, assess, and address AI-enabled disinformation campaigns aimed at attacking its credibility.
- The organization and its member states should continue to promote proactive and transparent communication strategies that clearly explain the OPCW's procedures, mandates, and limitations. This will strengthen resilience against delegitimizing narratives.
- States, international organizations, platform operators, and AI providers should work together to develop methods for detecting AI-based "swarm" campaigns and implement coordinated measures to flag or mitigate such activities, particularly during sensitive events.

Innovation at Drone Speed—How the War against Ukraine is Shaping Future Warfare

Over the last two years, we have reported on the rapid development of drones in the war in Ukraine. This trend still holds true and has even developed more momentum. Major developments include, amongst others, the rising importance of counter-drones hunting other drones, the significant increase in uncrewed ground vehicles (ground drones) and the fact, that Ukraine has become a world leader in drone and counter-drone technology, changing from an importer of security to an international provider.

- Given the poor prospects for the international regulation of drone warfare and the dynamic trends of technological development, special attention must continue to be paid to drone defense, both in the military and the civilian spheres.

Mirror Life Moves from Warning to Governance Debate

Although no self-replicating mirror organism is known to exist, mirror life has become an active governance concern. Since last year's CNTR Monitor, UNESCO, the UN Scientific Advisory Board, UK officials, funders, and the Mirror Biology Dialogues Fund have advanced discussions on moratoria, precursor technologies, and safeguards. Recent synthetic-cell progress and debate about ecological constraints reinforce the need for targeted rules that prevent self-replicating mirror organisms while preserving lower-risk work on non-replicating mirror molecules.

- Governments and international scientific bodies should maintain a strong norm against creating self-replicating mirror organisms.
- Policymakers, funders, and research institutions should develop governance for precursor technologies that could make mirror organisms easier to build.
- Lower-risk work on useful non-replicating mirror molecules should continue under rules that clearly distinguish it from work toward self-replicating mirror organisms.

Technology Preferences of Six Military Innovator Countries: An Updated Analysis of Security Policy Technology Strategies

The nature of warfare is constantly and rapidly changing – and new technologies play a significant role in this. But which technologies have been, and continue to be, regarded by states as particularly significant in military terms? An analysis of strategy papers from six selected states between 2016 and 2025 shows that the further development of traditional conventional weapon systems is important alongside the development of new military technologies such as improved command and control systems, novel sensors, AI-based analysis and networking. In addition, there has been a growing interest in (ballistic) missiles, particularly in recent years. However, the analysis also reveals that discussions on future military technology vary greatly across the states analyzed. Whilst, for example, Australia engages in high-quality discussions both in terms of quantity and depth, Germany lags behind both in terms of the number of documents and the depth of their content.

- The increased interest in ballistic missiles is a cause for concern and should be discussed at an international level, initiated by the EU countries.
- An open discourse on military and military-technical orientation is appropriate in democracies and also serves to secure legitimacy. Consequently, those countries that have so far shied away from entering into such a discourse, including Germany's Ministry of Defense, should intensify it.

FOCUS

Verification

Introduction: Verification in an Era of Rapid Technological Change

We live in an era of extraordinarily rapid technological change which impacts virtually all aspects of social life, including international security. On the one hand, some technological advances, such as drones, AI-assisted systems, significantly enhanced sensors, and new materials, are finding their way into military arsenals, significantly altering the conduct of warfare. Technological developments in areas like space, biology, biotechnology and chemistry may bring immense benefits, but some could also be misused for nefarious purposes, requiring thorough dual-use risk assessments. On the other hand, science and technology can also be harnessed to strengthen international peace and security, particularly through their use in arms control verification. Thus, there is an urgent need to critically examine how the rapid scientific and technological advances could reshape the verification landscape of the future.

In this CNTR Monitor, we adapt and broaden our definition of verification to reflect a changing security environment in which few arms control treaties with formal verification regimes remain in place and in which geopolitical tensions and shifts in the balance of power call for flexible approaches to both longstanding and emerging security problems. Significantly, in this environment, it is more important than ever to continue producing reliable and internationally accessible knowledge about activities that could undermine peace and security at the regional or global level. This in itself can act as a deterrent or means of dissuasion, establish accountability, help increase confidence in compliance with treaty obligations, detect violations of agreements, and prevent information gaps and reliance on worst-case assumptions in crucial areas. It is precisely this broad understanding of internationally accessible stability- and security-relevant knowledge production aimed at increasing transparency to which we refer in using the term verification. In this perspective, verification can be based on treaties, but does not exclusively rely on them, aims to establish whether states are acting responsibly, and could even include new actors, such as civil society, in such knowledge production.

Such a broad understanding of verification reveals important dynamics and points of friction that are likely to become increasingly relevant in the future. For instance, the balance between transparency and intrusiveness, long a matter of contention, may become even more consequential with advances in fields such as satellite imagery and AI-supported analysis. The relevance and function of negotiated restrictions of

data-gathering in formal verification protocols may shift in this area as public access to knowledge becomes increasingly fine-grained. The 2026 CNTR Monitor seeks to illuminate the interplay between these and other technological and political dynamics and offer recommendations how to react to them.

New Technologies for Verification Purposes

Many technologies are currently seeing rapid improvements that could—and in some cases already do—support verification. The most relevant elements include the improved safety of inspectors, data collection and data analysis, attribution methods, and cost reduction.

One area that warrants particular attention in this context is sensor technology. New sensors are becoming increasingly smaller while simultaneously offering higher resolution and greater robustness. They are easier to handle and more durable, for example because they operate at lower voltages. This reduces costs, which is especially important for nuclear safeguards. Particularly in light of the anticipated introduction of small modular reactors and other new reactor technologies, new sensor technologies, especially when deployed remotely, could help the International Atomic Energy Agency (IAEA) maintain a sufficiently dense yet affordable verification network.

See the chapter by Yan-Jie Schnellbach, Luis Pazos Clemens, and Adhitya Sekhar on novel particle probes, p. 40 ff.

Some of these advantages also hold true for future quantum sensing. Although quantum sensing has not (yet) produced any dramatic improvements in terms of resolution or robustness, it could significantly reduce the costs of analysis in some applications, help detect anomalies that trigger verification or investigation activities, and supplement—or, in the longer term, even replace—existing verification methods, for example, in identifying chemical substances or enabling inertial navigation in remote locations.

See the chapter by Markus Gräfe and Yan-Jie Schnellbach on quantum sensors, p. 24 ff.

Advances in satellite imagery also offer important benefits for verification. The resolution of optical, infrared, and radar images is continually improving, although, unlike in other fields, the associated cost trends are hard to predict. While the cost of delivering payloads to orbit per kilogram is falling, higher resolution depends on the surface area of the mirrors used in telescopes, and larger, flawless mirrors are—ceteris paribus—more difficult to manufacture and transport. By fusing satellite images from multiple sources using AI, even when imagery is sub-state-of-the-art quality, it is possible to reveal features that would otherwise remain undetectable even in the highest-quality single-source imagery available. While there is no evidence that AI-enabled data fusion from other sources produces effects comparable with those achieved with satellite data, improvements in data analysis can generally be expected.

See the chapter by Sarah McKee and Niklas Schörnig on data integration in satellite surveillance, p. 33 ff.

Yet, the use of AI in verification mechanisms remains a sensitive issue. The IAEA, for example, already employs machine learning in its analysis of large datasets and satellite imagery, while the Organisation for the Prohibition of Chemical Weapons (OPCW) has recently completed a review of the role of AI in its own work.[1] However, while AI has the potential to support verification, its use also carries risks of misinterpretation and "hallucinations" and raises second-order concerns about trust in AI, particularly with regard to the quality of training data, the potential for manipulation, and the explainability of AI-supported findings.[2] Measures can and should be taken to address these problems and enhance the credibility of AI-supported activities.

See the chapter by Thomas Reinhold and Fabian Unruh on the credible use of AI in verification, p. 18 ff.

In the biological field, genetic engineering attribution has been recognized as one possible means of supporting and complementing verification and attribution by providing information on the origins of genetic modifications. This in turn could, for example, help trace genetically modified pathogens back to their source, which is important in the event of a deliberate biological incident or a disease outbreak of unknown or am-

See the chapter by Kadri Reis on genetic engineering attribution, p. 48 ff.

biguous origin. Assuming that the available data are expanded, streamlined, refined, and adapted to real-world scenarios, genetic engineering attribution might support and complement other verification and attribution methods.

In summary, as the contributions to this Monitor show, many technological developments have the potential to support, enhance, and transform verification. While some may offer only incremental improvements or cost reductions, others—particularly those involving the use of AI to extract information from multiple or diverse data sources—could enable significant advances in generating new insights. As also discussed in this Monitor, seizing these opportunities must go hand in hand with recognizing and addressing potential challenges in applying new technologies to verification, exploring their role beyond traditional settings, and creating an ecosystem in which verification research can thrive.

The Fine Line Between Verification and Spying

In the traditional, narrow definition, verification is typically understood as the essential mechanism for increasing confidence that the parties to a treaty are complying with their obligations and enhancing the likelihood that noncompliant behavior will be detected. The high stakes involved in international security mean that without this provision, there would be a greater risk that a member state might secretly violate the agreement for its own advantage. However, the term "verification" should not be understood as implying a degree of absolute certainty that even the strictest surveillance regime cannot realistically provide. What verification can do, however, is provide a technical basis for assessing compliance, enhance trust, improve the effectiveness of agreements, and increase the political and material costs for noncompliant behavior. In principle, all the technologies described in this Monitor can or could be used to support and enhance verification and reduce uncertainty about the behavior of the parties to an agreement.

However, a number of further aspects require consideration, as overly intrusive verification could ultimately cause more harm than good. States therefore need to carefully balance effectiveness, intrusiveness, and national interests when designing verification mechanisms, a task that requires both technical expertise and political judgment. Importantly, verification generally does not aim to detect small or minor, often unintended, violations, but rather focuses on those actions considered militarily significant. If enhanced verification technologies provide increasingly detailed information on states' behavior and lead to the more frequent identification of smaller violations, including those resulting from inadequate national implementation of obligations, additional or different procedures may be required to address such issues collectively and cooperatively in order to avoid unnecessary tensions within a regime.

Moreover, the line between permissible verification and intended or unintended industry or military espionage might easily be crossed—either when highly intrusive verification methods collect additional information as a kind of "bycatch," or when such information jeopardizes core aspects of a verified party's security or proprietary information while providing only marginal benefits to the other parties. Even the perception of this risk could affect the level of acceptance of and willingness to participate in verification activities.

This situation is further complicated by the fact that some new technologies capable of enhancing verification can also be used unilaterally to gather information, potentially raising concerns about espionage or intrusive surveillance and reconnaissance. In the space domain, for example, autonomous approach functions provide the possi-

bility to obtain extensive information about "inspected" satellites, and includes a potential for sabotage. Recent advances in data fusion may enable states and nongovernmental actors alike to derive information from commercially available imagery that could potentially compromise national security. In the past, verification by national technical means (NTM)—meaning satellite surveillance of nuclear bases—was an important component of verification under the major, now-expired nuclear arms control treaties between the Soviet Union and the United States. These included obligations not to obstruct the other sides' activities. However, no regime currently exists that explicitly encompasses and regulates such unilateral information gathering. Depending on how they are used, technological advances in verification could therefore contribute to increased surveillance and, consequently, heightened tensions.

See the chapter by Sarah McKee on autonomous rendezvous and proximity operations, p. 55 ff.

See the chapter by Sarah McKee and Niklas Schörnig on data integration in satellite surveillance, p. 33 ff.

Other Approaches to Verification

Based on a definition of verification as a process of knowledge production, some contributions to this year's Monitor broaden the focus beyond traditional technology- and state-based approaches to encompass other aspects and forms of verification.

For instance, improving national implementation of existing regulations is regarded as one avenue complementary to verification. If adapted for arms control, systems for implementation review (SIRs) could be applied in several arms control fields covered in this Monitor. Rather than monitoring activities and installations, SIRs in arms control could provide forums for collectively reviewing, discussing, and improving national implementation measures, thereby contributing to the effectiveness of arms control regimes and regulations in a different way.

See the chapter by Owen Greene on SIRs, p.65 ff.

Specific features of outer space and cyberspace make verification particularly challenging. Both are characterized by a high degree of dual use, which creates distinct challenges in each domain. In outer space, the growing autonomy of spacecraft has consequences for verification, both by making it increasingly difficult to determine the nature of activities and by requiring verification activities to remain nonintrusive and retain their character as a form of collective knowledge production rather than unilateral intelligence gathering.

See the chapter by Sarah McKee on autonomous space craft and verification challenges, p. 55 ff.

While verification is thus possible in and from outer space in principle, the situation is more complex in cyberspace. Here, the strong dual-use nature of the technology and its ubiquitous application make classical verification highly challenging, though not entirely impossible.[3] Drawing on experience with the Chemical Weapons Convention, regulation could be based on intent rather than capabilities and place greater emphasis on attribution than on verification in the traditional sense.

See the chapter by Anja-Lisa Gonsior, Helene Pleil, and Christian Reuter on cyberspace regulation and traditional arms control, p. 60 ff.

Lastly, the role in verification of nongovernmental actors and experts outside government and intelligence services also deserves consideration. Bolstered by technological advances and the resulting greater accessibility and improved quality of publicly available data, including commercial satellite imagery, open-source intelligence (OSINT) and investigations by nongovernmental actors can contribute to providing additional information on verification-relevant issues and informing public debates and political decision-making, notwithstanding some problems and potential frictions in their application. As diverse as these examples are, they demonstrate that discussions of verification in the new security environment need to move beyond traditional conceptions and consider different approaches in the wider context of arms control verification.

See the chapter by Robin Möser on societal verification, p. 71 ff.

Designing an Ecosystem to Develop Verification Technologies

Most of the advances that promise to significantly enhance verification capabilities originate in basic research. Research in solid-state and atomic physics, optics, and information theory has driven the second quantum revolution, opening up new possibilities for its application to verification. Advances in radiation detection developed for fundamental particle physics have generated new capabilities that can be exploited for nuclear safeguards. Computer science is developing new AI systems and architectures that can be applied to a wide range of verification-relevant tasks. The capabilities of microbial forensics rely on progress in microbiology and genomics. In other words, fundamental advances that later shape verification technologies are typically developed without arms control in mind.

However, adapting and applying the results of frontier science to verification technologies that are ripe for implementation is far from straightforward. Every chapter of this year's CNTR Monitor Focus provides insights into the intricacies of verification at the intersection between science and politics, their mutual influence, and the ways in which technologies can be used for verification purposes. The distinct security aspects and specific characteristics of the weapons in question must be taken into account. The capabilities and limitations of such systems must be scientifically assessed by scholars with expertise in verification regimes and in the specific items, facilities, and general characteristics of evidence and its environments. Ultimately, these systems must be tested under realistic conditions.

It is in these areas that the authors of the CNTR Monitor identify the most significant obstacles to harnessing the best available science for the most effective verification. The development requirements for verification techniques are highly specific and target a small user base. Sustained financial support is needed to ensure that researchers and producers can continue to undertake these niche tasks. One example of a functional development infrastructure is the Member State Support Programme coordinated by the IAEA in the field of nuclear non-proliferation.[4] For most other arms control or disarmament domains, such ecosystems are rare or entirely absent.

First and foremost, bringing frontier developments into verification practice requires sustained and structured support for interdisciplinary research environments at universities and other institutions. Here, the applicability of frontier research to the development of novel verification techniques can be assessed, and research aimed at developing such techniques undertaken. This should be followed by development activities to advance verification technologies toward higher levels of technological readiness. Lastly and equally importantly, mission-driven validation in realistic settings, such as field trials, is crucial. Such trials could initially be carried out as standalone tests, but should ultimately be integrated into exercises that combine the application of verification methods with the political processes surrounding verification. Here, collaboration between governmental and nongovernmental actors from academia, applied research, and industry could contribute to the design of realistic scenarios while enhancing the technical feasibility and political acceptance of the verification measures in question.

Conclusion

Dedicating an entire volume to verification in the current geopolitical climate may, at first glance, seem counterintuitive. After all, many arms control agreements containing verification provisions have either expired or become dysfunctional, while traditional verification approaches often seem ill-equipped to deal with new political

realities and technological developments. Yet, some verification systems—such as those maintained by the IAEA or the OPCW—remain in place and may be more necessary than ever. These systems could benefit greatly from harnessing technological advances and coupling them with existing verification methods, where feasible and appropriate.

Despite these challenges, there are still developments that give cause for hope. In the biological field, where no verification system is currently in place but—contrary to the broader trend in arms control—political agreement exists to at least discuss verification options, technological assessments of potential verification methods represent important contributions to both the political and technical debate.

Although such debate remains the exception, and despite the absence of any signs that the security situation in Europe and worldwide is easing, it is important not to lose sight of the possibilities for verification. After all, if the broader political climate improves, solutions should be ready for the arms control challenges that will inevitably arise. Monitoring technological developments and exploring how they might be harnessed for arms control, disarmament, and non-proliferation purposes is therefore essential. This Monitor offers ideas for putting new technologies to constructive use in these areas.

Malte Göttsche

PRIF – Peace Research Institute Frankfurt/
Technical University of Darmstadt

Niklas Schörnig

PRIF – Peace Research Institute Frankfurt

Una Jakob

PRIF – Peace Research Institute Frankfurt

1 Organisation for the Prohibition of Chemical Weapons. (2026). *Artificial Intelligence. Report of the Scientific Advisory Board's Temporary Working Group on Artificial Intelligence* (SAB/REP/1/26). https://www.opcw.org/sites/default/files/documents/2026/03/Final%20Report%20of%20the%20SAB%27s%20TWG%20on%20AI%20FINAL%20VERSION.pdf

2 Schörnig, N. (2022). Artificial Intelligence as an Arms Control Tool: Opportunities and Challenges. In T. Reinhold, & N. Schörnig (Eds.), *Armament, Arms Control and Artificial Intelligence* (pp. 57-72). Springer. https://doi.org/10.1007/978-3-031-11043-6_5

3 See Reinhold, T. (2022). Arms Control for Artificial Intelligence. In T. Reinhold, & N. Schörnig (Eds.), *Armament, Arms Control and Artificial Intelligence* (pp. 211-226). Springer. https://doi.org/10.1007/978-3-031-11043-6_15

4 Schäfer, J. (2026). Technology diffusion in IAEA safeguards: infrastructuring through collaborative effort. *Zeitschrift für Friedens- und Konfliktforschung*. https://doi.org/10.1007/s42597-026-00174-9

Artificial Intelligence

Beyond the Black Box: Conditions for Credible Use of Artificial Intelligence in Verification

This chapter examines whether the often-cited "black box" character of AI constitutes a fundamental obstacle to its use in arms control verification or whether—and under what conditions—AI can become a credible verification tool. It argues that verification has always involved balancing transparency against the protection of sensitive information, and that the central challenge is not explainability itself, but credibility in a limited-trust environment. The chapter identifies the key technical and institutional requirements for credible AI use, including data protection, integrity, reproducibility, decentralization, transparency, and secure infrastructure. It further examines the governance trade-offs that arise when these requirements are negotiated and concludes that AI can strengthen verification regimes when embedded in robust institutional arrangements and accepted by participating actors.

Across almost all technical and administrative domains, there is a clear trend toward the increasing use of artificial intelligence (AI). Data volumes are growing, sensors are proliferating, and political decision-making is increasingly expected to be informed by continuous monitoring and rapid analysis. Arms control, disarmament, and non-proliferation verification are not exempt from this development. Not every verification procedure requires AI support, but it can help address some of the challenges that verification regimes encounter and can serve as an enabling factor for the deployment of modern technologies. Some of these challenges include an ever-growing number of facilities needing to be safeguarded by a limited human workforce, the increasing availability of sensors for verification purposes, and the accelerating pace of geopolitical developments requiring fast and reliable compliance assessments. Here, AI is particularly attractive as it promises to support exactly those tasks that are resource intensive for humans: screening large amounts of sensor data, detecting anomalies, aggregating heterogeneous information streams, and accelerating otherwise lengthy verification procedures to make verification more efficient, responsive, and potentially more objective.[1]

But verification is a highly political endeavor. Introducing AI into its processes raises questions about the explainability of the data processing, the credibility of its outcomes, the development and negotiation of the underlying algorithms, the selection and balance of the data used to train the AI, and even potentially unequal access of different participating actors to advanced models or computing infrastructure—all of which will be discussed in the following sections.

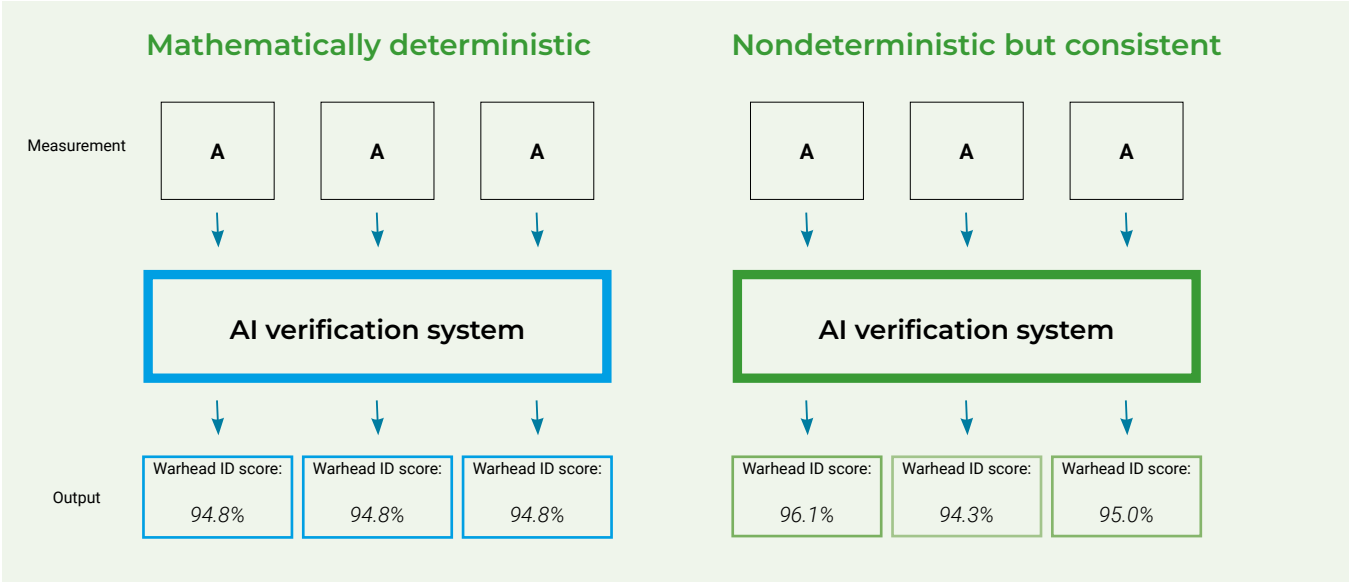
The Black Box Problem Reconsidered

The opacity of AI systems, in particular, seems to contradict the prerequisites for effective verification mechanisms. This so-called "black box" character of AI[2] has two closely interconnected dimensions: limited transparency and traceability of the system's internal workings and, consequently, limited explainability of the outcomes it produces. This is rooted in the logic of machine learning methods, which are composed of large numbers of interconnected mathematical operations. In practice, users may receive a classification or recommendation without being able to directly identify which specific characteristics of the input data, or which effects of the training data, led to the result. While the inner workings of simple neural networks for classification and regression tasks can be explored reasonably well with explainable AI (XAI) techniques,[3] this process is much more intricate for large language models (LLMs) and generative AI systems.[4] These AI techniques enable the processing of complex textual information corpora and are therefore well suited to verification tasks for which human analysis was previously considered indispensable. However, the algorithms used are stochastic in nature, meaning that data processing is not mathematically deterministic and identical inputs may not necessarily produce identical outputs. These characteristics are not merely technical inconveniences. They affect the core function of verification, which depends on processes capable of producing results that are considered reliable and credible. Artificial intelligence thus enters the verification debate not only as a promising tool, but also as a potential new source of contestation, with actors seeking to exploit its limitations to question verification findings or even justify noncompliance by challenging the reliability of the verification system itself.

The "black box" nature of AI stems from the limited transparency and traceability of its internal processes, as well as the limited explainability of its outcomes.

The stochastic nature of AI techniques presents a challenge because of the reliability and credibility expected from verification mechanisms.

Upon closer inspection, however, the picture is more nuanced. Verification usually involves balancing transparency against the protection of sensitive information. The more intrusive or sensitive the verification provisions, the more difficult it becomes to strike this balance. For example, national technical means were accepted earlier in nuclear arms control than on-site inspections involving physical measurements. States seek to avoid disclosing more than is strictly necessary for verification purposes, while still providing sufficient evidence to establish compliance. From this perspective, opacity is not entirely new. Even traditional verification technologies are often complex and not fully transparent to all parties. Consider nuclear warhead verification. High-resolution spectra of emitted particles could confirm the presence of a warhead and rule out a hoax object. However, these spectra would also reveal



sensitive design information, rendering their use politically unacceptable. While some approaches try to avoid gathering any sensitive information at all,[5] others apply what are referred to as "information barriers",[6] which intentionally reduce the resolution of the measured data while still preserving enough information for the verification task.

Verification usually involves balancing transparency with the protection of sensitive information, which means that perfect transparency is not required.

The mere fact that an AI system is not perfectly transparent therefore does not automatically disqualify it. Likewise, determinism in data analysis should not be overestimated. Deterministic systems can also systematically produce misleading results if they rely on inappropriate assumptions or biased data. In addition, traditional data analysis can also suffer from nondeterminism due to statistical fluctuations. Provided that systems with probabilistic elements can be adequately characterized, repeatedly tested, and validated to ensure consistent output behavior, nondeterministic AI is viable for verification.

For verification purposes, credibility is the politically relevant category.

For verification purposes, the politically relevant category is therefore not determinism but credibility, ensuring that participating actors perceive the system as producing results that cannot be easily dismissed as arbitrary, manipulated, or one-sided.

Conditions for Credible Use of AI in Verification

Given this shift away from the necessity of an explainable AI system, the question is which technical and institutional requirements AI systems used in verification must satisfy for parties to regard the resulting assessments as credible, even in a limited-trust environment.

Sensitive information must be protected at all stages of the life cycle of an AI model, from design to training to testing to implementation.

First, AI systems must be designed to protect sensitive information at all stages of the AI life cycle. Training modern AI models requires large volumes of data which, in an arms control context, may in some cases be open source, but will often originate from states, inspectors, and international organizations. Such data can contain information that extends beyond the scope of the treaty concerned and may reveal military capabilities, operational procedures, or technical characteristics that states are unwilling to disclose. Artificial intelligence models used for verification should therefore be developed according to the principle of data minimization. The technical parameters for this should ideally be developed and assessed by dedicated expert groups, thereby reducing the risk that highly technical issues become obstacles in political negotiations. Training data should be limited to the information necessary for the verification task, while remaining sufficiently informative to ensure that the AI's capabilities are not constrained. Where the boundary should be drawn here may not be clear-cut and will depend on both technical requirements and political circumstances. Rather than seeking to provide broad analytical capabilities, model architectures should be designed to answer narrowly defined verification questions. As research has shown that data can be extracted from trained models[7] despite safeguard measures, such measures must ensure that generated outputs do not exceed the agreed verification mandate. This design principle is especially critical with AI being such a fast-paced technology, which continually develops new capabilities as well as new vulnerabilities and attack vectors.

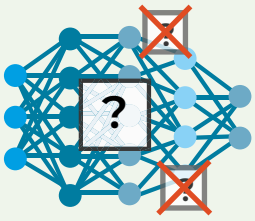
The integrity of both training data and AI models must be guaranteed to ensure credibility of the verification process.

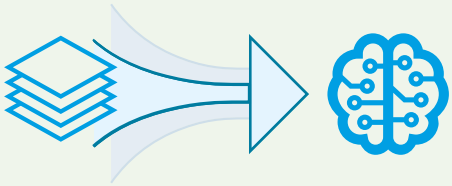
Second, the integrity of both data and models must be guaranteed. Verification ultimately depends on evidence. If the integrity of the underlying information can be called into question, the credibility of the entire process collapses. This challenge extends beyond protecting sensor data against tampering, as AI systems introduce additional attack surfaces through which adversaries may attempt to manipulate training datasets, inject misleading information into operational data streams, or modify model parameters—techniques commonly referred to as data or model poisoning.[8]

Conditions for Credible Use of AI in Verfication

1 Protect sensitive information at all stages of the AI life cycle

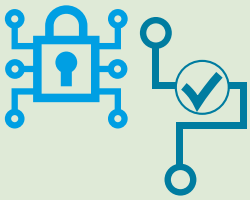
narrowly defined verification questions



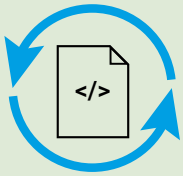


data minimization

2 Guarantee integrity of training data and models



3 Reproducibility as central design principle



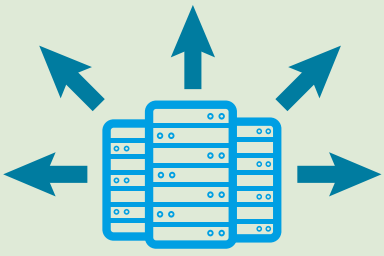
4 Decentralized verification architectures



5 Balanced transparency



6 Reduce asymmetries in hardware access



Consequently, verification architectures require strong mechanisms for ensuring provenance and traceability. Data streams should be cryptographically protected, stored in tamper-evident systems, and accompanied by verifiable chains of custody. Equally importantly, model versions and software updates must be documented and auditable, allowing changes to be tracked and reviewed. Secure data handling procedures already employed by organizations such as the International Atomic Energy Agency (IAEA) offer useful models that could be extended to AI-supported verification environments.

Reproducibility must be a central design principle, allowing all actors to independently assess limitations and evaluate results.

Third, reproducibility must become a central design principle, as it enables actors to test whether the model behaves as claimed, assess whether its limitations are properly understood, and independently reevaluate results. To achieve this, participating actors need to access source code and algorithms, detailed documentation of model architectures, hyperparameters, training data and procedures, preprocessing pipelines, and validation methods, so that they can recreate the training process under agreed conditions and obtain models with comparable behavior and performance characteristics. While exact bit-for-bit replication is often unrealistic given the stochastic elements of machine learning, comprehensive documentation and sharing of a specific technological setup can facilitate comparable outputs and statistically equivalent performance.

The verification architecture should be decentralized with functionally equivalent analytical capabilities operated by participating actors.

Fourth, verification systems should avoid excessive dependence on centralized infrastructures. Traditional verification mechanisms often rely on trusted international organizations such as the IAEA or the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO) to collect, process, and evaluate information. While such organizations play an indispensable role, the introduction of AI raises new concerns about covert manipulation of software, models, or data—whether through external cyberattacks or deliberate actions by participating actors. To mitigate this risk, participating actors could operate identical or functionally equivalent versions of the verification software and independently process the same datasets, thereby creating a decentralized verification architecture with equivalent analytical capabilities. Results could then be compared across multiple implementations and institutions and tested to see if different actors independently reach similar conclusions. This would, in turn, strengthen confidence in the verification process.

The requirement of transparency should focus on enabling actors to understand AI system designs, performance, and limitations instead of providing unrestricted access to all models and datasets.

Fifth, transparency must be understood as a practical rather than absolute requirement. Full transparency is neither politically realistic nor necessarily desirable, and it would probably undermine participating actors' intentions and willingness to compromise. The degree of transparency should therefore be sufficient to enable actors to understand what a system is designed to do, which data it uses, how its performance was evaluated, and where its limitations lie. This does not necessarily require unrestricted public access to all models and datasets. Approaches could include open-source model architectures, jointly developed reference implementations or curated training datasets, independent technical audits or certification procedures conducted by designated expert bodies, joint testing exercises, and benchmarking datasets.

Asymmetries in the international distribution of AI infrastructure create political and technical challenges that are more relevant to verification regimes than in other AI governance debates.

Lastly, the hardware environment deserves greater attention than it typically receives in AI governance debates. Advanced AI increasingly relies on specialized computing infrastructure, large-scale data centers, and access to high-performance processors. These resources are distributed unevenly across the international system. Such asymmetries create political as well as technical challenges.[9] Verification systems that rely on infrastructure controlled by only one or a small number of actors may create concerns about dependence, unequal influence, and hidden capabilities. Capability-building measures, together with support for or jointly managed computational facilities, can help allay concerns about technological asymmetries, thereby becoming important confidence-building measures and paving the way toward effective verification regimes.

Taken together, these requirements suggest that the main issue is not whether AI is perfectly explainable. Rather, the challenge lies in embedding AI systems in verification regimes in a manner that preserves reproducibility, integrity, transparency, and political acceptability. If these conditions are met, even complex AI systems may become credible verification tools. If they are neglected, however, even the most sophisticated AI systems may ultimately fail, because their results will not be accepted by the actors they are intended to convince.

Governance Trade-Offs

These requirements are demanding and create governance dilemmas of their own. States have incentives to shape datasets and model designs in ways that protect their own interests, while detailed knowledge of verification models could enable adversarial behavior or facilitate attempts to evade detection. Arms control agreements have long entailed negotiations over technical specifications, inspection procedures, and sensor capabilities. Similar debates surrounded issues such as, for example, the scope and accuracy of sensors used under the Open Skies Treaty.[10] Artificial intelligence should therefore not be viewed as a fundamentally different category of verification technology. Rather, it recasts familiar verification dilemmas in a new technical form. The question for policymakers and practitioners is therefore not whether AI should be used in verification, but under what conditions it is likely to strengthen rather than weaken the credibility of verification regimes.

For a short explainer on the Open Skies Treaty, see p. 132.

The central challenge here is not the existence of a "black box" as such, but ensuring that AI systems are designed and governed in a way that supports credibility in an environment characterized by limited trust. Reproducibility, integrity, transparency, decentralization, and robust security are the key prerequisites for the political acceptance of AI-supported verification. When these conditions are met, even complex AI systems can become credible and effective tools for arms control verification.

Thomas Reinhold

PRIF – Peace Research Institute Frankfurt

Fabian Unruh

PRIF – Peace Research Institute Frankfurt

1 International Atomic Energy Agency. (2026). *Development and implementation support programme for nuclear verification 2026–2027* (STR-409). International Atomic Energy Agency. https://www.iaea.org/sites/default/files/development-and-implementation-support-programme-for-nucear-verification-2026-2027.pdf

2 Kosinski, M. (2024, October 29). *What is black box artificial intelligence (AI)?* IBM Think. https://www.ibm.com/think/topics/black-box-ai

3 Phillips, P. J., Hahn, C. A., Fontana, P. C., Yates, A. N., & Greene, K. (2021). *Four principles of explainable artificial intelligence* (NISTIR 8312). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8312

4 Narayanan, M., Schoeberl, C., & Rudner, T. G. J. (2025). *Putting explainable AI to the test: A critical look at AI evaluation approaches.* Center for Security and Emerging Technology. https://doi.org/10.51593/20230003

5 Glaser, A., Barak, B., & Goldston, R. J. (2014). A zero-knowledge protocol for nuclear warhead verification. *Nature, 510*, 497–502. https://doi.org/10.1038/nature13457

6 UK–Norway Initiative. (2015). *Trust in verification technology: A case study: The UK–Norway information barrier* (Non-paper presented at the 2015 Review Conference of the Parties to the Treaty on the Non-Proliferation of Nuclear Weapons). University of Michigan, Consortium for Verification Technology. https://cvt.engin.umich.edu/wp-content/uploads/sites/173/2014/10/Trust_In_Verification_Technology-UKNI.pdf

7 Nakka, K. K., Frikha, A., Mendes, R., Jiang, X., & Zhou, X. (2025). PII-Scope: A comprehensive study on training data privacy leakage in pretrained LLMs. *Proceedings of the 14th International Joint Conference on Natural Language Processing and the 4th Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics*, 3731–3765. Association for Computational Linguistics. https://aclanthology.org/2025.ijcnlp-long.195/

8 Vassilev, A., Oprea, A., Fordyce, D., & Anderson, H. (2025). *Adversarial machine learning: A taxonomy and terminology of attacks and mitigations* (NIST AI 100-2). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-2e2025

9 Institute for Economics and Peace. (2026). *Global Peace Index 2026: Measuring peace in a complex world*. Institute for Economics and Peace. https://www.economicsandpeace.org/reports/

10 Dunay, P. (2004). *Open skies: A cooperative approach to military transparency and confidence building*. United Nations Institute for Disarmament Research. https://unidir.org/files/publication/pdfs/open-skies-a-cooperative-approach-to-military-transparency-and-confidence-building-319.pdf

Quantum Sensing

Verification with Quantum Sensing

Quantum sensors may eventually strengthen selected verification tasks, but the current evidence is uneven. A cold-atom gravity gradiometer has been tested above a known utility tunnel; navigation components have reached aircraft and ship trials; and nonlinear-interferometer spectroscopy has detected benign vapors in laboratory and open-path experiments. None of these constitutes a documented treaty-inspection deployment, and public comparisons with established instruments remain limited. The near-term priority is therefore not adoption but targeted, independently evaluated pilots with agreed protocols, managed-access designs, and clear rules for how a measurement could become usable evidence.

The key policy question is whether the sensor provides useful, validated information under realistic verification constraints.

Over the last couple of years, quantum research has moved beyond basic research in several areas, with quantum technologies now covering a wide range of devices and applications, from secure communication and computing to sensors using atoms, defects in diamond, or light. Some of these sensors are promoted for detecting objects over long distances or in noisy environments, including personnel, vehicles and facilities in the field. While they were not designed with arms control verification in mind, some may eventually support verification or lower its costs. Sensors generate observations that must be interpreted and combined with other evidence before states will accept them. A device that is more stable, selective, or deployable can shift a measurement from the laboratory to the field, from destructive sampling to noninvasive scanning, or from a fixed instrument to a mobile robot or drone. The key policy question is whether the sensor provides useful, validated information under realistic verification constraints. For the technologies discussed here, the evidence is promising in some areas but does not yet establish verification readiness.

In this text we discuss the pros and cons of the applications that appear most promising for future verification uses. First, we briefly describe what makes quantum sensing different from traditional sensing. We then highlight potential use cases before examining hurdles that may affect willingness to adopt new verification technology. We conclude that quantum technology is unlikely to revolutionize verification and that any practical benefit must be demonstrated against established alternatives.

Verification Frame: What Counts as Useful Evidence?

Verification is more complex than detection alone. Detection asks whether an instrument sees a signal; verification asks whether that signal can support a judgment about compliance with a treaty, declaration, inspection mandate, or agreed confi-

dence-building measure. Hence, a useful instrument must be linked to an inspection concept: what is being verified, where the measurement is allowed, what reference data are available, what level of uncertainty is acceptable, who handles the data, and what happens if the result is disputed. Even a technically strong result is not usable if the parties have not agreed to the instrument, procedure, and evidentiary role.

A useful instrument must be linked to an inspection concept.

This framing favors incremental use cases. Quantum gravimetry and magnetometry can reduce a search area or corroborate declared layouts. Quantum-assisted inertial navigation can document inspection routes and measurement locations where satellite navigation is unavailable. Nonlinear-interferometer spectroscopy can screen materials, gases, or residues for follow-up analysis. Each application still requires a legal basis, clear chain of custody, independent calibration, host-state protections, and confirmatory methods. The state of the art is best described as pre-verification: field or platform trials exist in adjacent applications, but the sensor classes discussed here have not been documented in treaty inspections.

How Quantum Sensors Help Solve Existing Verification Problems

Quantum sensors exploit controlled quantum states to measure external fields or optical properties with high precision. Cold atoms can act as matter-wave interferometers for gravity, acceleration, and rotation; atomic vapors and NV centers in diamond can serve as magnetometers; and parametric down-conversion in nonlinear crystals can link illumination at one wavelength with detection at another. These approaches are attractive because many verification-relevant signatures are weak: a tunnel slightly changes the local gravity gradient, a buried metal casing perturbs a magnetic field, and a chemical compound absorbs only at characteristic wavelengths. These technologies are dual-use: the same advances can serve civil, defense, or verification applications.[1]

The main contribution of this technology is to improve parts of the measurement chain. Quantum sensors may lower uncertainty, reduce drift, make a field measurement more repeatable, or provide a more selective screening signal. They must still deal with terrain, vibration, temperature drift, atmospheric absorption, sampling bias, adversarial denial and concealment, as well as operator error. Their policy value lies in broadening the toolbox available for verification tasks, but they cannot determine which measurement questions are worth pursuing in the field, which data can be shared or filtered, or which results require confirmatory laboratory analysis. Any claim of improvement must therefore be tested at the level of individual tasks and instruments, not inferred from the quantum label.

Any claim of improvement through quantum sensing must be tested at the level of individual tasks and instruments.

Quantum-Based Gravimetry: Supporting Local Underground Surveys

Quantum gravimeters and gravity gradiometers measure tiny variations in gravitational acceleration or in how that acceleration changes over short distances. Because gravity cannot be shielded, dense structures and hollow spaces can create measurable anomalies. In a controlled field test above a known utility tunnel, a cold-atom gravity gradiometer reported 1.5 to four times better statistical performance than commercial survey gravimeters, although overall repeatability was similar once systematic effects were included.[2] It was a static test on a known target, and each survey point took approximately ten minutes.

Quantum gravimeters could support the search for undeclared tunnels, cavities, shafts, large buried objects etc.

For verification, possible applications include the local search for undeclared tunnels, cavities, shafts, large buried objects, or geophysical changes associated with under-

ground nuclear testing. Detection depends on target size, depth, density contrast, sensor stability, survey geometry, and corrections for terrain and geology. Satellite gravimetry is valuable for large-scale Earth-system monitoring, but its spatial resolution is far too coarse to identify individual facilities or submarines.[3] The most realistic benefit is improved site-survey evidence, particularly when gravimetry is complemented by ground-penetrating radar, seismic methods, maps, and visual inspection.

For verification, gravimetry is most valuable as an anomaly-finding and corroboration tool.

In verification terms, gravimetry is most valuable as an anomaly-finding and corroboration tool. It can help inspectors decide where to focus more intrusive methods, while interpretation still requires contextual evidence. Any deployment would need pre-agreed survey grids, terrain models, uncertainty thresholds, and rules for handling geophysical data that may reveal unrelated infrastructure. Limited inspection time makes survey speed as important as sensitivity.

Quantum-Based Magnetometry: Local Magnetic-Anomaly Mapping

Magnetometers measure disturbances in magnetic fields. Quantum magnetometers based on atomic vapors, superconducting devices, or nitrogen-vacancy centers in diamond can be extremely sensitive. In security debates, they are often associated with submarine detection, but magnetic signatures decrease rapidly with distance and can be masked by geology, infrastructure, vehicles, cables, and natural field variations.

Magnetometry can support on-site inspections, e.g. to locating metallic components or magnetic field mapping.

The near-term verification value is local and inspection oriented. Magnetometry can help locate buried metallic components, cables, casings, drums, or equipment; check whether declared layouts match observed magnetic anomalies; or support on-site inspection methods that already include magnetic field mapping or other geophysical techniques in certain phases and according to defined procedures.[4] The potential quantum advantages are improved sensitivity, stability, vector measurement, and miniaturization. However, the literature reviewed here does not provide a verification-specific field comparison against high-end classical magnetometers. In many tasks, classical instruments may in fact be sufficient.

As evidentiary method, magnetometry is most effective when used to guide sampling, excavation, visual inspection, or other geophysical measurements. A verification protocol should therefore specify what counts as a reportable anomaly, what reference measurements are needed, and how inspectors avoid overinterpreting clutter in industrial or military environments.

Quantum-Assisted Inertial Navigation: Documenting Inspections When External Signals Fail

Inertial navigation determines the position of a vehicle or inspection platform by measuring its acceleration and rotation to pinpoint the location relative to an initial position. It does not need GPS, but small errors accumulate over time. Cold-atom accelerometers, gyroscopes, and clocks may reduce that drift and improve navigation where satellite signals are unavailable, jammed, spoofed, or not permitted. Public UK flight trials in 2024 and Royal Navy trials in 2025 show that relevant components can operate on mobile platforms, but the published reports do not provide an end-to-end performance comparison with conventional inertial navigation.[5]

Inertial navigation can be useful for inspection autonomy and documentation in verification.

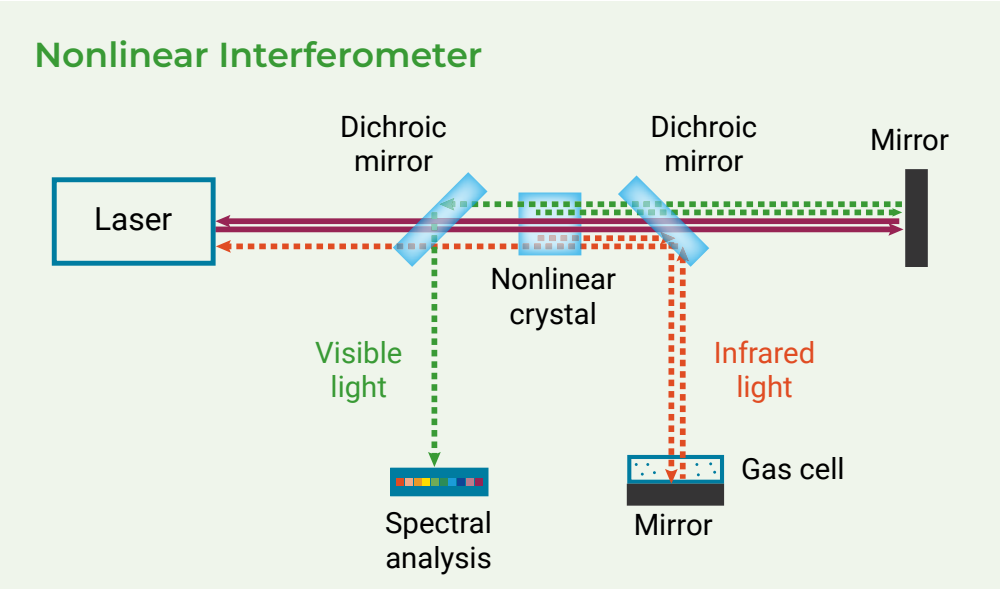
For arms control verification, the most relevant use is inspection autonomy and documentation. Using inertial navigation, robots, vehicles, drones, or inspectors could navi-

gate inside buildings, tunnels, ships, or cluttered sites where GPS is unavailable, compare actual routes and spaces with declared design information, and document where measurements were taken. This could improve repeatability, reduce reliance on host-provided infrastructure, and make subsequent reviews of inspection results more robust.

The verification caveat is that navigation data may itself constitute sensitive information. A high-resolution route record may reveal facility layouts, access constraints, or process flows. Quantum-assisted navigation would require map integration, classical inertial systems, procedural safeguards, host-state review where appropriate, and human oversight. It should be treated as part of the evidence-management system, not merely as a convenience for inspectors.

Nonlinear Interferometers: Chemical and Biological Screening with Light

Nonlinear interferometers are a particularly promising but still emerging quantum sensing option for verification. In a conventional interferometer, a beam splitter creates two paths that are later recombined. In a nonlinear interferometer, special crystals create linked light paths through parametric down-conversion. One path can pass through the sample at a useful wavelength, while the other can be measured with a detector that is easier to operate. Because the two paths are linked, the measured light carries information about what happened in the sample path even when the sample wavelength itself is not detected directly. This principle emerged from induced-coherence experiments and was later extended to imaging of light that is not directly detected.[6]



Scheme of a nonlinear interferometer for spectral sensing in the infrared while detecting visible light only. Main components are a laser that pumps a nonlinear crystal generating a photon-pair beam via spontaneous parametric down-conversion. The setting as an interferometric setup enables sensing with undetected light. While the infrared light can sense a gas in the gas cell, its partner light in the visible spectral range is detected with a spectrometer, revealing the infrared spectral properties of the gas.

The arms control relevance of this technology comes from wavelength separation. Many chemicals and biological materials have strong and selective absorption features in the mid-infrared fingerprint region. Conventional mid-infrared spectroscopy is powerful, but its sources and detectors are often costly, noisy, bulky, or require cooling. Nonlinear interferometry can probe samples in the mid-infrared while using mature visible- or near-infrared detectors. Reviews and experiments have demonstrated the feasibility of infrared spectroscopy with visible readout, gas spectroscopy using near-infrared detection, near-infrared grating-spectrometer readout of mid-infrared spectra, biological and hyperspectral microscopy with this approach, and ultra-broadband quantum infrared spectroscopy.[7] In 2024, an open-path experiment detected methane below 4 parts per million and common solvent vapors at roughly 9–14 parts per million in ambient air.[8] These figures show technical progress, but not yet an advantage over portable Fourier-transform infrared (FTIR) spectroscopy or readiness for inspections.

Wavelength separation makes nonlinear interferometers promising for detection of chemicals and biological materials.

The most realistic application of nonlinear interferometers in chemical weapons verification is screening and triage.

For chemical weapons verification, the most realistic application is a screening and triage instrument rather than a replacement for accredited laboratory analysis. During routine or challenge inspections, nonlinear-interferometer spectroscopy could help screen declared chemicals, precursors, solvents, degradation products, residues, exhaust streams, or air samples before resorting to more intrusive sampling. This should be framed carefully: OPCW-related analytical practice and training already includes FTIR, portable FTIR, Raman spectroscopy, and GC/MS for rapid identification and verification support.[9] Handheld infrared capability alone is therefore nothing new. The possible added value lies in combining mid-infrared probing with visible or near-infrared readout, lower detector burden, tailored spectral windows, and greater compatibility with managed-access screening.

For short introductions to the BWC and the CWC, see p. 127 ff.

For biological verification and investigations, the potential is more limited and should be described carefully. The Biological Weapons Convention (BWC) does not have a standing verification regime comparable to the Chemical Weapons Convention (CWC). Consequently, nonlinear-interferometer microscopy or spectroscopy would be relevant mainly to investigations, confidence-building, biosafety inquiries, or future verification concepts rather than routine BWC inspections.[10] Mid-infrared microscopy can distinguish broad classes of biomolecular material, such as lipids, proteins, carbohydrates, or culture media, and can reveal spatially resolved chemical patterns. Its value would lie in serving as a nondestructive screening and prioritization method: flagging samples, surfaces, or production residues for follow-up analysis, or supporting biosafety investigations in contexts where low light exposure and label-free measurement matter.

Potential Verification Uses and Maturity

Verification Question	*Sensor Option*	*Evidence Generated*	*Main Caveat*
Is there an undeclared underground void, shaft, tunnel, or large buried object near an inspected site?	Cold-atom gravity gradiometry or hybrid quantum/classical gravimetry	Controlled detection of a known tunnel, supporting more focused follow-up inspection	Static known-target test; survey speed, terrain, and geology remain important, and overall repeatability was similar to commercial instruments
Do declared layouts match local magnetic anomalies, or are buried metallic components present?	Quantum or high-end classical magnetometry, including atomic-vapor or NV-center approaches	Local magnetic-anomaly evidence that can guide visual inspection or other geophysical methods; no quantum-specific verification advantage yet been shown	Clutter, distance and background variation limit interpretation; classical instruments may be sufficient in many cases
Can inspectors document routes and measurement locations in GPS-denied or host-controlled spaces?	Cold-atom inertial sensors, gyroscopes, and quantum clocks integrated with classical navigation	Potentially more reproducible route and location records for inspectors, robots, or drones inside complex facilities	No published end-to-end advantage; still requires maps, classical inertial systems, drift correction, data minimization, and human procedures
Can chemical substances, vapors, or residues be screened before more intrusive sampling?	Nonlinear-interferometer mid-infrared spectroscopy with visible or near-infrared readout	Proof-of-principle detection of benign vapors, with prospective role in triage and managed-access screening	Must be compared with portable FTIR and tested with mixtures, humidity, and realistic backgrounds before use in inspections
Can biological material or process residues be prioritized for follow-up analysis?	Undetected-light microscopy or hyperspectral imaging	Label-free maps of broad biomolecular signatures in residues, samples, or surfaces	Not pathogen identification; must be paired with established genomic, immunological, culture-based, or mass-spectrometric methods

Note: Maturity varies considerably depending on exact instrument configuration. Most entries describe prospective verification roles, not documented treaty use.

These developments justify targeted research and carefully designed field trials. High-gain and loss-compensated schemes may improve sensitivity or robustness in selected configurations, yet most systems remain laboratory instruments rather than rugged field devices.[11]

Verification Design: From Instrument Readings to Accepted Evidence

A verification system is a political construct and thus more than a sensor. It includes the inspection mandate, managed-access rules, data handling, calibration, chain of custody, and dispute-resolution procedures. Such a regime typically verifies declarations of treaty-controlled substances, such as chemical weapons compounds or nuclear fissile material, and checks for any undeclared activities that might contravene an agreement. New technological methods therefore also have a political dimension that must be addressed before use. Technical validation alone does not guarantee acceptance: inspected states may reject a novel instrument because of mandate, confidentiality, safety, precedent, or other political concerns. Quantum sensing should therefore be introduced through joint validation involving verification parties and vendors, rather than through isolated technology demonstrations. Across all domains, trials should test not only sensitivity but also false positives and negatives, operating time, data security, operator training, ruggedness, and inspector usability.

Quantum sensing should be introduced through joint validation involving verification parties and vendors.

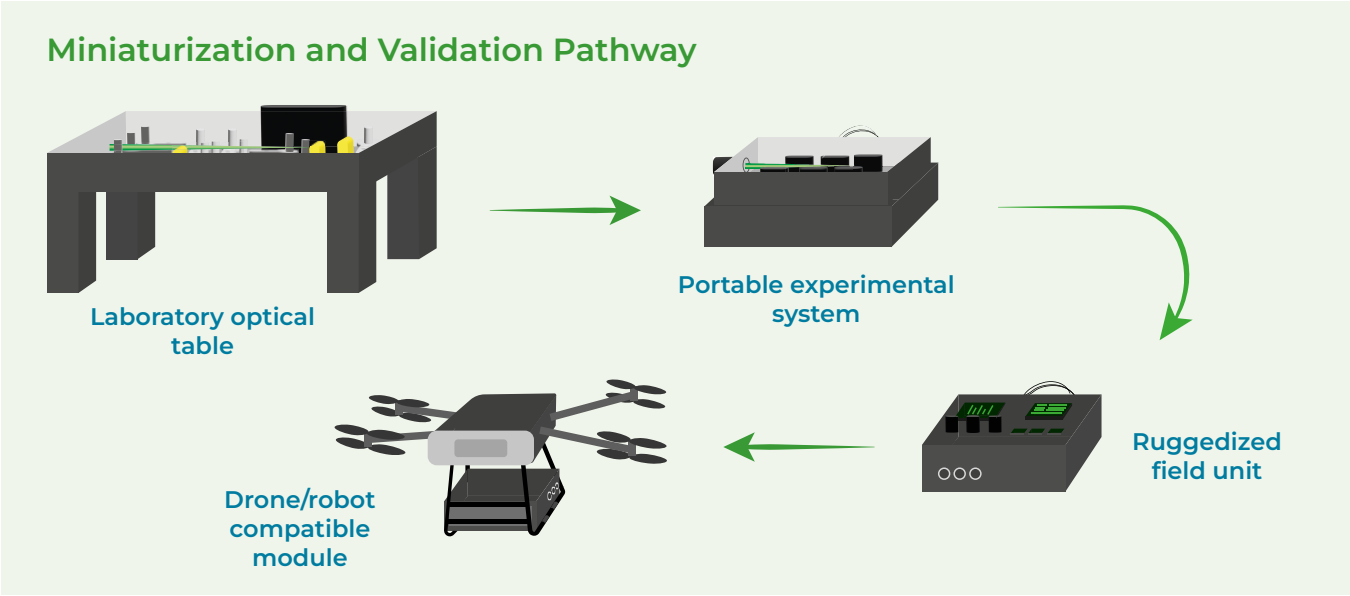
Illustrative pilots would make policy recommendations more concrete. For gravimetry and magnetometry, mock sites should include known voids, pipes, metallic clutter, and realistic terrain corrections. For inertial navigation, trials should test whether quantum-assisted navigation improves route reproducibility without exposing sensitive facility details. For nonlinear interferometers, tests should use safe chemical simulants, mixtures, humidity changes, and realistic background conditions, with field readings subsequently compared with established infrared spectroscopy and mass-spectrometry-based laboratory methods.

Managed access is important for all quantum sensors because a measurement can reveal information—such as facility layouts, buried infrastructure, route data, or industrial process signatures—that extends beyond the inspection question. This is particularly important for spectroscopic sensors. A host state or facility operator may object that a broad-spectrum instrument could reveal proprietary or militarily sensitive information unrelated to the inspection mandate. Nonlinear interferometers offer a design advantage here: the probe wavelength, detected wavelength, spectral range, and reference library can be engineered. In principle, a device could be built as a broad spectrometer for laboratory investigations or as a restricted-range instrument that only tests treaty-relevant bands. Physical filters, locked software, cryptographic logging, jointly held calibration samples, or jointly operated classifiers could reduce unnecessary data collection while preserving evidentiary value. The same filter logic can also be applied to classical photonic techniques.

Nonlinear interferometers can be designed both as broad spectrometers and as restricted-range instruments only for treaty-relevant testing.

The distinction between technology and data must be explicit. For decision-making, the crucial issue is not the full disclosure of proprietary sensor designs, but whether the data are reliable, properly calibrated, traceable, and generated in accordance with certified procedures. A verification organization does not need to reveal every technical detail of an instrument if performance has been validated by agreed test protocols, calibration standards, tamper-evident logging, and independent review. This is a general verification principle rather than a quantum-specific exception.

Access to technology is also likely to be mediated by commercial suppliers. A sharp technical divide between states is not the central concern at this stage; many systems



From laboratory setup to deployable system: Nonlinear-interferometer technologies will need to move from optical tables to compact, ruggedized units before they can support field verification, including future drone- or robot-compatible sensing modules.

will be provided by specialized firms, spin-offs, or start-ups rather than built domestically by each verification actor. Commercial quantum gravimeters and companies developing cold-atom sensors, quantum-diamond magnetic sensors, and compact cold-atom hardware already exist.[12] The policy task is therefore to certify instruments, publish nonsensitive performance data, enable shared training, and avoid dependence on opaque vendor claims.

The most important policy step is to define evidentiary thresholds before deployment.

The most important policy step is to define evidentiary thresholds before deployment. Inspectors, host states, and verification organizations should know whether a sensor result is a screening flag, corroborating evidence, a trigger for follow-up sampling, or part of a formal compliance assessment. Without this hierarchy, even technically impressive sensor results may be politically unusable.

Limitations, Risks, and Areas Requiring Caution

The main risk is premature confidence. Across all sensor types, performance depends heavily on the measurement setting, quality of calibration, availability of reference data, and the way results are interpreted. Quantum gravimeters can improve local underground mapping, but measurements can be slow, sensitive to vibration, and difficult to interpret in complex geology or built environments. Magnetometers can detect magnetic anomalies, while buried infrastructure, vehicles, cables, and natural field variations can produce misleading signals. Quantum inertial systems can reduce navigational drift, yet integration with maps, classical sensors, and inspection procedures remains essential.

The main risks at this point are premature confidence, confidentiality, as well as political and procedural rejection.

For nonlinear interferometers, laboratory spectroscopy is more mature than drone-mounted or open-path sensing, where alignment, humidity, background gases, sunlight, platform motion, and mixed samples can strongly affect performance. Detection limits for warfare-relevant chemicals should be established through realistic tests with safe simulants rather than inferred from ideal laboratory conditions. Compact lab-in-a-box systems are plausible future development goals, but do not yet reflect current capabilities. For biological applications, these sensors are best described as screening tools: they can help flag suspicious material but reliable agent identifica-

tion still requires established genomic, immunological, culture-based, or mass-spectrometric methods.

A second risk is confidentiality. More capable field sensors may increase the amount of collateral information collected during an inspection. Verification organizations should therefore design quantum-sensor pilots alongside data-minimization rules: restricted spectral windows, pre-approved maps, on-site processing, tamper-evident logs, joint inspector–host review, and deletion or redaction rules for data outside the mandate. Such measures are a precondition for making new measurements politically acceptable.

A third risk is political and procedural rejection. Even a well-validated instrument may be refused if parties consider it too intrusive, insufficiently transparent, or outside the agreed mandate. Technical trials should therefore be accompanied by early consultation on the acceptability of equipment, data access, and rules governing decision-making.

Conclusions: Incremental Change, Not Revolution

Quantum sensing may eventually matter for verification because it could make certain measurements more mobile, selective, stable, and robust. Gravimetry may improve local underground mapping; magnetometry may support searches for metallic artifacts; inertial navigation could make inspections less dependent on external signals; and nonlinear interferometers could bring mid-infrared screening closer to field use. The evidence reviewed here does not yet show that these systems are ready for treaty inspections or consistently outperform mature alternatives. The main recommendations are pragmatic: fund a small number of comparative pilots focused on concrete verification questions, specify evidentiary thresholds, certify the data chain, and agree data-governance and equipment-approval rules before deployment. If those trials show a meaningful advantage, quantum sensing could become a useful addition to the verification toolkit.

Markus Gräfe

Technical University of Darmstadt

Yan-Jie Schnellbach

PRIF – Peace Research Institute Frankfurt/
Technical University of Darmstadt

1 Krelina, M. (2025). *Military and Security Dimensions of Quantum Technologies: A Primer*. SIPRI. https://doi.org/10.55163/ZVTL1529
European Commission. (2025). *Quantum Europe Strategy: Quantum Europe in a Changing World* (COM(2025) 363 final). https://digital-strategy.ec.europa.eu/en/library/quantum-europe-strategy

2 Stray, B., Lamb, A., Kaushik, A., Vovrosh, J., Rodgers, A., Winch, J., Hayati, F., Boddice, D., Stabrawa, A., Niggebaum, A., Langlois, M., Yu-Hung, L., Lellouch, S., Roshanmanesh, S., Ridley, K., de Villiers, G., Brown, G., Cross, T., Tuckwell, G., Faramarzi, A., Metje, N., Bongs, K., & Hlynski, M. (2022). Quantum sensing for gravity cartography. *Nature, 602*, 590-594. https://doi.org/10.1038/s41586-021-04315-3

3 GFZ German Research Centre for Geosciences. *GRACE-FO | Gravity Recovery and Climate Experiment – Follow-On Mission*. https://www.gfz.de/en/section/global-geomonitoring-and-gravity-field/projects/grace-fo-gravity-recovery-and-climate-experiment-follow-on-mission

4 CTBTO. *On-site Inspection Techniques*. https://www.ctbto.org/our-work/on-site-inspection/on-site-inspection-techniques

5 Government of the United Kingdom. (2024, May 13). *Un-jammable quantum tech takes flight to boost UK's resilience against hostile actors*. https://www.gov.uk/government/news/un-jammable-quantum-tech-takes-flight-to-boost-uks-resilience-against-hostile-actors

Royal Navy. (2025, December 19). *Royal Navy carries out Arctic trial on GPS-free navigation with Imperial College London*. https://www.royalnavy.mod.uk/news/2025/december/19/20251219-royal-navy-trials-quantum-tech-with-imperial-college-londo

6 Zou, X. Y., Wang, L. J., & Mandel, L. (1991). Induced coherence and indistinguishability in optical interference. *Physical Review Letters 67*, 318-321. https://doi.org/10.1103/PhysRevLett.67.318
Barreto Lemos, G., Borish, V., Cole, G. D., Ramelow, S., Lapkiewicz, R., & Zeilinger, A. (2014). Quantum imaging with undetected photons. *Nature 512*, 409-412. https://doi.org/10.1038/nature13586

7 Fuenzalida, J., Giese, E., & Graefe, M. (2024). Nonlinear Interferometry: A New Approach for Imaging and Sensing. *Advanced Quantum Technologies 7,* 2300353. https://doi.org/10.1002/qute.202300353
Kalashnikov, D. A., Paterova, A. V., Kulik, S. P., & Krivitsky, L. A. (2016). Infrared spectroscopy with visible light. *Nature Photonics 10*, 98-101. https://doi.org/10.1038/nphoton.2015.252
Lindner, C., Kunz, J., Herr, S. J., Wolf, S., Kießling, J., & Kühnemann, F. (2021). Nonlinear interferometer for Fourier-transform mid-infrared gas spectroscopy using near-infrared detection. *Optics Express 29*(3), 4035-4047. https://doi.org/10.1364/OE.415365
Kaufmann, P., Chrzanowski, H. M., Vanselow, A., & Ramelow, S. (2022). Mid-IR spectroscopy with NIR grating spectrometers. *Optics Express 30*(4), 5926-5936. https://doi.org/10.1364/OE.442411
Kviatkovsky, I., Chrzanowski, H. M., Avery, E. G., Bartolomaeus, H., & Ramelow, S. (2020). Microscopy with undetected photons in the mid-infrared. *Science Advances 6*(42), eabd0264. https://www.science.org/doi/10.1126/sciadv.abd0264
Paterova, A. V., Maniam, S. M., Yang, H., Grenci, G., & Krivitsky, L. A. (2020). Hyperspectral infrared microscopy with visible light. *Science Advances 6*(44), eabd0460. https://www.science.org/doi/10.1126/sciadv.abd0460
Tashima, T., Mukai, Y., Arahata, M., Oda, N., Hisamitsu, M., Tokuda, K., Okamoto, R., & Takeuchi, S. (2024). Ultra-broadband quantum infrared spectroscopy. *Optica 11*(1), 81-87. https://doi.org/10.1364/OPTICA.504450

8 Neves, S., Kartiyasa, A., Ghosh, S., Gaulier, G., La Volpe, L., & Wolf, J.-P. (2024). Open-path detection of organic vapors via quantum infrared spectroscopy. *APL Photonics 9*, 096108. https://doi.org/10.1063/5.0220707

9 Organisation for the Prohibition of Chemical Weapons. *Verification*. https://www.opcw.org/about/technical-secretariat/divisions/verification
Organisation for the Prohibition of Chemical Weapons.(2026). *Call for Nominations for a Course for Analytical Chemists from Laboratories Supporting Customs Services* (S/2501/2026). https://www.opcw.org/sites/default/files/documents/2026/03/s-2501-2026%28e%29.pdf

10 Tuzmukhamedov, B. (2021). *Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on their Destruction. Introductory Note*. United Nations. https://legal.un.org/avl/ha/cpdpsbbtwd/cpdpsbbtwd.html

11 Hashimoto, K., Horoshko, D. B., & Chekhova, M. V. (2023). Broadband spectroscopy and interferometry with undetected photons at strong parametric amplification. *Advanced Quantum Technologies 8*(4), 2300299. https://doi.org/10.1002/qute.202300299
Gemmell, N. R., Flórez, J., Pearce, E., Czerwinski, O., Phillips, C. C., Oulton, R. F., & Clark, A. S. (2023). Loss-compensated and enhanced midinfrared interaction-free sensing with undetected photons. *Physical Review Applied 19*, 054019. https://doi.org/10.1103/PhysRevApplied.19.054019

12 Exail. *Quantum gravimeters*. https://www.exail.com/product/quantum-gravimeters
AOSense, Inc. *Sensors for a Quantum World*. https://aosense.com/
Qnami. *Quantum Sensing Leaders in Nanoscale Precision*. https://qnami.ch/
SBQuantum. *Quantum Diamond Magnetic Sensing Technology*. https://sbquantum.com/

Space

From Images to Insights: Data Integration in Satellite Surveillance

Satellites are important tools for both intelligence gathering and verification. While advances in optical, infrared, and radar technology have improved satellite resolution, further gains are subject to diminishing returns and disproportionally increasing costs. That said, recent developments in artificial intelligence (AI) and data fusion will make it possible to combine and integrate lower-quality data from different sources to generate higher-quality information. This could be used by both state and non-state actors and could have a significant impact on government secrecy policy—for better or for worse. Following an introduction to the historical use of satellites for verification purposes, the article goes on to outline the different imagery technologies along with their respective limitations. It then sets out the potential for broader verification, while also highlighting the potentially negative strategic implications of improved satellite reconnaissance. The text concludes with a number of general recommendations, as well as a specific recommendation for German space security policy.

Space Surveillance as "National Technical Means"

Since the 1960s, the United States, the Soviet Union, and, later, also China along with other actors, have used satellites to spy on each other, replacing more vulnerable solutions such as the high-altitude U2 "Dragon Lady" reconnaissance aircraft, which was famously shot down over Soviet territory in 1960. On the positive side, information gathered by satellites, encompassed by the term "verification by national technical means (NTM)"[1] in arms control treaties, has increased transparency by helping to verify or refute claims regarding missile deployments and military bases. Historically, these capabilities remained closely held secrets; neither the raw imagery nor the sensor specifications were shared, in order to prevent adversaries from gauging the technology's limitations.[2] Yet, as the nuclear superpowers appeared to be broadly compatible when it came to technological capabilities, it is reasonable to conclude that the information each gathered made a significant contribution to the stability of the nuclear balance.[3] The first bilateral nuclear arms control agreements—SALT (Strategic Arms Limitation Talks, 1972) and the ABM Treaty (Anti-Ballistic Missile Treaty, 1972)—relied solely on national technical means for verification, with only the later INF Treaty

National technical means (NTM): A nation's independent intelligence and monitoring assets, usually used in the context of arms control treaties.

For a brief explainer of SALT and the ABM, see p. 124.

(Intermediate-Range Nuclear Forces Treaty, 1987) and the START I (Strategic Arms Reduction Treaty, 1991) adding on-site inspections to NTM. At the same time, the treaties prohibited restrictions of NTM in the field of strategic weapons, including through concealment. However, monitoring another country's strategic arsenal has not been the only use of satellites. Today, reconnaissance satellites are deployed for a much wider range of purposes, including detecting and tracking troop movements on the ground, in the air, and, within limits, at sea; assessing the destruction of infrastructure ("battle damage assessment"); and providing early warning of missile launches.

Among the key players, the U.S., China, and Russia lead in sheer numbers, each operating a fleet of more than 100 military satellites, including specialized reconnaissance satellites. Of the countries with smaller fleets, France, Israel, Germany, the UK, India, Italy, Japan, and South Korea are known to operate their own optical and radar-imaging satellites. Although such information is generally classified, Brazil is reportedly operating a radar satellite through a private company capable of penetrating the jungle canopy, while Indonesia and South Africa are believed to have similar programs, although these can presumably be described as nonoperational.

Earth observation from space relies on a range of remote sensing techniques, including electro-optical (visual), infrared (thermal), and synthetic aperture radar (SAR). Section 1 in the box on Technical Background provides detailed information on each sensing type and the current state of the art. Taken individually, each sensing modality is subject to distinct physical and technical limitations. However, these limitations can both be overcome through and provide a rationale for integrating multiple sensing domains into a unified analytical framework, yielding spectacular new results.

For the technical background on taking images in space, see p. 38 f.

From Individual Images to Space Data Integration

As multiorbital constellations expand in scale and revisit frequency, space-based sensing is becoming increasingly central to the future of military air and multidomain operations.[4] This transformation carries a paradox familiar to network-centric warfare theory:[5] the integration of distributed sensing architectures enhances situational awareness and combat effectiveness, while simultaneously increasing dependence on a functioning and survivable satellite network.

Revisit frequency: The time a satellite system takes to rotate through its orbit and pass over the exact same geographic location on the rotating Earth again.

The role of satellite technologies in cross-platform collaboration and detection[6]—particularly in enabling real-time monitoring and targeting, as well as improving anti-stealth capabilities—has increased significantly over the past few years. China's recent investments in satellite constellations and autonomous sensing systems illustrate this broader shift. Systems such as the Ludi Tance No. 1 (LT-1) L-band synthetic-aperture radar satellite constitute enabling infrastructure: high-resolution optical imaging, combined with all-weather radar observation and increasing revisit rates, is contributing to unprecedented, persistent monitoring capabilities.

L-band: The 1 to 2 gigahertz (GHz) range of the radio spectrum, as designated by the Institute of Electrical and Electronics Engineers (IEEE).

The analytical significance of modern satellite networks is vividly illustrated by a striking 2022 report: imagery from the Chinese Jilin-1 constellation appeared to capture a U.S. F-22 Raptor in flight. While tracking a stealth fighter from orbit is not yet a total breakthrough in stealth detection per se, it signals a massive shift in capability. This success is driven by a new "enabling infrastructure"—comprising systems such as the aforementioned Jilin-1 optical constellation and LT-1 radar satellites. By integrating high-resolution optical data with all-weather radar and rapid revisit rates, these shared architectures are turning sporadic glimpses into persistent, real-time monitoring. The decisive step is **data fusion**: by combining optical, radar, infrared, and radio frequency monitoring systems, this transforms detection from a platform-centric problem into

Radio frequency (RF): The rate of oscillation of an electromagnetic wave or electrical signal, typically ranging from about 3 kilohertz (kHz) to 300 gigahertz (GHz) (1 Hertz = 1 oscillation per second).

a cross-platform correlation problem. Seemingly unrelated events can be integrated into a broader analytical picture, in which their combined pattern may be significant even though none of the individual events, considered in isolation, may be. Weak signals, environmental disturbances, thermal signatures, airfield activity patterns, and radio frequency (RF) anomalies may individually fall below actionable thresholds. When aggregated and analyzed across domains, however, they reduce ambiguity. Aircraft may minimize radar cross section in specific frequency bands and geometries, but they cannot eliminate all physical interactions with the electromagnetic and thermal environment. Engine exhaust produces infrared signatures; airframes perturb ambient radio frequency fields; and movement generates detectable pattern-of-life deviations. Much like assembling a jigsaw puzzle, putting the individual pieces—in this case, data from different satellites—together reveals the whole picture.

As orbital sensing layers become persistent and digitally integrated, these cross-domain perturbations become searchable. A complementary example of multidomain exploitation is provided by research on passive radio frequency monitoring. Chinese studies[7] have explored how signals emitted by low-earth orbit (LEO) communication satellites, such as Starlink, could be monitored to detect aircraft or UAVs through subtle perturbations in the ambient electromagnetic environment. This principle—that weak, noncooperative signals can measurably interact with airborne platforms—is supported by foundational work on the effects of solar radio bursts on aviation frequency bands,[8] as well as subsequent studies examining space weather impacts on drones[9] and applied research into RF interference mitigation in UAV communication systems.[10] These findings demonstrate that the RF spectrum is not a blank slate: it contains intrinsic weaknesses and ambient signals that cannot be eliminated. When combined with optical, radar, and other sensing modalities within a persistent, multiorbital informatics framework, these spectral observables create exploitable opportunities for detecting low-observable platforms.

Uncrewed aerial vehicle (UAV): An aircraft without a human pilot aboard, commonly known as a drone.

The more extensive the sensing network, and the more sophisticated the analytic layer, the fewer environmental variables remain exploitable for undetected operation. While these capabilities enhance monitoring and treaty verification, they also raise questions about where legitimate verification ends and intrusive surveillance begins. Arms control arrangements, such as the Open Skies Treaty, have addressed this tension by establishing mutually agreed limits on sensing capabilities, recognizing that enduring verification depends not only on technical effectiveness but also on political acceptance of the scope of observation.

For a short explainer on the Open Skies Treaty, see p. 132.

Enhanced Verification, Emerging Vulnerabilities: Rethinking Control in Data-Driven Conflict

From the perspective of arms control and verification, more accurate and more reliable information has always been regarded as positive. Sensing technologies have proliferated significantly, individually enabling more effective monitoring and verification. Data fusion has facilitated significant advances in verification and reconnaissance, particularly in monitoring movements of stealthy military assets. Based on this technology, states could, for example, verify that certain strategic assets such as stealth bombers have not been deployed near their border but remain at their distant home bases. But with the major bilateral nuclear disarmament treaties dead, satellite-based monitoring continues only informally through national means—without any treaty framework to mandate or structure it, thereby shifting the fine line between verification and spying more toward the latter than is (grudgingly) accepted.

States have consistently sought to keep certain secrets concealed, leaving their adversaries uncertain about the accuracy of their intelligence or whether they might still have an ace up their sleeve. With the advances in data-fusion technology, we are heading toward a state of "inevitable transparency"—with all its pros and cons. In addition, it might not only be opponents who gain a better understanding of security-relevant information; civilian actors, too, may be able to uncover sensitive knowledge and make it publicly accessible. This presents a challenge for remote sensing policy. Where comparable imagery can be obtained from international commercial providers, restricting domestic operators alone may do little to prevent access to sensitive information while imposing costs on legitimate civilian users. This raises the question: How will states react to this development?

Given that actors might balk at openly attacking both state-run and civilian space assets, resilience at the hardware layer can therefore generate vulnerabilities at the electromagnetic and informational layers. The focus of attacks would thus shift to data support functions, including ground stations or even AI computer centers responsible for the fusion of the data.

As a consequence, strategic competition is likely to shift away from kinetic anti-satellite engagement toward cyber and electronic means of interference.[11] Given that these activities are difficult to observe and attribute, they fall outside many of the assumptions underpinning traditional arms control verification. This suggests that future discussions may need to consider issues such as the resilience of verification systems, protection against interference with critical data flows, and measures to reduce the risk of false or escalatory warnings. However, states should be aware of the problems they might face in the not-so-distant future.

One possible confidence-building meausre would be a binding commitment to authenticate data entering verification pipelines.

As verification increasingly relies on AI-assisted fusion of data from multiple space-based and terrestrial sensors, the integrity of the underlying data becomes progressively more important. One possible confidence-building measure would be a politically binding commitment to authenticate data entering agreed verification pipelines—for example, through digital signatures or comparable cryptographic mechanisms—to reduce the risk that manipulated or fabricated data could contribute to false assessments or unnecessary escalation. This would also involve civilian actors providing publicly available data. The extent to which states could or should regulate the public sale of high-resolution imagery is open to debate—traditionally, shutter control policies have had to balance state security requirements with the need to foster a viable commercial market for imagery products. The recent case of Planet Labs withholding commercial imagery of the Middle East sets a new precedent, while also presenting policymakers with a fresh challenge. As Kelevitch[12] argues, the fact that shutter control can work effectively enough to restrict access to commercially available imagery presents a concern. The international reach of shutter control to restrict civilian monitoring during conflict became evident when U.S. authorities submitted a diplomatic démarche to the EU on July 13, which subsequently resulted in a delay in the provision of Copernicus imagery under the EU Space Threat Response Architecture.

See also the chapter on societal verification, p. 71 ff.

These developments raise questions about whether existing national frameworks remain adequate for an increasingly international commercial Earth observation market. In this context, Germany should reassess whether its National Data Security Policy for Space-Based Earth Remote Sensing Systems remains fit for purpose, particularly in balancing national security objectives with the scientific, commercial, and humanitarian benefits of commercial Earth observation. More broadly, shutter control policy should be revisited to ensure that measures are in place to preserve its legitimate security function while preventing its overapplication from inadvertently undermining the ability of the international community to independently monitor conflicts occurring on allied territory. Kelevitch recommends mandatory notification of allied

governments whose territories fall within restricted imaging areas, alongside a protocol requiring states to specify the geographic scope and duration of an image hold, allowing affected states to identify alternative sources of imagery where necessary.

In sum, as in other domains of verification, political will here is also significantly lagging behind technological possibilities, particularly when combining the latest sensor technology with AI. While these techniques could prove valuable for future arms control agreements, adversaries might perceive them as intrusive surveillance, potentially increasing the risk of attacks on satellite infrastructure or even on satellites themselves. So, what are the political consequences of these latest developments?

States should invest in AI-enabled intelligence.

When weighing the risks against the opportunities, the prospect of gaining significantly enhanced intelligence clearly outweighs the risks, particularly for emerging space powers such as Germany. States should therefore invest in AI-enabled intelligence. Rather than relying on collecting imagery alone, the strategic advantage increasingly depends on the integration and analysis of multisensor data. Policymakers should thus prioritize AI, data fusion, and analytical capacity. States should, however, bear in mind that non-state actors will also have increased AI capabilities. In light of this, it is important to maintain rigorous technology assessment. Governments should systematically evaluate claims of novel sensing technologies, allocating resources to research capable of distinguishing credible advances from strategic signaling, thereby avoiding technology races driven by the fear of being exposed.

Germany should reassess whether its National Data Security Policy for Space-Based Earth Remote Sensing Systems remains fit for purpose in an increasingly international commercial Earth observation market.

Sarah McKee

PRIF – Peace Research Institute Frankfurt

Niklas Schörnig

PRIF – Peace Research Institute Frankfurt

1 Aftergood, S. (2019, November 1). *"National technical means" leaves the lexicon.* Federation of American Scientists. https://fas.org/publication/ntm-obe/

2 A notable breach of this protocol occurred in 2019, when President Trump tweeted a mobile phone shot of a high-resolution briefing photo showing a failed Iranian missile launch. Deviating from the standard practice of publishing "degraded" versions for public release, the image's extreme clarity allowed analysts to calculate the sub-10cm resolution of the USA 224 satellite, effectively revealing classified performance data for the U.S. KH-11 reconnaissance fleet.

3 Grossfeld, E. (2025). Russia's declining satellite reconnaissance capabilities and its implications for security and international stability. *International Journal of Intelligence and CounterIntelligence, 38*(1), 1–30. https://doi.org/10.1080/08850607.2024.2330848

4 Strohmeyer, M. D. (2022). *DAF fellows published portfolio: Commercial data revolution for national security*. Air University. Starling, C. G., Wetzel, T. K., & Trotti, C. S. (2021). *Seizing the advantage: A vision for the next U.S. national defense strategy*. AF Fellows, Atlantic Council. https://www.atlanticcouncil.org/content-series/atlantic-council-strategy-paper-series/seizing-the-advantage-a-vision-for-the-next-us-national-defense-strategy/

5 Ginter, K. (2007). *Space technology and network centric warfare: a strategic paradox*. Strategy Research Paper. https://apps.dtic.mil/sti/pdfs/ADA469763.pdf

6 Minculete, G., & Păstae, V. (2024). Characteristic elements of combat drones in military actions: Unmanned aerial vehicles procured by Romanian armed forces. *Land Forces Academy Review, 29*(2), 274–289. https://doi.org/10.2478/raft-2024-0030

7 Chen, S. (2024, September 13). *Starlink radiation makes stealth target glow on Chinese radar*. South China Morning Post. https://www.scmp.com/news/china/science/article/3278209/starlink-radiation-makes-stealth-target-glow-chinese-radar

8 McKee, S. R., Cilliers, P. J., Lotz, S., & Monstein, C. (2023). The effects of solar radio bursts on frequency bands utilised by the aviation industry in Sub-Saharan Africa. *Journal of Space Weather and Space Climate, 13*, 4. https://doi.org/10.1051/swsc/2023001

9 Xue, D., Zhang, Q.-H., Feng, X., & Wang, C. (2026). Space weather and low-altitude drone economy. *Space Weather, 24*(1), e2025SW004803. https://doi.org/10.1029/2025SW004803

10 Malik, A., & Rao, M. (2025). Radio frequency transformer and reinforcement learning-based adaptive interference mitigation for aviation communication systems. In *2025 35th Irish Signals and Systems Conference (ISSC)* (pp. 1–6). IEEE. https://doi.

org/10.1109/ISSC67739.2025.11291303

11 Parsons, N. (2024, July 22). *Space data use in multi-domain operations*. BAE Systems Digital Intelligence. https://www.baesystems.com/en-uk/insight/space-data-use-in-multi-domain-operations

12 Kelevitch (2026, August 26). *Shutter control worked. That should worry the imagery industry more than anyone.* SpaceNews. https://spacenews.com/shutter-control-worked-that-should-worry-the-imagery-industry-more-than-anyone/

Technical Background: Taking Images in Space

Optical Imaging

Optical imaging is the oldest and most widely used form of satellite remote sensing. Optical satellites function as highly sophisticated digital cameras, capturing sunlight reflected from the Earth's surface in the visible and near-infrared portions of the electromagnetic spectrum. Optical sensors are passive instruments that rely entirely on natural illumination, making image acquisition dependent on daylight and atmospheric conditions. Given that optical sensors depend on reflected sunlight, clouds, fog, smoke, and airborne pollutants can obscure the Earth's surface. These sensors cannot acquire conventional images at night without artificial illumination.

Optical imagery provides highly intuitive, photograph-like representations of the Earth's surface and is widely used for mapping, environmental monitoring, disaster response, urban planning, and military reconnaissance. Modern optical sensors can distinguish fine spatial details, allowing analysts to identify buildings, vehicles, infrastructure, and changes in land use. Commercial satellites currently achieve resolutions of approximately 30 cm per pixel, while advanced military systems are generally estimated to achieve around 10 cm or better. Although these resolutions enable the identification of many technical objects, they remain insufficient to reliably read small text such as vehicle license plates under normal operating conditions.

Once the photo is taken, it needs to be compressed and transmitted. Since the bandwidth of a satellite is usually limited, a balance must be struck between accuracy (level of detail) and robustness (error resistance).[1] Accuracy describes how much detail the image shows. The stronger the compression, the smaller the file size, but the greater the loss of detail. High robustness guarantees that the file reaches Earth with as few errors and as readable as possible. The higher the robustness, the larger the file, because redundancies ensure that the image is preserved even in the event of data loss. In sum, image quality depends on both the available budget and the desired speed of image acquisition.

Infrared (IR)

Infrared imaging is based on the principle that warmer objects emit more infrared radiation than cooler ones. As a result, images showing the distribution of infrared radiation can be used to identify sources of heat and, for example, determine whether technical structures are in use. Unlike visible-light imaging, infrared measurements detect thermal radiation rather than reflected sunlight, enabling satellites to collect data both during the day and at night. Satellite infrared measurements are widely used to estimate surface and cloud temperatures, monitor weather systems, detect wildfires and volcanic activity, and study changes in the Earth's climate. By measuring infrared radiation at different wavelengths, satellites can also provide information about atmospheric temperature, humidity, and the distribution of gases and aerosols.

However, satellite-based infrared imaging has important limitations. Water vapor and carbon dioxide in the atmosphere absorb infrared radiation (the greenhouse effect), reducing the amount of radiation reaching the sensor. In addition, infrared sensors must be carefully shielded from the satellite's own thermal emissions, which can otherwise cause image noise and degrade measurement quality. In military applications, this challenge is often addressed using sophisticated cooling systems, such as those based on liquid helium. For space applications, minimizing image noise typically requires balancing sensor performance against constraints on satellite mass, power consumption, and cost.

SAR

Synthetic aperture radar (SAR) operates at microwave wavelengths, which are significantly longer than infrared radiation. Unlike optical and infrared sensors, SAR is an active sensing technology: the satellite transmits radar pulses toward the Earth's surface and records the reflected signals. Because microwave radiation penetrates clouds, rain, smoke, and, to some extent, vegetation, SAR can produce images regardless of weather conditions or time of day, making it one of the most reliable remote sensing technologies for continuous observation.

The principal disadvantage of radar is that longer wavelengths require much larger antennas than optical systems to achieve comparable spatial resolution. Synthetic aperture radar overcomes this limitation by exploiting the satellite's motion. Instead, rather than relying on the physical size of its antenna, the satellite transmits thousands of radar pulses as it travels along its orbit, using a side-looking geometry. By coherently combining the echoes received from successive positions, the system synthesizes a much larger effective antenna, or "synthetic aperture," than is physically present. This processing enables relatively compact satellites to produce high-resolution imagery, but it also requires them to move past the target rather than remaining stationary over a single location.

Modern SAR systems are capable of resolving remarkably fine details. Military satellites are generally estimated to achieve resolutions of around 10 cm or better, sufficient to identify features such as the configuration of vehicles or whether doors or hatches are open. Nevertheless, SAR imagery differs fundamentally from optical photography. Radar images represent the strength of reflected microwave signals rather than visible appearance, meaning that features such as painted markings or license plates typically cannot be distinguished even at very high spatial resolution.

Synthetic aperture radar also generates extremely large volumes of data. Since the satellite transmits and records thousands of radar pulses for every image, onboard processing and data compression are essential before transmission to ground stations. Artificial intelligence is increasingly being employed to prioritize, filter, and compress imagery, thereby helping to manage limited communication bandwidth, although this presents many of the same challenges associated with the onboard processing of optical imagery. Once almost exclusively a military capability due to its technological complexity and cost, SAR has become increasingly accessible through commercial providers. Advances in commercial Earth observation now allow sub-meter measurement accuracy. Optical satellites, such as Maxar's WorldView Legion constellation, can resolve objects as small as 30 cm, while SAR provider Umbra offers market-leading resolution from 25 cm to 1 m. Its imagery products are available under a Creative Commons license at highly attractive prices, and the company has partnered with European Space Imaging (EUSI) to deliver high-resolution, all-weather solutions for emergency and security applications.

1 Liu, Y., Wang, P., & Wang, S. (2025). Robustness plugin for military target recognition algorithms in on-orbit satellite environments. *Journal of Intelligent Computing and Networking, 1*(1), 54–62. https://doi.org/10.64509/jicn.11.24, p. 54.

Nuclear Physics

Novel Particle Probes in Nuclear Verification

In the context of the global resurgence of nuclear energy, innovation and expansion pose a challenge to the implementation of current safeguards. Research and academic institutions are therefore exploring new nuclear verification techniques, with non-destructive assay (NDA) approaches based on novel particle probes proving particularly attractive. This chapter presents some of the most promising techniques, following an overview of the main existing approaches. The chapter also discusses the dependence of these techniques on novel industrial technologies and their integration in potential verification scenarios. This analysis of the evolution of frontier verification techniques enables us to identify common technological drivers and approaches that can sustain their development.

Small modular reactor (SMR): A class of nuclear fission reactor with an electrical output of less than 300 MW, based on modular design principles to enable faster and lower cost construction, as well as more flexible and scalable deployment across a range of applications, including general electricity production, facility power, and industrial heat production.

Trends in the nuclear industry and the growing legacy of spent nuclear fuel underscore the need for long-term investment in tools for verifying nuclear material. See also the Advanced Reactors chapter in the CNTR Monitor 2025.

Since radioactivity is a defining feature of nuclear materials, the verification of nuclear inventory and facilities often relies on radiation detection devices. Simultaneously, radiation protection requires these materials to be biologically shielded to ensure safe operations, making reliable measurements—and hence verification—more challenging. In addition to the dangers that radioactive materials pose to living organisms and the environment, they can also be used for the—prohibited—construction of nuclear weapons, ranging from dirty bombs to nuclear warheads. Consequently, the whereabouts of nuclear material must be meticulously documented and verified. This is partially due to the need for long-term management of increasing quantities of spent nuclear fuel (SNF) in the absence of a clear pathway toward its final disposal. Simultaneously, nuclear energy is experiencing a global resurgence as a potential means of reconciling demands for energy sovereignty and low-carbon electricity. New nuclear energy programs have been launched by several countries, including Poland,[1] Türkiye,[2] and Egypt,[3] while existing programs are being expanded or revived following earlier phase-out decisions. One major development is the emergence of new deployment concepts,[4] including fleets of small modular reactors (SMRs) and microreactors for decentralized, industrial, transportable, or mobile nuclear power—potentially even for the maritime propulsion of cargo vessels. In addition to this, new novel advanced reactors (NARs or Gen IV reactors) using different fuel forms and coolants are also being developed, particularly by nuclear power start-ups in the U.S.[5] These new concepts challenge existing safeguards measures designed to ensure non-proliferation: larger, more widely dispersed fleets of reactors make inspection logistics more challenging, longer refueling cycles or sealed cores prevent access to fuel elements, and new coolants that cannot be detected by existing measurement techniques further hinder verification efforts. Accommodating these new concepts, designs, and deployment approaches would require major modifications to current safeguards methods as well as additional personnel. New nuclear verification techniques are therefore needed to

ensure that current safeguards remain effective, and research and academic institutions are investigating such approaches. These efforts depend on improvements in signal detection and data-processing techniques resulting from ongoing R&D. Non-destructive assay (NDA) approaches for the passive or active detection of nuclear materials are of particular interest because they offer nonintrusive procedures that can minimize the burden and risks to operators. Therefore, the development of novel techniques capable of enhancing existing surveillance systems through nondestructive, continuous, or remote monitoring is particularly promising in terms of future verification tools for nuclear facilities and material.

Non-destructive assay (NDA): NDA techniques encompass a range of methods and instruments that enable the on-site characterization of nuclear materials during inspections, often without the need to take samples. Wherever possible, NDA is preferred because it eliminates the need for international transport, prevents secondary waste during sample disposal, and often produces at least preliminary results during the inspection itself.

Current Probes: Existing Safeguards Tools

Established safeguards rely on the verification of operator-declared nuclear material accountancy. This is achieved through inspections that combine independent qualitative and quantitative measurements to verify the presence, composition, and quantity of nuclear material with containment and surveillance (C/S) techniques to maintain an unbroken continuity of knowledge (CoK) between inspections. The radioactive emissions of the material, usually in the form of neutrons, α-, β-, or γ-rays, are used by passive detectors to characterize the inventory during verification. As nuclear materials also interact with particles in characteristic, isotope-specific ways, active detection methods exploit their responses to different types of radiation to gather even more information. Inspectors often rely on NDA methods to perform measurements in the field without needing to take samples. Current methods employed for NDA rely primarily, but not exclusively, on the detection of three main probes: γ-rays, neutrons, and Cherenkov radiation.

- **γ-rays** are emitted by most nuclear material, and their energy spectra contain characteristic information about the sample. γ-ray spectrometry determines the presence of specific isotopes. Depending on their energy, they can penetrate lower density materials, allowing limited measurements to be made through thin shielding. However, dense materials—including nuclear materials such as uranium or plutonium—can block a significant fraction of these emissions, as the material's high density effectively shields against its own radiation. High γ-ray emission levels from the fission products of these two elements make it even more difficult to measure their quantities directly using traditional γ-ray spectrometry.

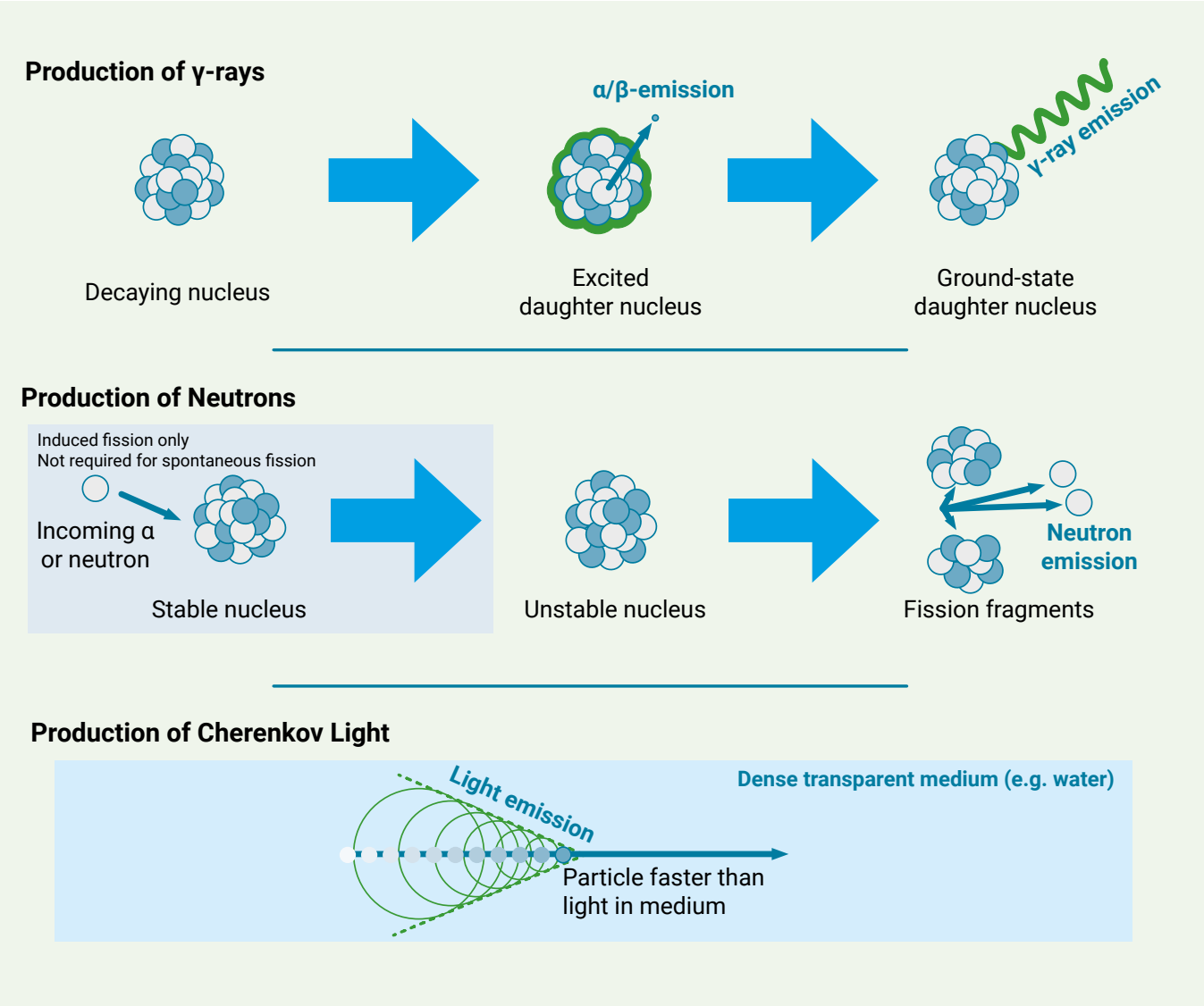
Nondestructive verification techniques for nuclear material often rely on measuring radiation emitted by the materials themselves, especially γ-rays and neutrons.

- **Neutrons** are emitted primarily through spontaneous or induced fission and α-ray-induced reactions. Coincidence and multiplicity counting techniques can be used passively or actively to determine the mass of fissile material. Neutrons are effectively shielded only by materials with a high neutron-capture cross-section, such as water or borated polyethylene, making measurements possible through certain types of shielding.

- **Cherenkov radiation** (photons) is emitted when β-rays, produced directly or indirectly by spent nuclear fuel, travel faster than the speed of light in water in storage ponds—the optical equivalent of a sonic boom. By detecting this UV light, the presence of spent fuel in the assemblies can be verified from above the ponds. Many materials—including water—are transparent to Cherenkov light.

Recent Developments: New Instruments and Approaches

While the potential increase in the number of nuclear facilities resulting from the deployment of SMRs and NARs was first anticipated in the 2010s, with more concrete



The physical processes underlying the main signals used for current nuclear verification techniques: γ-ray emission from nuclear decay, neutron emission from spontaneous or induced fission (the latter requiring an incoming α-particle or neutron), and Cherenkov light from fast charged particles.

plans emerging in the 2020s, the challenges posed by nuclear waste have long been. To address these challenges, new technologies such as PGET scanners (see Table 1) are currently being developed. Other approaches under consideration include enhancing existing methods, such as remote detector control and data transfer, and modifying algorithms to address reactor remoteness and new fuel compositions, as well as using new particle types:

Passive neutron albedo reactivity (PNAR) scanners: This technique is a modifies existing neutron detection techniques by using the neutrons emitted by the nuclear material itself to characterize its response to neutrons and reconstruct a tomographic image of spent fuel rods.[6] While not yet in use, there are plans to combine PNAR and PGET into a single device capable of producing highly detailed characterizations of fuel rods prior to their encapsulation in long-term storage casks.

Muography: Cosmic-ray muography uses high-energy particles in the Earth's upper atmosphere to produce density maps of the interior of objects or structures, resulting in X-ray-like imaging. As a naturally occurring form of highly penetrating radiation, it enables imaging without additional radiation exposure. There are two muography

approaches: 1) absorption radiography using a single detector positioned below the object of interest to produce 2D images and 2) scattering tomography that enables a 3D reconstruction but requires detectors to be positioned both above and below the interrogated object.[7] Muography excels at imaging and positioning within large static or even geological structures[8] (e.g., underground mine shafts) and smaller objects composed of materials with different densities, such as spent-fuel casks containing nuclear material.[9] However, deployment scenarios are limited by the incoming muon flux, angular distribution, and possible solid-angle coverage of a detector station. Muography is thus already being used on a small scale to image nuclear waste containers, although the process currently requires days to weeks. Future applications are imaging of nuclear facilities, such as reactors or disposal facilities.

Particle Probe	*Example Measurement Devices*	*Typical Applications*
γ-rays	• **HM-6 field spectrometer:** hand-held digital gamma-ray spectrometer using CZT semi-conductor detector. • **Electrically cooled germanium spectrometer (ECGS):** portable high-resolution spectrometer based on a high-purity germanium detector. • **Passive gamma emission tomography (PGET) scanner:** tomographic imaging system composed of CZT detectors.	• Dose-rate measurement • Source search • Isotope identification • Enrichment-level determination & Pu/U attribute verification • Spent fuel verification • Partial defect detection
Neutrons	• **Active well coincidence counter (AWCC):** transportable detector composed of ^{3}He tube counters circularly arranged and embedded in polyethylene. Embedded AmLi sources provide a uniform distribution of interrogating neutrons. • **High-efficiency neutron counter (HENC):** large-scale neutron coincidence and multiplicity counter based on ^{3}He tubes. • **Uranium neutron coincidence collar (UNCL):** cart-mounted compact polyethylene collar containing ^{3}He tubes and an AmLi source.	• Verification of ^{235}U in highly enriched samples • Plutonium quantification inside waste drums • Determination of ^{235}U content in fresh (unirradiated) fuel assemblies
Cherenkov radiation	• **Robotized Cherenkov viewing device (RCVD):** autonomous floating detector equipped with a downward-pointing UV camera that provides real-time measurements of Cherenkov radiation.	• Analysis of individual spent fuel assemblies in storage ponds

Table 1: Overview of particle emissions used for verification measurements and typical associated applications.[10]

Novel Particle Probes: Frontier Developments for 2030 and Beyond

Research and academic institutions are uniquely positioned to assess the feasibility of employing novel particle probes for nuclear non-proliferation and to develop initial concepts for new verification techniques. Four of the most promising novel particle probes currently under investigation are discussed below.

CZT detector: Semiconductor radiation sensor that uses cadmium zinc telluride to directly convert radiation particles into electrical signals at room temperature. It is primarily used for the detection of gamma rays and X-rays.

Antineutrino monitoring: Antineutrinos are subatomic particles emitted primarily during radioactive decay, particularly of fission products in reactors, making them a clear signature of nuclear facilities. They also pass through all materials unhindered, making them impossible to shield or falsify—but also very hard to detect. Antineutrino detectors have advanced significantly over the last few decades. Potential applications include serving as safeguards for SMRs and advanced reactors by remotely monitoring the power level and isotopic inventory of operating reactors without the need for on-site inspections; verifying the absence of operating reactors within a certain area; and conducting nondestructive assays of interim storage casks.[11] Several applied research projects are underway to develop demonstrator devices,[12] while basic research is exploring new detection approaches with higher sensitivity.[13]

Ongoing research is investigating techniques that enable measurements through the thick radiation shielding often required for nuclear material, such as reactor vessels or storage casks.

Neutron beam interrogation: Neutron beams are being proposed to probe the structure of nuclear material by measuring their attenuation with spatially resolving detectors to create neutron radiographs. The energy-resolved detection of neutron transmission through a sample also enables the spatial reconstruction of its composition owing to the unique attenuation features of individual isotopes. Several approaches to this technique are being developed, with the most ambitious exploring the feasibility of obtaining tomographic images showing the 3D distribution of different isotopes.[14] Potential applications include characterizing nuclear material, such as fuel, within heavily shielded enclosures, such as storage casks.

Optical/laser spectroscopy: There has also been growing interest in the use of laser-based optical probes as an alternative to particle probes employing radioactive beams. Lasers can excite the electronic energy levels in samples, which subsequently relax through the emission of photon signals. This allows for nondestructive, real-time, in situ analysis of trace quantities of nuclear material, while requiring significantly less sample preparation for laser ionization spectroscopy.[15] This has also enabled the use of Light Detection and Ranging (LiDAR) imaging instead of Radio Detection and Ranging (RADAR) imaging, offering greater detail and accuracy, particularly for the kind of short-range static targets typically encountered by nuclear inspectors.[16] One possible verification application is an NDA tool allowing identification of trace amounts of nuclear material in the field.

Nuclear resonance fluorescence (NRF): NRF is a process in which a nucleus absorbs and resonantly reemits a γ-ray. The energy of the emitted γ-ray depends on the nuclear structure and can therefore be used as a signature to identify specific isotopes. A γ-ray excitation beam, typically with an energy larger than 1 MeV, can thus be used to determine the composition of a sample or, for verification purposes, to directly measure the isotopic quantity of fissile material in a nonfissile environment. The γ-ray energies involved are high enough for them to travel through thick shielding materials, making active NRF probing a promising alternative to passive techniques that are more sensitive to background and shielding. This enables the characterization of nuclear material in heavily shielded, opaque environments, such as NARs using non-water coolants or SNF stored in shielded casks.[17]

Integration of Novel Verification Techniques

If integrated into nuclear verification procedures, novel particle probes could contribute to the overarching objective of obtaining more information more efficiently. Their remote operation and automatization would reduce inspector burden by requiring less supervision time and fewer personnel, while also reducing radiation exposure and at the same time enabling more continuous and informative measurements. Specific challenges arising from the potential large-scale deployment of SMRs and microre-

actors could also be addressed with novel particle probes. Operations such as the refueling of remotely located small reactor arrays could be monitored using in situ antineutrino detectors with data-transmission capabilities. Fuel-loaded reactors being transported on land or water could also be inspected at specified locations, such as production facilities or designated ports, using passive methods such as antineutrino monitoring or active ones such as NRF or neutron-beam interrogation. Active methods, particularly NRF, are expected to be well suited to inspecting sealed reactor designs with high levels of radiation opacity, which may, in some cases, be further increased by high-density coolant or fuel materials. This approach is currently constrained by the high cost and large footprint of the necessary equipment and is therefore expected, in the near term, to be deployed only at specific locations, such as reactor production facilities or spent-fuel storage sites. With further advances in laser, accelerator, and detector technology, however, these methods could conceivably become portable in the longer term. Bilateral tailoring of future reactors and verification techniques, as encouraged by the concept of safeguards by design, should also play a central role in the ongoing development of methods based on novel particle probes.

Technological Drivers and Sustained Innovation

Beyond their intended use for nuclear safeguards, the recent developments and projected future technologies examined in this chapter share several common innovation drivers:

On the technological side, one major milestone has been the upgrade from photomultiplier tubes (PMTs), first developed in 1930s, to silicon photomultipliers (SiPMs). The widespread adoption of SiPMs, first in research,[18] then for applications,[19] has overcome the drawbacks of PMTs while leveraging the massive economies of scale of the modern semiconductor industry, since SiPMs are produced in the same foundries used to produce chips. Verification scenarios typically present low-signal situations where SiPMs provide superior performance owing to their improved spatial and timing resolutions, higher photoefficiency, and immunity to magnetic-field interference, albeit at the cost of a slightly lower energy resolution. Additionally, their smaller size, lower operating voltages (< 1 kV), reduced power consumption, physical ruggedness, and tolerance to vibrations make SiPM-based detectors more portable, adaptable and reliable for on-site use. Almost all the detectors listed here have benefited directly from the maturation of this technology.

Advances in scintillator development have driven a shift from liquid to plastic scintillators, which offer more durability, lower toxicity and flammability, increased portability and scalability, and the potential for bespoke geometries. These properties are all highly desirable for detectors intended for deployment across a range of field verification environments.[20] Handheld detectors could benefit from this in particular.

The adoption of transition-edge sensors has also enabled developments in microcalorimetry, which identifies tiny shifts in temperature that are indirectly produced when a particle is detected. This could potentially enable the identification of rare decays that would be obscured by traditional detectors, which measure light flashes or electrical currents directly. This is particularly advantageous for nuclear forensics or in scenarios requiring the analysis of complex mixtures of nuclear material.

Significant progress in robotics tools and applications has also facilitated the safer deployment of nuclear monitors, with growing potential for remote operation, as demonstrated by the RCVD. Continuous advances in signal processing, mainly driven by communication technology, has resulted in a variety of viable off-the-shelf compo-

Photomultiplier tube (PMT): Vacuum tube containing a photocathode that emits electrons upon exposure to electromagnetic radiation such as UV or visible light. A dynode structure under high voltage (1,000+ V) multiplies these electrons, which are then measured at an anode, resulting in an electric charge signal.

Silicon photomultiplier (SiPM): Solid-state semiconductor device comprising many pixels of avalanche diodes, capable of producing an electrical signal in response to single light photons. These devices are extremely compact, with more than 1,000 pixels per mm2, and operate at low voltages ranging from 20 to 100 V.

Foundry: Shorthand for semiconductor fabrication plant. This refers to factories that use photolithography to create integrated circuits from silicon wafers. These devices are commonly used as cutting-edge microprocessors (including AI chips), solid-state memory, photodetectors, embedded processors, or microcontrollers.

Scintillator: Material that emits light when exposed to ionizing radiation, either through direct ionization, as in the case of alpha or beta rays, or through secondary emissions induced by gamma rays or neutrons. This light emission can then be detected by light sensors, which usually provide a signal proportional to the ionizing energy deposited in the scintillator.

nents, reducing the need for expensive custom application-specific integrated circuits for fast processing. Advances in data-handling technologies, including artificial intelligence (AI) and machine learning (ML)-assisted methods, enable the analysis of more complex and noisy datasets with many measurement channels, while simultaneously improving the efficiency of both data transfer and storage. Muon tomography, PGET, and PNAR have particularly benefited from these advances, leading to high-fidelity image reconstruction of measured objects.

Modern signal processing and an increase in computing resources have enabled tomographic and multimodal imaging.

With regard to innovations in techniques, one major milestone has been the incorporation of multimodal imaging. By integrating sensors specifically tailored to different particle probes, the resulting system becomes sensitive to a broader range of radioactive by-products, as envisaged in the planned combination of PGET and PNAR. Nevertheless, there is still scope for stronger integration of complementary techniques in this area.

Adapting widely used industrial technologies to physics-related detection applications has enabled this wave of new detector advancements. Beyond general collaboration between industry and research, medical physics stands out as one of the major drivers.[21] Here, imaging and detection technologies demonstrate high societal impact, strong R&D, and continued training opportunities. At the same time, the application of technological advances with high industrial R&D investment in areas not served by these industries created a fertile environment for maturation. To sustain innovation for verification purposes, supporting interdisciplinary environments that foster maturation is likely to be the most effective way of leveraging existing scientific and R&D efforts. Explicitly funding collaborations and projects at the intersection of security applications and high-investment R&D could create similar maturation dynamics to those seen in medical physics. While there has already been an increase in security- and dual-use-focused R&D programs, it is important to ensure that nuclear verification is included in these efforts. At the same time, the civil security aspect must be maintained. Since verification relies on trust and transparency, new trusted verification technology results are unlikely unless these security research initiatives are also firmly anchored in civilian and public research.

Luis Pazos Clemens

PRIF – Peace Research Institute Frankfurt/
Technical University of Darmstadt

Adhitya Sekhar

Technical University of Darmstadt

Yan-Jie Schnellbach

PRIF – Peace Research Institute Frankfurt/
Technical University of Darmstadt

1 OECD Nuclear Energy Agency. (2025). *Advancing Poland's nuclear energy programme.* https://www.oecd-nea.org/jcms/pl_105146/advancing-poland-s-nuclear-energy-programme

2 World Nuclear News. (2017). *Akkuyu construction formally starts.* https://www.world-nuclear-news.org/Articles/Akkuyu-construction-formally-starts

3 World Nuclear Association. (2025). *Nuclear power in Egypt.* https://world-nuclear.org/information-library/country-profiles/countries-a-f/egypt

4 Schnellbach, Y. (2025). Advanced reactors and international security. In Göttsche, M., Reis, K., & Daase, C. (Eds.). (2025). *New Realities of AI in Global Security. CNTR Monitor – Technology and Arms Control 2025* (pp. 57–61). PRIF – Peace Research Institute Frankfurt. https://doi.org/10.48809/cntr2025en

5 U.S. Department of Energy, Office of Nuclear Energy. (2025). *X-energy starts building construction for advanced nuclear fuel facility.* https://www.energy.gov/ne/articles/x-energy-starts-building-construction-advanced-nuclear-fuel-facility

U.S. Department of Energy, Office of Nuclear Energy. (2026). *NRC issues construction permit for TerraPower's Natrium advanced reactor*. https://www.energy.gov/ne/articles/nrc-issues-construction-permit-terrapowers-natrium-advanced-reactor
Oklo Inc. (2025). *Oklo breaks ground on first Aurora Powerhouse.* https://oklo.com/newsroom/news-details/2025/Oklo-Breaks-Ground-on-First-Aurora-Powerhouse/default.aspx
Oak Ridge National Laboratory. (2026). *ORNL, Kairos Power partner to advance deployment of next-gen nuclear energy.* https://www.ornl.gov/news/ornl-kairos-power-partner-advance-deployment-next-gen-nuclear-energy

6 Tupasela, T., et al. (2021). Passive neutron albedo reactivity measurements of spent nuclear fuel. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 986*, Article 164707. https://doi.org/10.1016/j.nima.2020.164707

7 Kaiser, R. (2018). Muography: Overview and future directions. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 377(2137)*, 1–9. https://doi.org/10.1098/rsta.2018.0049
Tanaka, H. K. M., et al. (2023). Muography. *Nature Reviews Methods Primers, 3*, Article 88. https://doi.org/10.1038/s43586-023-00270-7

8 Tanaka, H. K. M. (2020). Muometric positioning system (μPS) with cosmic muons as a new underwater and underground positioning technique. *Scientific Reports, 10*, Article 18896. https://doi.org/10.1038/s41598-020-75843-7

9 Andreetto, P., et al. (2025). Results on re-verification tests of spent fuel casks with muon tomography: MUTOMCA project. *Journal of Applied Physics, 138*, Article 124903. https://doi.org/10.1063/5.0272975
Niederleithinger, E., et al. (2021). Muon tomography of the interior of a reinforced concrete block: First experimental proof of concept. *Journal of Nondestructive Evaluation, 40*, Article 65. https://doi.org/10.1007/s10921-021-00797-3

10 International Atomic Energy Agency. (2011). *Safeguards techniques and equipment* (International Nuclear Verification Series No. 1, Rev. 2). https://www.iaea.org

11 Akindele, O., et al. (2022). *Nu Tools: Exploring practical roles for neutrinos in nuclear energy and security* (PNNL-31870). https://doi.org/10.2172/1826602

12 Bowden, N. S. (2023). The Mobile Antineutrino Demonstrator Project. *Proceedings of the INMM Annual Conference.* https://resources.inmm.org/annual-meeting-proceedings/mobile-antineutrino-demonstrator-project

13 LiquidO Consortium. (2021). Neutrino physics with an opaque detector. *Communications Physics, 4*, Article 273. https://doi.org/10.1038/s42005-021-00763-5
Ackermann, N., et al. (2025). Direct observation of coherent elastic antineutrino-nucleus scattering. *Nature, 643*, 1229–1233. https://doi.org/10.1038/s41586-025-09322-2

14 Zimmer, M., Scheuren, S., Kleinschmidt, A., et al. (2022). Demonstration of non-destructive and isotope-sensitive material analysis using a short-pulsed laser-driven epi-thermal neutron source. *Nature Communications, 13*, Article 1173. https://doi.org/10.1038/s41467-022-28910-1
Wolfertz, A., Losko, A., Long, A. M., et al. (2024). Energy-resolved fast-neutron radiography using an event-mode neutron imaging detector. *Scientific Reports, 14*, Article 30487. https://doi.org/10.1038/s41598-024-61234-5

15 Kwapis, E. H., et al. (2024). Laser ablation plasmas and spectroscopy for nuclear applications. *Applied Spectroscopy, 78*(1). https://doi.org/10.1177/00037028231211559
Garrett, L. (2025). *Nanosecond laser-induced breakdown spectroscopy for novel nuclear diagnostics*. https://doi.org/10.7302/25562

16 Sanchez Belenguer, C., Wolfart, E., Bostrom, G., Sequeira, V., Velnom, L., et al. (2025). A decade of 3D laser scanning for design information verification: Lessons learned from Onkalo. *ESARDA Bulletin, 67*. https://doi.org/10.3011/ESARDA.IJNSNP.2025.4

17 Ludewigt, B. A., Quiter, B. J., & Ambers, S. D. (2011). *Nuclear resonance fluorescence for safeguards applications* (LBNL-4350E). Ernest Orlando Lawrence Berkeley National Laboratory.

18 Renker, D., & Lorenz, E. (2009). Advances in solid state photon detectors. *Journal of Instrumentation, 4*, Article P04004. https://doi.org/10.1088/1748-0221/4/04/P04004

19 Gundacker, S., & Heering, A. (2020). The silicon photomultiplier: Fundamentals and applications of a modern solid-state photon detector. *Physics in Medicine and Biology, 65*(17), Article 17TR04. https://doi.org/10.1088/1361-6560/ab7b2d

20 Hamel, M. (2021). *Plastic scintillators*. Springer International Publishing.

21 Demarteau, M., Lipton, R., Nicholson, H., & Shipsey, I. (2016). Particle and nuclear physics instrumentation and its broad connection. *Reviews of Modern Physics, 88*, Article 045007. https://doi.org/10.1103/RevModPhys.88.045007

Biotechnology

Genetic Engineering Attribution: Contribution to Verification and Practical Limitations

This chapter examines whether genetic engineering attribution (GEA) can meaningfully contribute to biosecurity and verification. It asks what current methods can actually infer from engineered biological material, how far results from curated datasets (e.g., plasmid repositories) can be applied to real-world incidents, and what this means for verification and compliance. The analysis suggests that sequence-based attribution may support screening, prioritization, and the narrowing of hypotheses, but that its value remains limited by incomplete data, uneven metadata, and the gap between repository benchmarks and realistic forensic settings. For this reason, the priorities are to improve data standards, evaluate attribution methods under more realistic conditions, and use sequence-based results only in combination with other forms of evidence.

Plasmid: A small, usually circular piece of DNA that can exist separately from a cell's main chromosome. In biotechnology, plasmids are commonly used to carry and introduce genetic material into cells.

Efforts to attribute the origin of biological agents are not new. Microbial forensics was developed to investigate biological incidents, identify the agent involved, where it came from, and who may have been responsible. While microbes have been used for harmful purposes for centuries, the 2001 anthrax letter attacks[1] in the United States gave particular urgency to the development of forensic capabilities to support investigation and attribution. These events highlighted the need for analytical tools and knowledge bases that could provide investigative leads, identify likely sources, and help establish responsibility with a high degree of scientific confidence. In practice, this meant combining genomic analysis with epidemiological information, environmental and clinical samples, and documented chains of custody.

This task remains highly relevant. Biological events can have major political and security consequences, yet attribution often remains difficult and contested. The prolonged disputes over the origin of COVID-19, including WHO-related reporting as late as July 2025, more than five years after the peak of the pandemic, illustrate how scientifically and politically contentious such questions can become.[2] Even where substantial data exist, uncertainty may persist for years, especially when key evidence is inaccessible, incomplete, or contested. Attribution, therefore, has direct implications for accountability, deterrence, and international stability.

In recent years, these questions have gained renewed significance in the context of modern biotechnology. As genetic engineering has become increasingly standardized, accessible, and data rich, attention has shifted toward whether engineered bi-

ological constructs contain detectable traces of their origin. This has given rise to growing interest in genetic engineering attribution, or GEA,[3] as a possible extension of microbial forensics into an era of plasmid design, synthetic biology, and large biological sequence repositories.

Why Attribution Matters

In this chapter, verification is understood in a broad sense: not as confirming stockpiles or declared facilities, but as supporting assessments of whether suspicious biological activity is consistent with prohibited development, use, or concealment. GEA contributes to this narrower evidentiary task by helping determine whether a construct appears engineered, what methods may have been used, and what further inquiries should follow.

Attribution offers several possible benefits for security and verification.[4] First, identifying the likely origin of a biological construct can help guide response efforts by clarifying which actor may be involved, their likely capabilities, and their motives. Second, by facilitating the identification of responsible parties, attribution can enable civil, criminal, or diplomatic consequences. Third, the realistic prospect of attribution may help deter misuse. In this sense, attribution has implications beyond the incident itself, since it may influence future behavior through the possibility of later scrutiny or consequences.

These points are particularly relevant in discussions about the verification of biological weapons obligations. Unlike some other arms control domains, biology offers relatively few opportunities for straightforward verification through inspection or material accounting alone.[5] Biological materials can be replicated, modified, concealed, and redistributed. This increases the importance of methods that can extract information from the agent or construct itself. In biotechnology, conclusions about origin therefore often rely on indirect, data-driven evidence, including sequence data and associated metadata, rather than on direct inspection alone.

Methodological Signatures and Sequence Clues

The central premise of GEA is that genetic engineering leaves traces. Design decisions are rarely random; they are shaped by training, available tools, routine laboratory practices, preferred methods, and access to particular equipment and materials. Taken together, these decisions may create a kind of "methodological signature" that can help indicate who was likely responsible for a given modification.[6]

In practice, such traces can take different forms. Some are more direct, such as the insertion of foreign genetic material, the use of recognizable laboratory components, or the presence of short sequence scars associated with particular cloning techniques. Others are more indirect, including codon usage patterns, backbone selection, promoter choice, guanine-cytosine (GC) composition, or other sequence-level regularities that reflect design habits rather than explicit markers. The attraction of GEA lies in the possibility that such patterns may be detectable at scale using statistical and machine-learning methods, enabling investigators to infer whether, and perhaps how, a biological construct was engineered.

Databases such as Addgene[7] have played a major role in advancing this line of work because they provide large, curated collections of plasmid sequences together with information about depositors, authors, backbone elements, and other construct features. Consequently, plasmid repositories became one of the main testbeds for GEA.

Plasmid design: The planned arrangement of genetic elements within a plasmid, such as inserted genes, promoters, and other functional components. Different design choices can reflect the purpose of an experiment and the methods used to build it.

Promoter: A DNA sequence that controls when and how strongly a gene is switched on. Promoter choice is an important part of genetic design.

Codon usage patterns: Different ways of writing the same genetic instructions in DNA. These patterns can sometimes provide clues about how a sequence was designed.

Backbone: The basic structural part of a plasmid that contains the elements needed for it to function and be maintained in a cell. Additional genetic material is inserted into this underlying framework.

Plasmid repository: A collection or database of plasmids that researchers can deposit, document, and often share with others. Repositories may also contain information about who submitted a plasmid, its sequence, its components, and related publications.

GEA is based on the idea that engineering can leave detectable traces in biological sequences, but current methods rely heavily on curated repository data.

Engineered biology is increasingly designed, stored, and shared in digital form, leaving behind sequence records, repository entries, and other traces of laboratory work in addition to physical material. These traces made attribution more tractable as a measurable computational problem. Researchers could then test whether recurring patterns in sequences and metadata were associated with particular contributors or engineering contexts. Early benchmark results suggested that such signals do exist, and that repository-based attribution is possible under structured conditions.[8] In practice, these studies indicated that likely lab-of-origin patterns could be recovered with considerable accuracy from curated plasmid datasets.[9] This framing also attracted attention beyond academia, including through programs such as DARPA's Bio Attribution Challenge, which approaches attribution as a large-scale computational problem using genomic and environmental datasets designed to simulate realistic scenarios.[10]

Repositories and Sequence Archives as Data Sources

Addgene: A nonprofit repository that collects, stores, and distributes plasmids contributed by research laboratories. Its records often include sequence information, depositor details, plasmid maps, and links to related publications.

Repositories and sequence archives support attribution in different ways. Plasmid repositories such as Addgene preserve laboratory-made constructs together with relatively rich contextual information, including depositor identity, plasmid maps, backbone descriptions, and links to publications. This makes them particularly useful for benchmarking attribution methods, as engineered sequences can be analyzed alongside their documented design context.

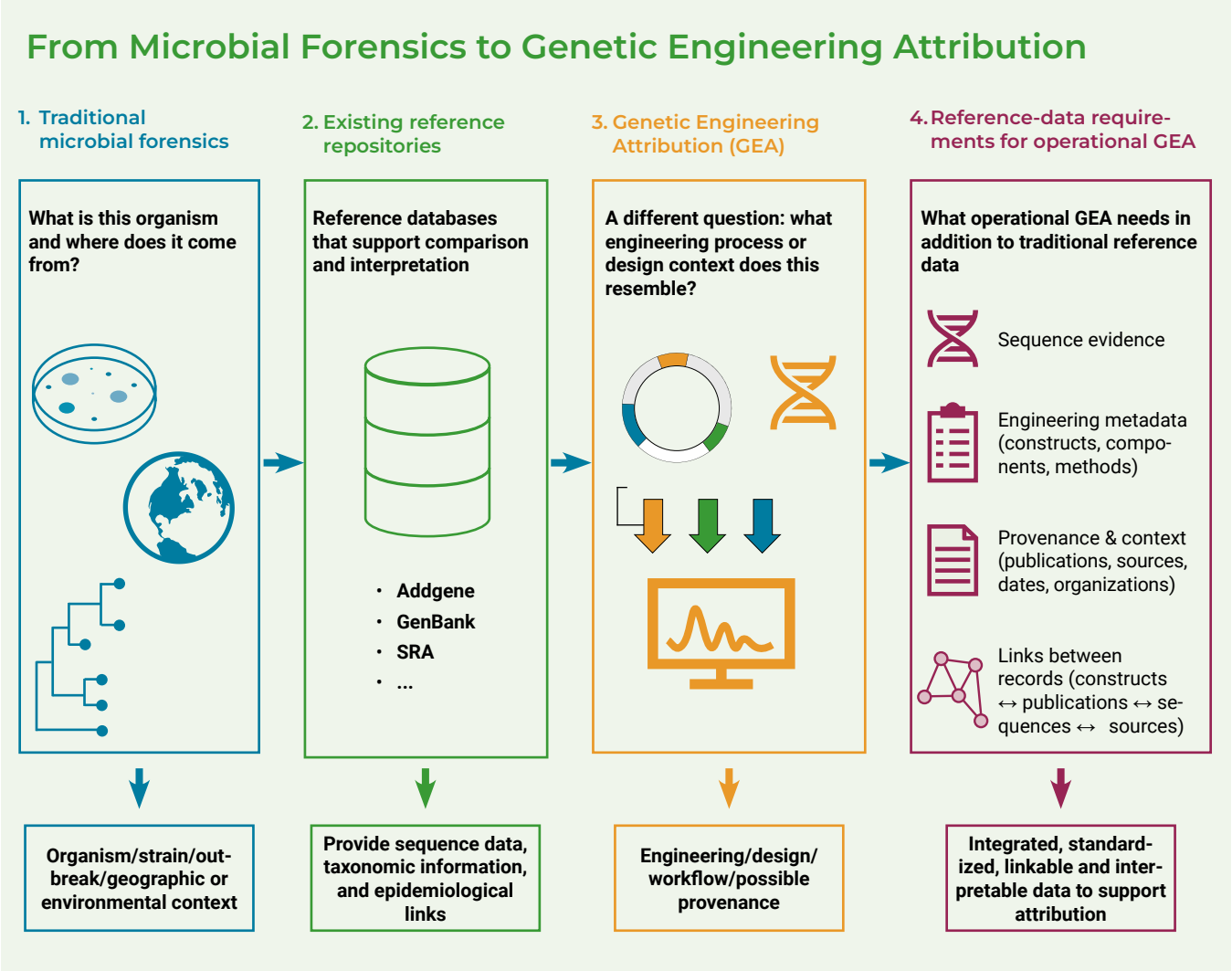
Broad-coverage sequence archives, such as GenBank, ENA, and DDBJ (all part of the International Nucleotide Sequence Database Collaboration),[11] serve a different function. These databases provide much broader range of biological sequences, including natural and engineered material, but often with less standardized information on whether and how a sequence has been engineered, as well as much weaker laboratory provenance. Raw-read repositories such as SRA add another important layer, because real-world samples may be fragmentary, mixed, or environmental rather than complete and well-annotated constructs. Metadata resources such as BioProject and BioSample can provide useful contextual information, but they do not, on their own, offer the structured engineering descriptions on which repository-based GEA often relies.

Different biological databases preserve different types of sequence and metadata information, which means they are suited to different attribution tasks.

Different databases make it possible to address different attribution questions, but each also imposes its own constraints. Addgene-like repositories are useful for determining whether attribution is possible under curated conditions. Resources such as GenBank, ENA, or SRA are useful for illustrating how difficult attribution becomes once sequences are heterogeneous, incomplete, and no longer embedded in a clearly documented engineering setting.

These repository-based results are promising, but they should be treated with caution. Public repositories provide a particularly favorable setting for attribution research, and success in these datasets does not mean that the same methods will perform reliably in real-world incidents. They contain (nearly) complete sequences, known labels, and relatively structured metadata, whereas real-world incidents are typically much messier. Samples may be incomplete, degraded, mixed with natural background material, or collected long after the relevant event. Metadata may be missing or inconsistently recorded, and investigators often do not begin with a clearly defined set of potential actors. In other words, repositories help demonstrate what is technically possible, but not what is operationally reliable.

A somewhat different question is whether engineered DNA can be distinguished from natural sequences at all. One example is the detection tool Synsor,[12] which uses dif-



ferences between natural plasmids and engineered vectors to classify sequences as either natural or engineered. This suggests that engineering-related signals can sometimes be detected, but remain dependent on the underlying dataset and they do not by themselves indicate who was responsible. It is one thing to infer that a sequence may have been engineered, another to identify the laboratory, actor, or process responsible. That said, a system that can flag suspicious sequences may still be useful, even if it cannot identify their origin. Such a tool could support screening, triage, and the narrowing of hypotheses, while helping investigators prioritize samples for follow-up analysis. This would still be a meaningful contribution, but it is different from claiming that sequence analysis alone can establish responsibility.

A more challenging question is whether such signals can go beyond merely indicating engineering to also help distinguish between possible sources. Some engineering practices may leave small traces in DNA, including differences in GC content (the proportion of guanine and cytosine bases in the sequence) and other broad patterns in sequence composition. Yet these differences are often limited in their discriminatory value. A preliminary look at GC-related variation suggests that simple compositional features can reflect some variation between groups of constructs, but this is not strong enough for reliable source attribution. In many cases, the variation within the sequences associated with a single depositor is of a similar magnitude to the variation observed between different depositors. This makes reliable attribution difficult when relying on simple sequence features alone.

Engineered vector: A DNA molecule that has been deliberately modified to carry genetic material into a cell. Engineered vectors are used as tools in genetic engineering and can include plasmids and other delivery systems.

A further limitation concerns the structure of the available data. Repositories may contain large numbers of plasmids, but these are not evenly distributed across contributors. Many laboratories are represented by only a handful of sequences, while a smaller number account for a substantial proportion of the total dataset. This unevenness makes robust inference difficult, and raises a broader question of how far repository-based results can be applied to the threat scenarios that matter most for biosecurity and verification.

Promises and Limits of GEA

These findings suggest that GEA remains promising, but also constrained. Biological constructs do appear to contain information about their design history. Even with no explicit watermark or identifier, traces of laboratory practice may remain visible in the sequence. This means that attribution is not purely speculative. Some forms of sequence-based inference are possible, and they may become more useful as datasets and analytical methods improve.

GEA can support screening and the narrowing of hypotheses, but its signals are often weak, overlapping, and insufficient on their own for strong attribution claims.

The limitations are equally important. In practice, attribution-relevant signals are often weak, overlapping, and highly context dependent. Any meaningful inference is therefore likely to require a combination of multiple partial indicators, potentially aided by machine learning or other statistical methods. Even then, the outcome is more likely to be probabilistic than definitive. This is particularly important in forensic or verification contexts, where the question is whether the pattern carries enough evidentiary weight to support a meaningful judgment. Standardized components reduce variability, sparse datasets limit confidence, and deliberate obfuscation remains possible. A malicious actor could, in principle, adopt common design practices, reuse widely shared elements, or intentionally imitate characteristics associated with another laboratory or country. This creates not only technical uncertainty but also political risk, since attribution claims can carry serious diplomatic consequences.

For this reason, GEA should not be seen as a standalone solution to biotechnology verification.[13] Its role is better understood as one source of evidence among several, useful for generating leads, narrowing possibilities, and supporting broader assessments. In combination with other measures, such as DNA synthesis screening, documentation requirements, epidemiological analysis, environmental sampling, or intentional traceability mechanisms, it could make a meaningful contribution. On its own, however, it is unlikely to provide the level of certainty that strong attribution claims require. A detectable signal is not the same as usable evidence.

Verification and Compliance

For a short introduction to the BWC, see p. 127 f.

Although GEA methodologies remain at an early stage, they are still relevant to discussions on verification and compliance under the Biological Weapons Convention. Their main value lies not in providing definitive proof, but in strengthening the technical basis for assessing suspicious biological activity. In a field where verification has long been considered challenging, even partial attribution capabilities may prove valuable.

In practical terms, GEA would serve as a trigger for further inquiry rather than a final judgment. A sequence-based finding might indicate that a sample is probably engineered, that a particular vector or design pattern was used, or that the construct resembles material associated with a known repository, laboratory practice, or technical community. Each of these findings would then point investigators toward specific follow-up questions, helping to structure the investigation and prioritize investigative resources. This role is relevant both to routine oversight, where it may help flag inconsistencies, and in post-incident investigation, where it may help narrow down possible methods, sources, or technical pathways.

At the same time, expectations should remain realistic. The presence of detectable engineering signatures does not mean that origin can always be established with confidence. If GEA is to become more useful for verification, progress will depend not only on better algorithms, but also on stronger data standards, broader reference datasets, more realistic evaluation settings, and clearer standards for interpreting attribution evidence. More consistent metadata on construct type, vector system, backbone, and engineering method would improve analysis. Benchmarks that include partial, mixed, or degraded samples would provide a more realistic test of performance. Sequence-based outputs will need to be interpreted alongside other forms of evidence rather than treated as self-explanatory.

GEA can support both routine oversight and post-incident investigation, but only as one input in a broader evidentiary process.

The task ahead is therefore twofold: to improve the science of attribution, and to integrate it carefully into broader governance and verification discussions without overstating what present methods can deliver. As biotechnology and artificial intelligence continue to evolve, the pressure to develop stronger attribution tools will only increase. These capabilities may support response, accountability, and deterrence, but only if their limitations are understood as clearly as their promise.[14]

Policy Recommendations

1. Improve data standards for engineered biological sequences. National and international efforts to strengthen biosecurity and research governance should support more consistent documentation of engineered biological material in repositories and sequence archives. This should include better systems for recording construct type, vector system, backbone, engineering method, and relevant depositor context.

2. Evaluate attribution methods under realistic conditions before operational use. Public-sector support for research and capability development should prioritize testing genetic engineering attribution methods under conditions closer to real-world incidents, including partial, degraded, mixed, or poorly documented samples. Evaluation should distinguish clearly between engineering detection and source attribution, and should include uncertainty reporting suited to high-stakes use.

3. Use genetic engineering attribution as one component of a broader evidentiary framework. Sequence-based attribution should be treated as one source of information among several, not as standalone proof. In practice, it is most useful when combined with epidemiological, environmental, and investigative evidence, including in discussions on compliance and verification.

Kadri Reis

PRIF – Peace Research Institute Frankfurt

1 The United States Department of Justice. (2010, February 19). *Amerithrax Investigative Summary*. https://www.justice.gov/archive/amerithrax/docs/amx-investigative-summary.pdf

2 Scientific Advisory Group on the Origins of Novel Pathogens (SAGO). (2025). *Independent assessment of the origins of SARS-CoV-2*. World Health Organization. https://www.who.int/publications/m/item/independent-assessment-of-the-origins-of-sars-cov-2-from-the-scientific-advisory-group-for-the-origins-of-novel-pathogens

3 Alley, E. C., Turpin, M., Liu, A. B., Kulp-McDowall, T., Swett, J., Edison, R., Von Stetina, S. E., Church, G. M., & Esvelt, K. M. (2020). A machine learning toolkit for genetic engineering attribution to facilitate biosecurity. *Nature Communications, 11*(1), 6293. https://doi.org/10.1038/s41467-020-19612-0
Lewis, G., Jordan, J. L., Relman, D. A., Koblentz, G. D., Leung, J., Dafoe, A., Nelson, C., Epstein, G. L., Katz, R., Montague, M., Alley, E. C., Filone, C. M., Luby, S., Church, G. M., Millett, P., Esvelt, K. M., Cameron, E. E., & Inglesby, T. V. (2020). The biosecurity benefits of genetic engineering attribution. *Nature Communications, 11*(1), 6294. https://doi.org/10.1038/s41467-020-19149-2

4 Lewis et al. (2020)

5 Cropper, N., Rath, S., & Teo, R. (2022). *Creating a Verification Protocol for the Biological Weapons Convention: A MODULAR-INCREMENTAL APPROACH*. Nuclear Threat Initiative. https://www.nti.org/wp-content/uploads/2022/06/Creating-a-Verification-Protocol_FINAL_June2022.pdf

6 Lewis et al. (2020)

7 Kamens, J. (2015). The Addgene repository: An international nonprofit plasmid and data resource. *Nucleic Acids Research, 43*(D1), D1152–D1157. https://doi.org/10.1093/nar/gku893

8 Alley et al. (2020)
Crook, O. M., Warmbrod, K. L., Lipstein, G., Chung, C., Bakerlee, C. W., McKelvey, T. G., Holland, S. R., Swett, J. L., Esvelt, K. M., Alley, E. C., & Bradshaw, W. J. (2022). Analysis of the first genetic engineering attribution challenge. *Nature Communications, 13*(1), 7374. https://doi.org/10.1038/s41467-022-35032-8

9 Lewis et al. (2020)

10 Biological Technologies Office. (2026). *Bio-attribution Challenge*. Defense Advanced Research Projects Agency. https://www.darpa.mil/research/challenges/bio-attribution

11 Arita, M., Karsch-Mizrachi, I., & Cochrane, G. (2021). The international nucleotide sequence database collaboration. *Nucleic Acids Research, 49*(D1), D121–D124. https://doi.org/10.1093/nar/gkaa967

12 Tay, A. P., Didi, K., Wickramarachchi, A., Bauer, D. C., Wilson, L. O. W., & Maselko, M. (2024). Synsor: A tool for alignment-free detection of engineered DNA sequences. *Frontiers in Bioengineering and Biotechnology, 12*, 1375626. https://doi.org/10.3389/fbioe.2024.1375626

13 Lewis et al. (2020)

14 Lewis et al. (2020)

Space

Trust but Verify: Autonomous Spacecraft and the Verification Challenge

Autonomy in satellite navigation and rendezvous and proximity operations (RPO) presents new opportunities for orbital verification, constellation resilience, and protection against disruptions to ground segments. At the same time, autonomy also reduces the number of decision points that can be directly attributed to intentional human action. Since attribution of irresponsible or aggressive behavior in orbit is already inherently challenging, the integration of autonomous systems further complicates strategic interpretation. In geopolitically sensitive orbital regimes such as geosynchronous orbit (GSO), poorly governed autonomous operations could increase the risk of miscalculation and escalation between space actors.

Enabling Orbital Verification: Innovations in Spacecraft Autonomy

On November 25, 2024, China announced the launch of what it described as the world's first operational "self-driving" Earth-observation satellites, equipped with autonomous navigation capabilities. These Siwei satellites are designed to independently maintain and adjust their orbital configurations without routine ground control intervention. Following China's announcement, developments in Europe during 2025 further illustrated the broader trend toward increasing spacecraft autonomy. For example, researchers at Julius-Maximilians-Universität Würzburg demonstrated an AI-based satellite attitude controller capable of autonomously executing spacecraft orientation maneuvers in orbit, representing an important step toward adaptive and intelligent spacecraft control systems.[1] A satellite with the capability to reposition itself when there is an opportunity to capture a better measurement offers the potential for in-orbit monitoring of objects and maneuvers.

Rendezvous and proximity operations (RPO): Orbital maneuvers where two spacecraft intentionally approach the same orbit and maneuver in close proximity, sometimes culminating in physical docking.

An autonomous satellite with advanced maneuverability capabilities can improve operational resilience against cyberattacks targeting the ground segment by maintaining mission operations with reduced dependence on ground-based control.[2] Furthermore, on-board autonomy can extend a satellite's operational lifetime by enabling real-time orbital maneuvers to avoid space debris and other satellites conducting potentially hazardous or aggressive maneuvers. This autonomy represents a crucial step toward distributed satellite constellation architectures that underpin resilient space defense postures and, by extension, deterrence by denial.[3]

Attitude control: The process of changing and maintaining the orientation of a spacecraft, to control which direction it is facing in space.

Advances in autonomous satellite navigation and RPO, combined with sensors tailored to internationally recognized verification missions, could enable spacecraft to conduct in-orbit verification activities, opening up verification approaches that were previously considered beyond practical technical reach. The UN First Committee could task an open-ended working group, in consultation with technical verification bodies such as the IAEA, with developing proposals for a fleet of internationally sanctioned inspector satellites.

Recent U.S. research[4] has proposed using detectors to exploit naturally occurring cosmic-ray protons to infer the presence of fissile material aboard satellites, with possible applications in the development of future nuclear warhead verification concepts. Although this work demonstrates the feasibility of the underlying concept through modeling, it remains exploratory and does not yet constitute a mature nor internationally accepted verification capability. The political feasibility of orbital verification processes within the international community remains fiercely contested, with no agreed norms governing their conduct, scope, or legal status. In the present context, the capabilities afforded by current technology outpace international governance and cooperation in this area.

Departure burn: A rocket engine firing that provides the velocity needed for a spacecraft to leave its current orbit and begin its journey toward another destination.

Hibernation orbit: An orbit in which a spacecraft remains for an extended period with minimal activity while awaiting a future mission phase, conserving fuel and power until operations resume.

Flyby: A spacecraft trajectory that brings a satellite into close proximity with another object while continuing on a non-capture path.

Existing On-Orbit Inspection Capabilities

AeroAstro Inc. Escort satellite project[5] proposed a *small satellite inspector* concept in which an inspector satellite would trail a primary satellite stationed in geostationary equatorial orbit (GEO) to analyze its performance. The inspector satellite's orbit would be optimized to slow its approach, thereby enabling safe RPO by reducing the risk of collision. The inspector could then perform a departure burn and move into a hibernation orbit upon completion of an inspection flyby. Several projects have operationalized the mechanism of an Escort-type inspector satellite, demonstrating autonomous RPO and in-orbit satellite servicing. Notably, Northrop Grumman's **MEV missions** performed life-extension and proximity operations in GEO, Starfish Space's **Remora mission** demonstrated autonomous RPO in low-Earth orbit (LEO), and Astroscale's **ELSA-d mission** validated safe close-proximity approaches and docking maneuvers. Europe's **e.Inspector** project, a space debris inspection technology, is under development with a demonstration mission scheduled for 2027.

Orbital Behavior as an Object of Verification

The hovering/trailing companion satellite concept has also been generalized to other proximity operations, commonly referred to as "guardian" satellites. While both inspector and guardian satellites rely on controlled proximity operations, their design philosophies diverge fundamentally. Inspector satellites prioritize carefully shaped approach trajectories and strict collision avoidance to enable nonintrusive inspection. In contrast, guardian satellites may adopt persistent close-range positioning, where the capacity for rapid interception, or even collision, is not merely a risk but an intentional feature leveraged as a form of deterrence.[6] Evidently, the RPO associated with guardian satellites is more applicable to defense strategies than to orbital verification.

Differentiating between peaceful and harmful RPO intent requires discerning orbital mechanics, particularly to identify anomalous behavior. Newman and Tate[7] remark that RPO with nefarious intent might exhibit a higher relative velocity, which is uncharacteristic of routine operations in that orbit. Kelecy et al.[8] employ machine learning techniques to infer the intent of RPO behavior from observed relative motion. Their work advances an innovative approach based on real-time imagery without requiring

Orbital Regimes and Associated Counter-space Capabilities



	LEO Low Earth Orbit	MEO Medium Earth Orbit	GEO Geosta-tionary Orbit	GSO Geosynch-ronous Orbit	HEO Highly Elliptical Orbit
Orbital conge-stion	Very high	Low–Moderate	High	High	Very low
Attack logic	Quick strike	Targeted attack	Deny prime orbit	Staying nearby	Specific oppor-tunity
Illustrati-ve capabi-lities	• Kinetic attack (ASAT) • Electronic attack (jam-ming) • Brief proximity attacks	• Electronig war-fare (jamming) • Interception/ disruption of specific instru-ments	• Proximity at-tacks/ invasive inspection • Electronic attacks & di-rected energy attacks	Same capabili-ties as GEO but with varying attack objectives	• Interception • Electronic attacks & mis-sile-warning disruption

Terminology note:
This paper uses GSO as an abbreviation for geosynchronous orbit in the general technical sense. When referring specifically to the International Telecommunication Union (ITU) Radio Regulations, however, GSO denotes geostationary-satellite orbit, consistent with ITU terminology. To avoid ambiguity, references to the ITU regulatory framework should retain the ITU usage.

detailed tracking analytics. This approach could enable the verification of orbital behavior whilst maintaining the necessary information barriers, an important principle of effective verification practice. Space-faring states should establish agreed limits on verification sensors and inspection activities (information barriers) to ensure that orbital verification can be conducted without enabling intrusive surveillance.

Dual-Use Orbital Behavior and the Feasibility of Verification

Autonomous RPO operations could compress what would previously have required multiple levels of authorization[9] into a single automated decision-making process, thereby complicating the attribution of culpability and liability. This challenges the concept of liability under the Outer Space Treaty. The accompanying uncertainty heightens the risk that routine or ambiguous maneuvers could be interpreted as hostile. Some analysts argue that the central concern is the potential for aggression[10] rather than the mere existence of the capability, suggesting that risk mitigation should focus on observable behavior, such as trajectories and operational patterns. Verification-related maneuvers, trajectories, and approach velocities should be transparent and observable to minimize ambiguity and reduce perceptions of threat.

For a short introduction to the Outer Space Treaty, see p. 123.

Autonomous RPO could enable an aggressor to preposition a spacecraft in a seemingly benign orbit for peaceful purposes, only to activate it later and maneuver into the orbital plane of a target. From there, it could follow an optimized trajectory to conduct a close-approach attack. A commonly cited example of such co-orbital behavior is the Russian spacecraft Kosmos 2499, which was initially cataloged as space debris before later demonstrating active maneuvering capabilities. Based on repeated close approaches and orbital matching with its associated Briz-KM upper stage, several secondary studies (orbital tracking analyses) have interpreted the mission as demonstrating capabilities consistent with a potential co-orbital anti-satellite system. Although no official designation confirms a weapons-related purpose, the case highlights the ambiguity that arises from the dual-use nature of RPO. Autonomy may then amplify the strategic instability inherent in such scenarios by enabling spacecraft to transition rapidly between maneuvering states.

Constraints on Verification Imposed by Orbital Regimes

Nuclear command, control, and communications (NC3): The overarching framework of hardware, software, and personnel that allows national leaders to monitor threats, authorize the use of nuclear weapons, and transmit launch orders to military forces in the field.

These risks are not uniform across orbital regimes. Attacks involving RPO may be more plausible in GEO than in LEO, as high-value, dual-use assets in GEO occupy predictable orbital slots,[11] making them more susceptible to targeting. These satellites often support strategic communications and early-warning functions, meaning that interference (electronic, cyber, radiofrequency) could be interpreted as a precursor to a disarming strike, carrying significant escalatory risks. In this context, the distinction between conventional and nuclear signaling becomes increasingly blurred. More specifically, Egeli[12] argues that U.S. doctrine already implies the possibility of nuclear retaliation in response to nonnuclear attacks on NC3 systems.[13] Although often described as a U.S. "threat," this is better understood as an inferred signal derived from broader doctrinal ambiguity. This concentration of strategic value and escalation sensitivity in GEO stands in contrast to the more distributed and resilient architecture of LEO systems.

This does not, however, preclude the use of RPO in LEO altogether. Autonomous systems may still enable selective targeting, intelligence collection, or coercive signaling. However, these applications are more consistent with limited or ambiguous interfer-

ence than with the high-impact, prepositioned attack scenarios that are more feasible in GEO. As a result, while autonomy increases the operational flexibility of RPO across orbital regimes, its most destabilizing applications are likely to manifest in GEO rather than LEO, with important implications for the feasibility of orbital verification.

Sarah McKee

PRIF – Peace Research Institute Frankfurt

1 Julius-Maximilians-Universität Würzburg. (2025, July 11). *World premiere in space: Würzburg AI controls satellite.* https://www.uni-wuerzburg.de/en/news-and-events/news/detail/news/world-premiere-ai-control/

2 Colmenero, F. J., Ocón, J., Alonso, M., Jalvo, R., & Ramos, J. (2026). ERGO: An autonomy framework for space robotics and beyond. *Engineering Proceedings, 133*(1), 107. https://doi.org/10.3390/engproc2026133107
Tasdighi, A. (2025). Cyber-resilient autonomous spacecraft: A multi-domain resilience framework for deep space missions. *American Journal of Computer Science and Technology 8*(1), 1–15. https://dx.doi.org/10.2139/ssrn.5085627

3 Mazarr, M. J. (2020). Understanding deterrence. In F. Osinga & T. Sweijs (Eds.), *NL ARMS Netherlands Annual Review of Military Studies 2020: Deterrence in the 21st century—Insights from theory and practice* (pp. 13–28). T.M.C. Asser Press. https://doi.org/10.1007/978-94-6265-419-8_2

4 Danagoulian, A. (2026). Verification of the Outer Space Treaty with cosmic protons. *Nature, 655*, 585–590. https://doi.org/10.1038/s41586-026-10783-2

5 Jacobovits, A., & Vaneck, T. (2003). AeroAstro's Escort: A microsatellite for on-orbit inspection of space assets. In *Proceedings of the Small Satellite Conference* (SSC03-IV-7). Utah State University. https://digitalcommons.usu.edu/smallsat/2003/All2003/30/

6 Egeli, S. (2021). Space-to-space warfare and proximity operations: The impact on nuclear command, control, and communications and strategic stability. *Journal for Peace and Nuclear Disarmament, 4*(1), 116–140. https://doi.org/10.1080/25751654.2021.1942681

7 Innovation News Network. (2026, February 17). *How close is too close? Rendezvous and proximity operations in crowded orbitals*. https://www.innovationnewsnetwork.com/rendezvous-and-proximity-operations-in-crowded-earth-orbit/66203/

8 Kelecy, T., Abernethy, S., Jones, F., Gerber, E., & Wurzel, H. (2022). Predicted intent inferred from real-time rendezvous and proximity behavior. In *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference* (pp. 1–16). Advanced Maui Optical and Space Surveillance Technologies Conference. https://amostech.com/TechnicalPapers/2022/Conjunction-RPO/Kelecy.pdf

9 Johnson, K., Roberts, T. G., & Weedon, B. (2022). Mitigating noncooperative RPOs in geosynchronous orbit. *Æther: A Journal of Strategic Airpower & Spacepower, 1*(4), 79–94. https://www.jstor.org/stable/48711887

10 Staats, B. (2022). Mitigating security risks and potential threats of emerging Rendezvous and Proximity Operations. *Astropolitics, 20*(1), 64–92. https://doi.org/10.1080/14777622.2022.2080547

11 Johnson et al. (2022)

12 Egeli (2021)

13 Acton, J. M. (2018). Escalation through entanglement: How the vulnerability of command-and-control systems raises the risks of an inadvertent nuclear war. *International Security, 43*(1), 56–99. https://doi.org/10.1162/isec_a_00320

Cyber

What Cyberspace Regulation Can Learn from Traditional Arms Control

Mechanisms from chemical weapons disarmament could inform future regulation of cyberspace. Despite growing cyber threats and the increasing military use of cyber operations, international regulation remains fragmented and largely nonbinding. Drawing on interdisciplinary expert interviews, this chapter identifies key challenges for cyberspace regulation, including dual-use technologies, verification difficulties, rapid technological change, and limited political will. The CWC—particularly its General-Purpose Criterion, attribution-related practices, independent expertise, and related informal practices—could provide valuable lessons and support a shift from technology-centered regulation toward behavior- and context-based governance in cyberspace.

For a short introduction to the CWC, see p. 128 f.

Can lessons from traditional Arms Control (AC), such as in relation to chemical weapons, offer useful guidance for future regulation of cyberspace despite the apparent distinctiveness of the weapon system involved? Drawing on expert interviews from the fields of cybersecurity, AC, and policy, this chapter argues that important conceptual lessons can be learned from traditional AC, as illustrated here using concepts from the Chemical Weapons Convention (CWC).[1]

Why Look at Traditional Arms Control for the Regulation of Cyberspace?

We are seeing a rise in cyber incidents worldwide, increasingly linked to hybrid warfare, with cyberattacks becoming ever more targeted and destructive.[2] Cyber operations have become an increasingly influential instrument in international security and military conflict, where state and non-state actors employ offensive cyber capabilities for espionage, disruption, and attacks on critical infrastructure.[3] While appropriate approaches to regulating cyberspace are urgently needed, no binding international AC or regulatory regime currently exists for cyberspace. Existing governance approaches mainly rely on voluntary norms, confidence-building measures, and informal cooperation. The key fora and institutions in the field are the Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace in the Context of International Security (GGE), the Open-Ended Working Group on Developments in the Field of ICTs in the Context of International Security (OEWG), and the UN Treaty on Cybercrime, overall constituting a rather informal and fragmented normative regulatory structure.[4]

Arms control agreements have proven successful in many areas. "Arms control aims at preventing conflicts and fostering stability in inter-state relations by either reducing the probability of usage of a specific weapon or regulating its use and thus, reducing the costs of armament".[5] Approaches to AC can be diverse and may address aspects beyond the regulation of use. While scholars are skeptical about simply applying traditional AC measures to emerging technologies,[6] we ask: What would effective AC measures for cyberspace look like? Instead of comparing regulatory approaches, we focus on the similarities in the challenges involved in regulating weapons; here "established arms control regimes are confronted with comparable challenges, including dual-use considerations and technological change".[7]

Dual-use technology: Technology that can serve both civilian and military purposes; be used offensively or defensively; or be misused for unintended purposes.

We recognize that contemporary AC "captures the wide range of control initiatives that can be deployed to limit the instruments of armed violence",[8] including soft law approaches, such as informal norms and practices that are already established in this domain.[9] Here, recent studies indicate a shift in cyber governance toward informal regulation and norms-based approaches.[10] At the same time, norms-based regulation has always been an essential part of traditional AC. By assessing similar challenges across different weapon technologies, we aim to derive lessons for regulation and examine them against the backdrop of the growing importance of soft law approaches. As "informal governance mechanisms are already evolving to fill that governance gap",[11] such mechanisms often serve as precursors to the formalization and standardization of these norms. This particularly concerns incorporating domain-specific expertise by non-state and private actors and building on their established practices. With a view to achieving chemical weapons disarmament, engineers and scientists played a central role in drafting the CWC.[12] Since the CWC is considered one of the most successful AC regimes, we will use it here as an example.

The Challenges of Regulating Cyber Weapons

One reason for the regulatory gap lies in the specific characteristics of cyberspace, compounded, *inter alia*, by the absence of a universally accepted international definition of what constitutes a *cyber weapon*. Building on our past work,[13] we defend the validity and appropriateness of using the term *cyber weapon*, drawing extensively on scholars who have examined in detail the conceptualization of bellicose cyber operations as *cyber weapons*.[14] As such, we consider definitional approaches that focus on the purpose and intent of cyber tools to be the most useful,[15] i.e., "any computer equipment or device that is designed, intended or used to have violent consequences; that is, to cause death or injury to persons or damage or destruction of objects".[16] Similarly, the definition of chemical weapons in the CWC centers on the capacity of toxic chemicals to cause harm.[17] Moreover, cyber technologies are dual-use, meaning they have both civilian and military applications or can be used for both offensive and defensive purposes. Dual-use cyber technologies can include cryptography and encryption software, intrusion software and spyware, as well as artificial intelligence (AI) and its harmful potentials.[18] Furthermore, cyber technologies evolve rapidly, with software being continuously updated, modified, and integrated into diverse applications. It would therefore appear difficult to transfer traditional AC mechanisms based on counting, monitoring, or destroying weapons stockpiles to the cyber domain. Verification of actions in cyberspace presents a particular challenge, since commercial, state, and civilian applications are difficult to differentiate, quantitative limitations cannot be imposed, and cyber weapons are hard to track due to their non-physical nature. The private and commercial industry sector, in particular, is gaining importance in the cyber domain through its ownership of important infrastructure, operational rights, and state-of-the-art knowledge.[19]

The regulatory gap is due in part to the specific characteristics of cyberspace and the absence of an accepted international definition of what constitutes a cyber weapon.

From Chemical Disarmament to Cyberspace Regulation: What Can We Learn?

Common challenges in regulating both chemical substances and cyber technologies include the dual-use dilemma, verification difficulties, rapid technological change, and the significant role played by private actors. Experts highlight the fact that cyber weapons are harder to regulate than chemical weapons because they lack physical materiality. Verification therefore represents one of the most difficult challenges to effective cyberspace regulation. Traditional inspection regimes—such as the one established by the Organisation for the Prohibition of Chemical Weapons (OPCW)—would require intrusive access to sensitive systems and could expose state secrets or defensive capabilities. At the same time, technological change in cyberspace occurs much faster than in traditional AC domains. In addition, the private sector plays a much larger role because major parts of cyberspace infrastructure are privately owned and operated.

Verification represents one of the most difficult challenges to effective cyberspace regulation.

The success of the CWC was closely linked to historical experiences and the political conditions prevailing at the time of its negotiation, shortly after the end of the Cold War. Chemical weapons had already caused large-scale human suffering, creating a strong normative taboo against their use.[20] Experts argue that these conditions are not yet present for cyber weapons. Cyber capabilities also provide considerable strategic advantages for military operations, espionage, and intelligence gathering, reducing incentives for comprehensive restrictions.

Focusing on Context of Use Instead of Technologies

Despite these differences, experts identify several mechanisms from chemical weapons AC that could inform cyberspace regulation. One particularly relevant concept is the General-Purpose Criterion (GPC) contained in the CWC. Rather than prohibiting only specific substances, the GPC defines chemical weapons according to their intended harmful use. Applied to cyberspace, this approach would shift attention away from regulating specific technological applications—such as software or technical tools—and toward harmful behavior, intended effects, and prohibited targets such as critical infrastructure.

Experts, therefore, suggest that, rather than attempting to ban individual technologies, cyberspace regulation should define unacceptable actions and contexts of use. Behavior-based regulation is more adaptable to technological innovation, may help prevent regulatory loopholes, and could provide a means of addressing the dual-use problem.

Attribution and Verification of Cyber Operations

Attribution: The process of identifying the actor responsible for an incident or activity through technical, contextual, and political analysis.

While traditional AC verification mechanisms are difficult to apply in cyberspace, experts repeatedly emphasize the growing importance of attribution. In the case of a suspected (cyber)attack, attribution refers to the identification of the actor responsible for the attack. Regarding cyber operations, attribution has become considerably more feasible in recent years due to improved technical capabilities and investigative expertise, and is already being carried out by the private sector and independent research institutes. Davis et al.,[21] for example, demonstrate how stateless attribution is already being established in practice. Building on this, future cyber AC could rely less on intrusive inspections and more on trusted attribution structures. For example, Reinhold et al.[22] have developed a method for eliminating matching cyber weapons stockpiles without the need for intrusive measures. In cyberspace, stockpiles refer to the collection of knowledge of vulnerabilities and exploits.[23] To adapt the *managed access* approach from chemical weapons disarmament verification for cyberspace would require technical expertise for fact finding and ensuring protection of sensitive data, as well as political credibility. At the same time, as with chemical weapons disarmament in general, this should take place within a multi-stakeholder framework.

An independent institution, for instance under the auspices of the UN, could analyze cyber incidents, provide evidence, and support international accountability processes under the umbrella of a globally recognized attribution mechanism. Attribution alone would not solve all challenges. Questions regarding standards of evidence and political interpretation, as well as access to sensitive or classified data, would remain contested. Nevertheless, attribution represents one of the most realistic avenues for practical cyber verification.

The Path to Cyberspace Regulation

Cyberspace regulation is unlikely to begin with comprehensive disarmament treaties comparable to the CWC. Instead, progress will probably emerge through informal norms, i.e., soft law practices, comprehensive norm-building processes, and incremental governance mechanisms.[24] However, soft law instruments and norm building are important prerequisites for potential future binding AC agreements. More and more research indicates shifts in the effectiveness of weapons regulations, suggesting that emerging disruptive technologies, in particular, might be better regulated by a combination of institutionalized regulation and informal norms (see, e.g., Schliebs & Vohs[25]). Existing initiatives, such as UN expert groups, codes of conduct, and national responsible state behavior norms, already contribute to this process. Ultimately, the comparison with chemical weapons disarmament demonstrates that some of its elements could be applied to the regulation of cyberspace, but this would require long-term norm development, political will, and institutional trust. Cyber AC is therefore likely to develop incrementally, through a combination of behavioral norms, attribution practices, and international cooperation mechanisms, rather than through the immediate adoption of comprehensive bans. The next steps to advance this process could, therefore, involve integrating existing approaches to cyber attribution by incorporating the perspectives and practices of governments, the private sector, and nongovernmental organizations. Here, an agreed set of standardized norms and practices focusing on the context of use of cyber technologies against the backdrop of the GPC could be promising, with the goal of establishing an independent attribution mechanism.

Progress in cyberspace regulation will probably emerge through informal norms.

Anja-Liisa Krichbaum

PEASEC/Technical University of Darmstadt

Christian Reuter

PEASEC/Technical University of Darmstadt

Helene Pleil

interface – Tech analysis and policy ideas for Europe e.V.

1 This chapter is based on our recent publication: Gonsior, A.-L., Pleil, H., & Reuter, C. (2026). 'Difficult but possible': An exploratory approach to the regulation of cyber weapons using lessons learned from chemical weapons arms control. *Journal of Cyber Policy*, 1–18. https://doi.org/10.1080/23738871.2026.2665269

2 Dunn Cavelty, M., & Wenger, A. (2020). Cyber security meets security politics: Complex technology, fragmented politics, and networked science. *Contemporary Security Policy*, *41*(1), 5–32. https://doi.org/10.1080/13523260.2019.1678855

3 Center For Strategic & International Studies. *Significant Cyber Incidents*. https://www.csis.org/programs/strategic-technologies-program/significant-cyber-incidents
United Nations Secretary-General, & United Nations Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security. (2021). *Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace in the Context of International Security* (A/76/135). https://digitallibrary.un.org/record/3934214?v=pdf

4 Henriksen, A. (2019). The end of the road for the UN GGE process: The future regulation of cyberspace. *Journal of Cybersecurity*, *5*(1). https://doi.org/10.1093/cybsec/tyy009

5 Reinhold, T., & Reuter, C. (2019). Arms Control and its Applicability to Cyberspace. In C. Reuter (Ed.) *Information Technology for Peace and Security* (2nd edn, pp. 207–231). Springer Vieweg. https://doi.org/10.1007/978-3-658-25652-4_10

6 Ford, C. A. (2020). *International Security in Cyberspace: New Models for Reducing Risk* (Volume 1, Number 20; Arms Control and International Security Papers). Office of the Under Secretary of State for Arms Control and International Security.
Wisotzki, S., & Kühn, U. (2021). Crisis in arms control: An introduction. *Zeitschrift Für Friedens- Und Konfliktforschung, 10*(2), 183–194. https://doi.org/10.1007/s42597-022-00074-8

7 Gonsior, Pleil, & Reuter (2026)

8 Cooper, N., & Mutimer, D. (2011). Arms Control for the 21st Century: Controlling the Means of Violence. *Contemporary Security Policy, 32*(1), 3–19. https://doi.org/10.1080/13523260.2011.556820

9 Mačák, K. (2017). From Cyber Norms to Cyber Rules: Re-engaging States as Law-makers. *Leiden Journal of International Law, 30*(4), 877–899. https://doi.org/10.1017/S0922156517000358

10 Ford (2020)
Futter, A. (2020). *What does cyber arms control look like? Four principles for managing cyber risk* [GLOBAL SECURITY POLICY BRIEF]. European Leadership Network. https://www.europeanleadershipnetwork.org/wp-content/uploads/2020/06/Cyber-arms-control.pdf
Meyer, P. (2011). Cyber security through arms control—An approach to international cooperation. *The RUSI Journal*.

11 Hagemann, R., Huddleston, J., & Thierer, A. D. (2018). Soft Law for Hard Problems: The Governance of Emerging Technologies in an Uncertain Future. *Colorado Technology Law Journal*. https://ssrn.com/abstract=3118539

12 Matoušek, J. (2010). The Chemical Weapons Convention and the Role of Engineers and Scientists. *Kemija u Industriji*, 59(2), 69–78.

13 Gonsior, Pleil, & Reuter (2026)

14 Maybaum, M., & Tölle, J. (2016). *Arms Control in Cyberspace – Architecture for a Trust-Based Implementation Framework Based on Conventional Arms Control Methods*. 8th International Conference on Cyber Conflict, Tallinn. https://ccdcoe.org/uploads/2018/10/Art-11-Arms-Control-in-Cyberspace.pdf
Reinhold, T., & Reuter, C. (2021). Towards a Cyber Weapons Assessment Model – Assessment of the Technical Features of Malicious Software. *IEEE Transactions on Technology and Society*.

15 Arimatsu, L. (2012). A Treaty for Governing Cyber-Weapons: Potential Benefits and Practical Limitations. *4th International Conference on Cyber Conflict, NATO CCD COE Publications*.
Reinhold, T. (2020). Cyberspace as Military Domain: Monitoring Cyberweapons. In D. Feldner (Ed.), *Redesigning Organizations* (pp. 267–278). Springer International Publishing. https://doi.org/10.1007/978-3-030-27957-8_20

16 Maybaum & Tölle (2016), p. 161

17 Robinson, J. P. P. (2008). Difficulties facing the Chemical Weapons Convention. *International Affairs, 84*(2), 223–239.

18 Riebe, T., Schmid, S., & Reuter, C. (2024). Dual-Use Information Technology: Research, Development and Governance. In C. Reuter (Ed.), *Information Technology for Peace and Security* (2nd edition, pp. 169–188). Springer Vieweg. https://doi.org/10.1007/978-3-658-44810-3_8

19 Larsen, J. A., & Smith, S. (Eds). (2023). *Arms control at a crossroads: Renewal or demise?* Lynne Rienner Publishers, Inc.

20 Price, R. (1997). The Chemical Weapons Taboo, Ithaca: Cornell University Press

21 Davis, J., Boudreaux, B., Welburn, J., Aguirre, J., Ogletree, C., McGovern, G., & Chase, M. (2017). *Stateless Attribution: Toward International Accountability in Cyberspace*. RAND Corporation. https://doi.org/10.7249/RR2081

22 Reinhold, T., Kuehn, P., Günther, D., Schneider, T., & Reuter, C. (2023). ExTRUST: Reducing Exploit Stockpiles with a Privacy-Preserving Depletion System for Inter-State Relationships. *IEEE Transactions on Technology and Society*. https://doi.org/10.48550/ARXIV.2306.00589

23 Reinhold et al. (2023)

24 Abbott, K. W., & Snidal, D. (2000). Hard and Soft Law in International Governance. *International Organization, 54*(3), 421–456. https://doi.org/10.1162/002081800551280
Gutierrez, C. I., & Marchant, G. E. (2021). *A Global Perspective of Soft Law Programs for the Governance of Artificial Intelligence*. Social Science Research Network (SSRN).
Hagemann, R., Huddleston, J., & Thierer, A. D. (2018). Soft Law for Hard Problems: The Governance of Emerging Technologies in an Uncertain Future. *Colorado Technology Law Journal*. https://ssrn.com/abstract=3118539
Peters, A., & Pagotto, I. (2006). Soft Law as a New Mode of Governance: A Legal Perspective. *NEWGOV: New Modes of Governance*. https://doi.org/10.2139/ssrn.1668531

25 Schliebs, M., & Vohs, V. (2026). 'This Is My Last Resort': Approaches to Overcome the Stalemate in Autonomous Weapons Regulation Through National Legislation and Industry Self-Regulation. In B. Boutin, T. K. Woodcock, & S. Soltanzadeh (Eds), *Legal, Ethical, and Technical Dilemmas in Military Artificial Intelligence* (pp. 287–310). T.M.C. Asser Press. https://doi.org/10.1007/978-94-6265-759-5_12

Verification

Systems for Implementation Review

Systems for implementation review (SIRs) are international or regional institutions and mechanisms through which parties to a multilateral agreement share information on implementation, compare their implementation activities, review one another's performance, learn from implementation experience, address noncompliance or poor compliance, and, where useful, adjust their commitments. They complement specific verification measures where these exist, and can help address problems of poor compliance where they do not. Experience, particularly from environmental regimes, shows that SIRs can be used to make a significant difference to the implementation and effectiveness of complex multilateral agreements. The main question addressed in this chapter is *whether and how the further development and use of SIRs can improve the dynamism and effectiveness of existing and emerging multilateral arms control and disarmament regimes*. This chapter argues that greater attention to developing and using SIRs in multilateral arms control regimes could make a difference, briefly illustrating this argument with reference to the Biological Weapons Convention (BWC), Chemical Weapons Convention (CWC), the Non-Proliferation Treaty (NPT), and humanitarian arms control agreements, as well as initiatives to promote international cooperation on the peaceful uses of outer space and to address the use of cyberweapons.

The development and effectiveness of multilateral agreements depends not only on the international configurations of power and interests but also on how the agreement and its institutions have been designed and developed. Verification systems and SIRs are both examples of the importance of institutional design. Compared with verification, SIRs have received relatively little attention in the fields of arms control and disarmament. At a time when progress on arms control is particularly difficult, it is important to explore how the development or increased use of SIRs could be helpful.

Systems for implementation review are international or regional institutions and mechanisms through which parties share information on implementation, compare implementation activities, review one another's performance, address noncompliance or poor compliance, and, where useful, adjust their commitments. These systems operate at the international level, but focus on implementation performance at the national level. They include mechanisms for regularly exchanging national reports on prog-

ress in implementation; international expert groups that advise on opportunities and challenges related to implementation; fact-finding or problem-solving mechanisms to address emerging problems with national implementation; peer-review processes; arrangements for lesson learning; and feedback mechanisms to facilitate possible adjustments to commitments.

SIRs complement verification systems by focusing on improving implementation rather than detecting noncompliance.

While verification systems focus on detecting noncompliance, SIRs are designed to help parties improve their implementation performance, recognizing that inadequate implementation by some parties may be due to a range of factors, many of which can be addressed through timely mobilization of international attention and assistance. For this reason, SIRs have often been less politically sensitive and easier to negotiate.

Evidence on the Significance of SIRs for International Regime Effectiveness

The significance of SIRs was first identified by research on the effectiveness of international environmental agreements. It was one of the main findings of a major three-year international research project titled "The Implementation and Effectiveness of International Environmental Commitments. Theory and Practice" (IEC) conducted in 1993–1996 by researchers based at the International Institute of Applied Systems Analysis (IIASA) in Austria.[1] The study, which comprised detailed analysis of a range of case studies, showed that the development and operation of SIRs was an important factor in improving implementation and effectiveness. It also found that weak or absent SIRs were a reliable indicator that an environmental agreement was probably making little difference to practice even if much effort went into its negotiation.

The impact of an environmental agreement can be reliably indicated by the strength of its SIRs.

The findings of this IEC research project attracted wide international attention and significantly influenced international negotiations on the further development of, among others, the climate change, stratospheric ozone protection, regional seas, and several biodiversity regimes.[2] Much diplomatic and technical effort was devoted to further developing SIRs in these regimes over the subsequent decade and beyond, with substantial success.

SIRs have contributed to environmental regimes by providing an international operating context and facilitating a global network of technical experts.

Well-developed SIRs have made two key contributions to environmental regimes. First, they provided an international operating context containing extensive detailed information directly relevant to assessing progress on national implementation efforts. Second, they facilitated dynamic global networks of national and international technical experts to identify and resolve emerging problems with national implementation. The expert networks operated within diplomatically approved international mechanisms and reported periodically to the Secretariat and states parties on their activities. However, the experts were also able to work together directly and flexibly without constant diplomatic involvement. The Montreal Protocol, climate change regime, Baltic Sea agreements, and CITES (Convention on International Trade in Endangered Species), for example, created specific incentives for states parties to report problems with national implementation. Such reporting served not only as a first step toward mobilizing assistance, but, in some cases, also opened the possibility of obtaining an exemption from a particularly challenging commitment if the problem could not be resolved with the help of international experts.

In this context, relevant national technical experts and implementing agencies became active participants in the wider international regime. This significantly increased the dynamism of the regime and also created incentives for successful national experts to engage constructively, for example by offering them opportunities for career advancement and access to international funding. In the process, clear

evidence emerged that SIRs not only improved national implementation of treaty commitments but also enabled levels of implementation that often went well beyond minimum requirements. Moreover, where SIRs are well developed, it becomes much easier to identify and question any information gaps in national implementation reports, thereby helping to deter noncompliance or misinformation, much like verification systems.

SIRs in Existing Arms Control and Disarmament Regimes

Although explicit awareness and discussion about the significance of SIRs remain relatively limited within the arms control and disarmament community, many multilateral arms control agreements include provisions for national implementation reporting, technical expert groups to advise and support national implementation efforts, peer-review mechanisms, and other types of SIRs. This implies that there is already a basis for further developing SIRs in ways that can help increase the dynamism and effectiveness of existing arms control and disarmament regimes. To illustrate this, the following paragraphs briefly discuss the opportunities and constraints associated with selected existing agreements.

The Biological Weapons and Toxins Convention (BWC)

For a short introduction to the BWC, see p. 127 f.

For many years, disputes relating to Article 5 and the development of a verification system have hindered progress toward strengthening the BWC. It is widely acknowledged that, due to a combination of commercial, technical, and security reasons, negotiating an effective BWC verification system would be complex and difficult. In this context, numerous initiatives have sought to develop useful SIRs, including information exchanges, confidence-building and transparency visits, peer-review mechanisms, a Science and Technology Review Working Group, and an International Agency for Biological Security to support national implementation of BWC commitments.[3] If approved, these initiatives would have helped to strengthen the BWC regime. However, they have generally been blocked by a coalition of states, including Russia, which may perceive such SIRs as a threat to their efforts to keep the BWC weak. This case illustrates both the potential for SIRs to enhance the dynamism of the BWC regime and the obstacles to improving regime effectiveness when key parties are determined to block such efforts. Nevertheless, if and when political opportunities arise to approve initiatives within the BWC, significant progress is likely to be achieved most rapidly on SIRs, whereas an acceptable verification regime would take considerably longer to negotiate.

SIRs represent the most likely avenue for rapid progress amidst the difficulty around negotiating an acceptable BWC verification system.

The Chemical Weapons Convention (CWC)

For a short introduction to the CWC, see p. 128 f.

The Chemical Weapons Convention (CWC) has well-developed systems for both verification and implementation review, supported by the treaty's Organisation for Prohibition of Chemical Weapons (OPCW). Over the past decade, the operation of these systems has been constrained and contested by efforts by Syria, Russia, and their allies to obstruct OPCW investigations into the production and use of chemical weapons by the Assad regime during the Syrian Civil War. However, the tide appears to have turned on this issue, as international evidence-gathering efforts and site visits have increased dramatically since the Assad regime was toppled in 2024.[4] In the meantime, the dynamism of chemical and biochemical research and development presents state parties with increasingly complex challenges when it comes to effective regulation and control, making SIRs a vital mechanism for maintaining regime effectiveness. In this context, the work of Temporary Working Groups under the Scientific Advisory Board, together with the ongoing development of an Action Plan on National Implementation of CWC commitments, provides an important focus. Multiple SIRs mechanisms and resources will be needed to promote the effectiveness of this Action

Multiple SIRs will need to synergize to ensure the effectiveness of an upcoming Action Plan on National Implementation of CWC commitments.

Plan, some in specialist or emerging areas of concern. We recommend that careful attention is devoted to the design of these systems, enabling them to work in synergy, thereby contributing to improved national performance and the timely identification and response to emerging problems across all states parties.

The Nuclear Non-Proliferation Treaty (NPT)

For a short introduction to the NPT, see p. 122.

The Nuclear Non-Proliferation Treaty (NPT) has a complex structure, in which the International Atomic Energy Agency (IAEA) has primary responsibility for operating the nuclear safeguards regime, while the NPT Review Conference process provides the main forum for reviewing progress on other key treaty commitments, including peaceful nuclear technology transfer to developing countries and progress toward disarmament by the five nuclear weapons states recognized under the treaty. In response to significant criticism, the IAEA nuclear safeguards regime has been strengthened over the past two decades. Although mainly designed for verification purposes, in practice, the present nuclear safeguards regime already includes a range of reporting, review, and lesson-learning mechanisms that could also be used more purposefully as SIRs. They could provide a basis for global or regional initiatives to reduce underlying vulnerabilities to proliferation—for example, by helping to limit the use of highly enriched uranium in medical or research reactors or further improving the security of nuclear stockpiles. Further, if the five nuclear weapons states were to renew their interest in at least paving the way for nuclear disarmament, they could participate in international SIRs tailored to this purpose, while also implementing confidence-building measures among themselves.

Although more focused on verification, the present nuclear safeguards regime does include measures that could be consolidated into more purposeful SIRs, especially with the support of nuclear weapons states.

Humanitarian Arms Control

We now turn to "humanitarian" arms control agreements, such as the Ottawa Convention, the UN Programme of Action on Small Arms and Light Weapons, and the Arms Trade Treaty. This broad category of multilateral agreements has a long history, but expanded particularly rapidly between 1995 and 2015. During this period, opportunities were taken to elaborate on the implications for specific aspects of human insecurity of long-established commitments under International Humanitarian Law and International Human Rights Law, including the indiscriminate effects on civilians of use of anti-personnel mines and the high levels of death and injury in many countries and areas associated with inadequately controlled small arms. Systems for implementation review are particularly relevant to the operation of these humanitarian arms control regimes because of the importance of engaging with the wide range of governmental and intergovernmental agencies, officials, and experts directly involved in addressing the complex, context-specific challenges of national implementation of commitments.

SIRs have been customized for humanitarian arms control agreements to promote progress both in the implementation of agreed norms and in the mobilization of international cooperation.

The effectiveness of humanitarian arms control agreements typically depends on two factors: (i) the implementation of agreed norms and responsible practices by governments and other agencies, particularly those that have previously been neglectful or lacked the necessary national awareness and regulatory systems; and (ii) the mobilization of international cooperation to support programs of action in areas severely affected by the problems addressed by the agreement. Systems for implementation review have been customized to promote progress on both of these aims. For example, the Ottawa Convention established commitments for states parties, on the one hand, to ban or severely restrict the production, use, and international transfer of anti-personnel mines and, on the other, to mobilize international programs to clear such mines from conflict-affected areas where they pose risks to local populations. The first aim requires SIRs involving national arms-licensing authorities, customs and military authorities, as well as the associated civilian governance and oversight bodies, while the second requires SIRs involving officials and experts from national, regional, and international aid agencies, civil society and mine clearance organizations. Most humanitarian arms control agreements are presently at risk of being marginalized

amid global geopolitical tensions, but renewed focus on supporting and developing the SIRs associated with them could mitigate the impact on their effectiveness, particularly in areas of severe need.

The Role of SIRs in Developing International Cooperation on the Uses of Outer Space, Cyber Weapons, and Emerging Technologies of Concern

It is widely recognized that the existing international treaties and institutions governing uses of outer space are inadequate to address the many security and safety issues involved. However, efforts to remedy this by negotiating new multilateral agreements have been hindered. Attention has instead shifted to the more limited aim of developing soft law international codes of conduct or good practice. Although there are some excellent proposals and initiatives in this area, it is proving very challenging to even begin negotiating them in the present international geopolitical climate.

See also the chapter by Sarah McKee, p. 55 ff.

A renewed focus on SIRs could be particularly helpful here. One proven advantage of these systems in other issue areas is that they provide mechanisms through which networks of national and international experts from industry and civil society can translate broad, diplomatically agreed principles into more specific good-practice guidelines. Although these international expert views on what constitutes a useful code of conduct initially lack any legal or diplomatic standing, they can gradually become a key international reference point. Even if some states or national officials refuse to formally recognize or participate in the process, its results can still be robust and influential, provided that the process is sufficiently broad-based and technically competent.

It is important to link such SIRs to diplomatically recognized bodies, such as authorized Technical Working Groups. In the case of outer space, these include the UN Technical Working Group under the UN Committee on the Peaceful Uses of Outer Space, as well as its counterparts under the UN Conference on Disarmament agenda on Prevention of an Arms Race in Outer Space, and within its Open-Ended Working Groups. Specific research or advisory group initiatives might be launched by informal coalitions of the willing, the findings of which may then be considered by the official multilateral working groups. This approach is obviously no panacea, but nevertheless merits close attention.

SIRs can connect networks of experts from industry and civil society with diplomatically recognized bodies like Technical Working Groups to help translate broad agreement principles into specific good-practice guidelines.

For similar reasons, creating carefully tailored SIRs appears to be an important focus of efforts to promote international discussions on codes of conduct or good-practice guidelines for addressing difficult issues, such as cyber weapons and the responsible development and use of "emerging technologies," such as artificial intelligence (AI).

Policy Recommendations

The German Foreign Ministry and its European Union and international partners should devote greater attention to the development and more purposeful use of SIRs in their efforts to:

- improve the dynamism and effectiveness of existing multilateral arms control and disarmament agreements, including the CWC, BWC, NPT, and humanitarian arms control agreements; and

- develop international codes of conduct and good-practice guidelines as part of their soft law approaches to promoting international cooperation on safety and security issues in outer space, as well as to limiting the risks posed by the use of cyber weapons and misuse of emerging technologies.

Owen Greene

University of Bradford

1 Victor, D. G., Raustiala, K., & Skolnikoff, E. R. (Eds.). (1998). *The implementation and effectiveness of international environmental commitments: Theory and practice*. MIT Press.

2 Greene, O. (1998a). The systems for implementation review in the ozone regime. In D. G. Victor, K. Raustiala, & E. R. Skolnikoff (Eds.), *The implementation and effectiveness of international environmental commitments: Theory and practice* (pp. 89–136). MIT Press.
Greene, O. (1998b). Implementation review and the Baltic Sea regime. In D. G. Victor, K. Raustiala, & E. R. Skolnikoff (Eds.), *The implementation and effectiveness of international environmental commitments: Theory and practice* (pp. 177–220). MIT Press.
Lanchbery, J. (1998). Long-term trends in systems for implementation review in international agreements on fauna and flora. In D. G. Victor, K. Raustiala, & E. R. Skolnikoff (Eds.), *The implementation and effectiveness of international environmental commitments: Theory and practice* (pp. 57–88). MIT Press.
Greene, O., & Lanchbery, J. (1995). Developing the Framework Convention on Climate Change. In J. B. Poole & R. Guthrie (Eds.), *Verification 1995: Arms control, peacekeeping and the environment* (pp. 357–367). Westview Press.

3 Lentzos, P. (2025). Biological weapons and security threats. In Stockholm International Peace Research Institute (SIPRI), *SIPRI Yearbook 2025* (pp. 289–302). Oxford University Press.
Shearer, M. P., Mui, A. K., Fogarty, A. S., & Vahey, R. A. (2025). The BWC at 50: Envisioning the next 50 years of the Biological and Toxin Weapons Convention. *Health Security, 23*(5), 289–291. https://doi.org/10.1177/23265094251387593

4 Jakob, U. (2025). Chemical weapons and security threats. In Stockholm International Peace Research Institute (SIPRI), *SIPRI Yearbook 2025* (pp. 273–287). Oxford University Press.

Verification

Societal Verification

Against the backdrop of the erosion of formal verification environments and the decline in treaty-based options, societal verification has attracted renewed interest. Societal verification—a concept that assigns non-state actors, including NGOs, think tanks, and researchers, a role in arms control and non-proliferation monitoring using open-source information—holds promise for practical application. Today, non-state experts with the expertise to analyze satellite imagery produce information on specific proliferation hot spots—largely states perceived as Western adversaries—that is then widely disseminated. Based on interviews with state officials and non-state analysts, this contribution provides an overview of the opportunities societal verification activities offer, as well as the practical limitations faced by those implementing the concept.

With the NPT under increasing strain, the JCPOA defunct, and New START having expired in February 2026, formal verification mechanisms are steadily disappearing. For arms control verification—the process of collecting, analyzing, and checking data to help assess whether a state is fulfilling its obligations under an international agreement or treaty—knowledge production using national technical means (NTM) and contractually regulated inspection regimes continues to play a key role. However, analyses by independent non-state actors—including NGOs, think tanks, and independent analysts—using open-source intelligence (OSINT) have increasingly come to the fore, broadening the debate. This information has largely been disseminated through open-access reports and analysis, media statements, and other forms of public engagement concerning states such as North Korea, Iran, and China. While non-state actors do not officially verify arms control agreements or perform actual verification tasks, they do gather verification-relevant information and share it publicly. In democratic states, societal verification has enabled OSINT analysts outside the government–intelligence nexus to contribute to the dissemination of knowledge, using, for example, commercial satellite images, trade data, and shipping information. In doing so, they provide the public with information that complements government accounts, at times even revealing previously undisclosed developments, including Iranian reconstruction activities following U.S. and Israeli strikes and on-the-ground activities in China related to the construction of missile silos.

For a brief introduction to the NPT, JCPOA, and New START, see p. 122 ff.

Although formal verification mechanisms are weakening, independent non-state actors using OSINT are contributing to the growth of open public debates and may also be influencing closed government discussions.

Broadening the availability of reliable information on state nuclear programs—particularly amid growing global power competition and in crisis situations—could raise public awareness, promote a more balanced narrative, and, at times, counter the tendency of citizens to assume worst-case scenarios when forming their opinions. It can also serve as a stepping stone toward mitigating nuclear risks and help pave the way for future formal arms control treaties. The monitoring of North Korean nuclear facilities and tracking of Iranian sites by non-state actors illustrate how satellite imagery

analysis produced by NGOs feeds into public debate and media reporting, while also being used by government officials.[1] However, the availability of these OSINT analyses and the associated media coverage tells us little about their ultimate impact on policy debates or government decisions. This would be a worthwhile avenue for future research.

While nuclear verification serves as the empirical anchor in this contribution, the production of verification knowledge also extends to chemical weapons monitoring, environmental treaty compliance, and humanitarian crisis assessment—fields in which open-source analysis is playing an increasingly important role. Thus, the broader trends analyzed here are, to some extent at least, transferable to other verification domains.

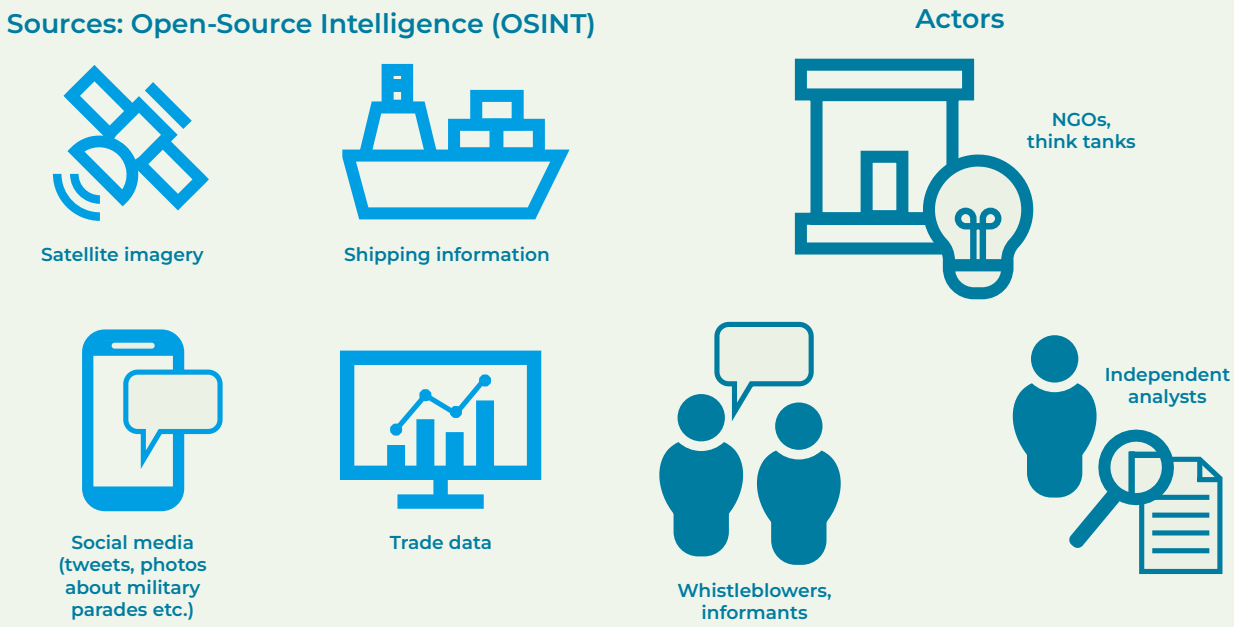
Trajectory of a Concept

As formal treaty regimes erode, societal verification—an interdisciplinary approach that gives non-state actors, such as whistleblowers, scientists, and experts, a role in arms control verification by drawing on open-source intelligence[2]—has acquired increasing practical relevance. Remote open-source monitoring is no longer merely a supplement to official verification; it is increasingly providing valuable verification-relevant information. The literature on societal verification is conceptually rich[3] but empirically thin. Existing studies establish the normative case for involving civil society in verification but fail to systematically map who is currently undertaking this work, what tools they are using, and under what conditions.

Civil experts who draw on OSINT can provide strong support to official channels for arms control verification. However, there is no existing framework to standardize and synergize their findings with those of the intelligence community.

First articulated by Szilard[4] and subsequently developed by Rotblat[5] and others, societal verification posits that nongovernmental actors can and should contribute to arms control verification through OSINT. The objective of societal verification activities is to confirm a state's compliance with its treaty obligations by relying on publicly accessible information. Early conceptual thoughts revolved around two models: "whistleblowing" and "civil society reporting".[6] Today, however, there is a general consensus that the focus has shifted from individual citizens to well-established expert communities performing societal verification tasks.

What is Societal Verification?



Opportunities

There is considerable potential for societal verification and the involvement of experts in detecting and demonstrating the absence of treaty violations in response to allegations against a state. Indeed, there is room for "(…) non-governmental organizations and individuals to engage in tasks associated with treaty verification, including monitoring compliance and supporting investigations with additional data".[7] Non-state OSINT analysts bring information into the public realm that, although not necessarily new to the intelligence community, can offer additional perspectives and assessments on conflicts and proliferation hot spots, complementing those released by the government. This helps to address crucial lacunae and raises citizens' awareness. From the interviews conducted in March 2026, two points stand out that illuminate the impact of verification-relevant knowledge produced through OSINT-based assessments by non-state actors.

Informing the Public

While the intelligence community, by its very nature, does not publish its sensitive findings, open-source analysts do, thereby helping to keep the public informed. They also raise awareness on issues that might not otherwise be at the center of the broader public attention. Indeed, the importance of non-state OSINT-based analyses derives from them being publicly available and accessible to all. It is here that non-state actors bridge the gap by providing expert analysis, which, together with other sources of information—such as journalism, government press releases, and social media—helps to create a more complete picture. This is highly beneficial, as increasing the information available to the public enables them to make more informed decisions, gain additional perspectives and context on particular policies, and develop a broader understanding of a conflict.

OSINT-based analyses increase public awareness and understanding of conflicts by combating misinformation and disinformation and providing expert insights that are both accessible to the public and can be referenced by officials.

Thus, think tanks and expert NGOs have an important role to play here in disseminating informed analyses and providing fact-based, balanced (counter)narratives to those advocating drastic action based on threat-escalation scenarios. Yet the reverse can also occur: analysts may assess conflicts and threat scenarios as more serious than governments do and consequently advocate tougher measures. Although intelligence communities may have reached the same conclusions, perhaps even earlier than those monitoring developments through available OSINT data, current and former government employees acknowledge the importance of the work of think tanks and NGOs and regularly access, read, and draw on their publications. They regularly engage with and request briefings from non-state analysts. Until a year ago, Washington had established communication channels that were routinely used to bring together government officials with non-state OSINT analysts from NGOs and think tanks. However, with the recent cuts to U.S. State Department and think tank funding, the prospects of continuing collaboration look rather dim. This has been a direct and widely felt consequence since the start of Trump's second term, lamented by government and independent actors alike.

Yet, U.S. State Department officials and International Atomic Energy Agency (IAEA) analysts continue to consult and engage with the work of non-state actors. One reason is pragmatic and relates to the types of information government officials are permitted to cite publicly. Here, NGOs and think tanks that track global nuclear arsenals, such as the Federation of American Scientists (FAS) and the Stockholm International Peace Research Institute (SIPRI) provide useful publicly accessible estimates. The *Status of World Nuclear Forces*[8] published by SIPRI, for example, provides detailed figures on developments in military stockpiles. That said, even within government, discussions are necessary when reports and analyses by non-state actors emerge, as not everyone has the same level of insight or security clearance. Public reporting, on the other hand, is something everyone working inside government can read.[9]

Stimulating Domestic Debate

Findings published by non-state actors can stimulate public discussion on topics otherwise confined to intelligence communities and empower citizens to engage in policy-relevant debates.

Reports and analyses by U.S. non-state actors are opening up domestic debate. If these actors publish something—for example, on Iran, Chinese silo fields, or the return of U.S. missiles to Lakenheath in the UK—in media outlets such as the *New York Times* or the *Washington Post*, this can spark a discussion among a range of stakeholders. While the intelligence community and some government agencies were already aware of China's silo fields, it was only after the Federation of American Scientists published its findings that the issue entered the broader domestic debate.[10] Analyses by non-state actors can bring issues to the fore that might otherwise have gone undiscussed, thereby prompting a much more open debate. In this particular case, it also had repercussions in China with the Chinese Foreign Affairs Ministry releasing a statement in response to the FAS numbers.[11] A non-state analyst's work might resonate with different segments of the broader population, while their analyses and statements might also elicit a response from the government. Whether or not this always makes a useful contribution to curbing nuclear-related threats is debatable, however, since back-channel diplomatic efforts may be underway in parallel without public knowledge. In such cases, the untimely release of OSINT analysis by non-state actors can undermine bilateral efforts by turning public opinion against concessions to a state perceived to be in breach of an agreement. Yet, societal verification activities illuminate important topics that would otherwise, in all likelihood, have remained exclusively within the domain of classified government discourse. Although, as noted above, this may not always be beneficial—particularly with regard to the timing of OSINT releases—in democracies, it undoubtedly helps citizens to hold their respective governments accountable. This also concerns domestic issues, such as plutonium pit production in the U.S. and NGOs' monitoring of the associated environmental impacts.[12] It thereby provides a domestic forum for open debate accessible to all. This is a highly salient point, as non-state OSINT analyses extend the conversation, previously confined to closed government circles, to a broader public, including voters. This, in turn, may indirectly influence the direction of policies in the making.

Limitations

OSINT-based analysis does come with limitations: asymmetries that bias and imbalance narratives, findings that are more indicative than conclusive, and reports that lack structured impact on official outcomes.

In a convincing account, Al-Sayed sets out the potential asymmetries inherent in publicly available information focusing on verification, which was often instrumentalized to advance "US and Western strategic interests".[13] At the same time, publicizing such information risks reproducing preconceived normative stereotypes about countries perceived to be Western adversaries, such as North Korea, Iran, and China. Moreover, it could create a "hype" around specific issues that would subsequently consume a disproportionate share of public attention. Kretzschmar et al.[14] thus caution that societal verification is far from unbiased and initiatives employing the concept should be aware of unintended consequences as the concept comes "with its own set of uncertainties".[15]

Besides questions of interpretive authority, there are also practical limitations. Several authors stated that societal verification efforts rarely produce "smoking gun" evidence. As held by Hinderstein and Hartigan,[16] in most instances, "(…) these tools are likely to convey information that is only indicative of potential violations or misuse" and can thus serve as no more than supplemental information. The concept, therefore, is best understood as complementary to state-level and international agency mechanisms, broadening the toolbox of available approaches and enhancing their capacity to alleviate suspicions and confirm violations.[17]

Lastly, there is a structural obstacle concerning the integration of the data produced into existing and traditional information streams for arms control verification. Currently, practitioners of the concept face a gap: they possess the tools to detect treaty

violations, but the data they gather does not, at least officially, inform official government analyses. Put differently, NGO experts lack official protocols to publicize their observations without taking the public by surprise, releasing reports that could impact both bilateral and multilateral relations, potentially distorting the debate.

Outlook

Analyses drawing on non-state OSINT produced by expert communities and NGOs fulfill an important role. By disseminating information about developments surrounding global nuclear stockpiles, illicit construction activities, and the acquisition of sanctioned military equipment, they provide "(…) much needed transparency and data verification functions".[18] Open-source analysis enriches the domestic debate by providing citizens with new information that complements government accounts.

Open-source analysis adds transparency and completeness to public debate. Governments should view it as a complementary tool and a sign of democratic strength. Open dialogue is necessary for its structured incorporation into treaty verification regimes.

However, access to the satellite imagery that non-state analysts rely on in their work has been noticeably restricted. In March 2026, for instance, U.S. satellite imagery providers Planet Labs and Vanator imposed an indefinite blackout on high-resolution images of the Middle East. Providers announced that this blackout followed consultations with the U.S. Government and was intended to prevent information about U.S. military personnel in the region from falling into the hands of its adversaries. It remains to be seen whether these blackouts set a precedent.

In democratic societies, states must accept that information is continuously released to the public by actors outside the government. This is simply a reflection of the technical possibilities available today. What is also needed going forward is improved dialogue between the government and non-state actors using OSINT to broaden the debate, thereby enabling policymakers to make better use of information generated outside the official government–intelligence nexus. These two streams of information should not be seen as competing with one another, but ideally as complementary, working together for the benefit of the public. Moreover, OSINT is already used as one of several sources supporting the IAEA's safeguards efforts and is permitted under Comprehensive Nuclear-Test-Ban Treaty (CTBT) rules. In sum, drawing on OSINT-based analyses and information produced outside the government–intelligence nexus within treaty regimes could, arguably, help address the current impasse in non-proliferation and arms control by broadening the discussion to include civil society perspectives.

For a short explainer on the CTBT, see p. 122.

Robin E. Möser

PRIF – Peace Research Institute Frankfurt

1 Author interviews with current and former U.S. State Department officials. The underlying research for this chapter was made possible by the VeSPoTec Hub (*Verification in a complex and unpredictable world: social, political and technical processes*) research consortium, which provided both financial support and a stimulating forum for exchanging and discussing initial ideas.

2 Al-Sayed, S. (2022). Revisiting Societal Verification for Nuclear Non-Proliferation and Arms Control: The Search for Transparency. *Journal for Peace and Nuclear Disarmament, 5*(2), 496–506.
Blackaby, F. (1998). Societal verification. In J. Rotblat (Ed.), *Nuclear weapons: The road to zero*. Routledge.
Rotblat, J. (1992). Towards a nuclear weapon-free world: Societal verification. *Security Dialogue, 23*(4), 51–61.

3 Al-Sayed (2022)
Gastelum, Z. N. (2013). *Societal verification for nuclear nonproliferation and arms control*. Sandia National Laboratories.
Al-Sayed, S., Glaser, A., & Mian, Z. (2023). Societal Verification of Nuclear Disarmament in the 21st Century: A Workshop Report. *Journal for Peace and Nuclear Disarmament, 6*(2), 365–75.
Hartigan, K., & Hinderstein, C. (2013a). Societal verification: Intellectual game or international game-changer? *ESARDA Bulletin, 49*.

4 Szilard, L. (1946). Can we avert an arms race by an inspection system? In D. Masters & K. Way (Eds.), *One world or none: A report to the public on the full meaning of the atomic bomb* (pp. 167–179). New Press.

5 Rotblat. (1992)

6 Al Sayed et al. (2023)

7 Revill, J., & Garzón Maceda, M. (2024). The role of open source data and methods in verifying compliance with weapons of mass destruction agreements. In J. Revill & M. Garzón Maceda (Eds.), *Open source investigations in the age of Google* (pp. 241–258).

8 Kristensen, H., Korda, M., Johns, E., Knight-Boyle, M., & Kohn, K. (2025). *Status of world nuclear forces*. Federation of American Scientists. https://fas.org/initiative/status-world-nuclear-forces/

9 Author interviews with current and former U.S. State Department officials

10 Kristensen, H., & Korda, M. (2021). China's nuclear missile silo expansion: From minimum deterrence to medium deterrence. *Bulletin of the Atomic Scientists*. https://thebulletin.org/2021/09/chinas-nuclear-missile-silo-expansion-from-minimum-deterrence-to-medium-deterrence/

11 Interview with Matt Korda. (2026, April)

12 Asplund, C. D., & von Hippel, F. (2023). Dealing with a debacle: A better plan for US plutonium pit production. *The Bulletin of the Atomic Scientists*. https://thebulletin.org/2023/04/dealing-with-a-debacle-a-better-plan-for-us-plutonium-pit-production/

13 Al-Sayed (2022)

14 Kretzschmar, S., Ostermann, L., & Rademacher, L. (2026). Verification as socio-epistemic challenge: the case of nuclear archaeology for nuclear disarmament verification. *Zeitschrift für Friedens- und Konfliktforschung*. https://doi.org/10.1007/s42597-026-00157-w

15 Kretzschmar et al. (2026), p.18

16 Hartigan, K. L., & Hinderstein, C. (2013b). *The opportunities and limits of societal verification*. Nuclear Threat Initiative, Institute of Nuclear Materials Management. https://media.nti.org/pdfs/The_Opportunities_and_Limits_of_Societal_Verification.pdf

17 Podvig, P., & Decker, E. (2022). The role of open-source data in verification. In P. Podvig (Ed.), *Exploring options for missile verification*. UNIDIR.

18 Kristensen, H., & Korda, M. (2024). Monitoring nuclear weapons developments with open source intelligence. In J. Revill & M. Garzón Maceda (Eds.), *Open source investigations in the age of Google* (pp. 66–86).

TRENDS

Artificial Intelligence

Artificial Intelligence and Semiconductors—Key Developments Since 2025

Semiconductors are no longer merely an economic concern; they have increasingly become a matter of national and international security. Competition over artificial intelligence (AI) capabilities is closely tied to control over semiconductor supply chains, critical raw materials, and the infrastructures required to support advanced computing. Strengthening resilience across semiconductor ecosystems—including manufacturing capacity, energy provision, and international partnerships—has therefore become a key strategic objective for states and firms alike.

Since the publication of the previous CNTR Monitor, several developments have reinforced and, in some areas, intensified these structural dynamics. The past year has been characterized by three overarching trends: the geopoliticization of critical raw materials, new bottlenecks in AI hardware supply chains, and the continued fragmentation of global technology governance. At the same time, the rapid expansion of AI infrastructure has created new strategic vulnerabilities related to energy supply, data center capacity, and the geographic concentration of advanced semiconductor manufacturing.

Critical Minerals as Strategic Leverage

Critical minerals are increasingly being used as strategic tools for geopolitical leverage, with competition evolving beyond just chip fabrication and processing capabilities.

One of the most notable developments is the increasing use of critical minerals as instruments of geopolitical leverage. In recent years, China, which controls a dominant share of global rare earth mining and processing capacity,[1] has expanded export controls and licensing requirements for several rare earth elements and related processing technologies that are essential for semiconductor manufacturing, advanced electronics, and defense systems. In late 2025, a conflict arose that illustrates this impact. The Dutch government intervened in Nexperia, a semiconductor company headquartered in the Netherlands, amid concerns that the firm's Chinese owner, Wingtech, might transfer strategic technology and production capabilities from Netherlands to China, raising national and European economic security concerns. In response, the Chinese Ministry of Commerce imposed export restrictions on Nexperia chips manufactured or packaged in China, causing almost immediate disruptions to global semiconductor supply chains. The dispute had a direct effect on other countries, including Germany, where many manufacturers rely on Nexperia components for automotive and industrial electronics.

This development illustrates that strategic competition over semiconductor technologies extends well beyond chip fabrication itself. Control over upstream inputs is becoming an increasingly important component of technological power. As a result,

supply chain resilience in the semiconductor sector must now also account for vulnerabilities in the extraction, processing, and trade of critical minerals.

Emerging Bottlenecks in AI Hardware

The rapid expansion of AI infrastructure has created new structural bottlenecks in the semiconductor ecosystem. While earlier policy debates focused primarily on the supply of advanced logic chips, the recent surge in AI data center construction has revealed shortages in other critical components. For instance, modern computing systems rely on random-access memory (RAM) as fast, temporary working memory enabling processors to access and manipulate data during computation. The dominant form of RAM today is dynamic random-access memory (DRAM). Artificial intelligence systems require a specialized high-performance variant of this known as high-bandwidth memory (HBM), which provides the exceptionally high data transfer rates needed by AI accelerators. As demand for HBM has surged, it has placed increasing pressure on the broader DRAM ecosystem and manufacturing capacity. These developments have also affected global memory supply chains more generally, including those for NAND flash technologies, which provide the nonvolatile, long-term data storage used, for example, in USB sticks or memory cards. Consequently, tightening conditions in memory markets are increasingly affecting a wide range of digital products, including smartphones, personal computers, storage systems, and industrial electronics. In practice, this means that the strategic competition over AI hardware is now increasingly extending beyond the production of leading-edge processors to encompass memory technologies, advanced packaging, and specialized semiconductor manufacturing processes, with growing repercussions for other IT sectors.

Rapidly expanding AI infrastructure is now affecting global memory supply chains, adding to the pressures on advanced logic chip production.

Fragmentation of Technology Governance

Another notable development is the increasing volatility of technology governance regimes. Export control frameworks targeting advanced AI chips and semiconductor manufacturing equipment remain a central instrument of geopolitical competition, particularly in the context of the U.S.–China technology rivalry. However, these regulatory frameworks have become more dynamic and politically contested.

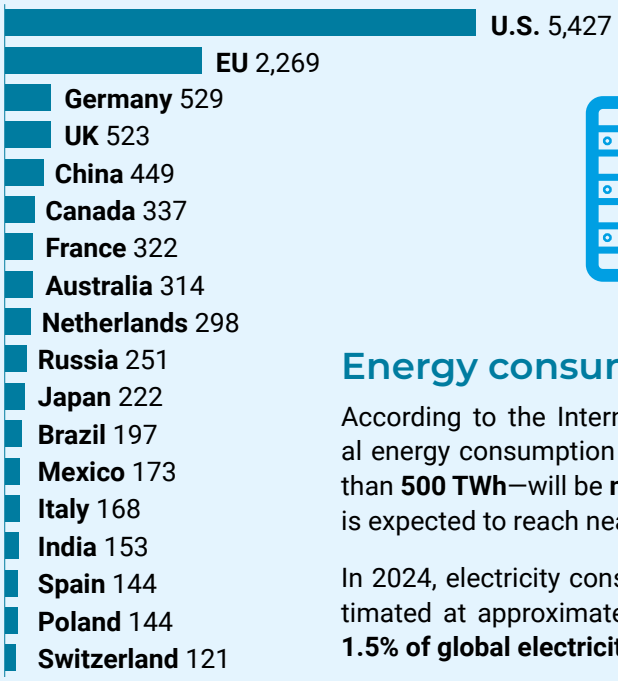
Changes to export licensing regimes, tightening restrictions on semiconductor equipment, and more comprehensive investment screening mechanisms have created a policy environment characterized by frequent adjustments and regulatory uncertainty. For governments and companies alike, this complicates long-term planning for semiconductor manufacturing investments, AI infrastructure development, and technology partnerships.

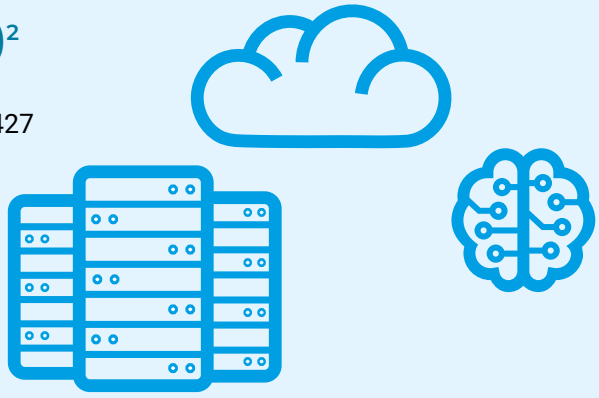
At the same time, concerns have emerged over the effectiveness of Western export controls as a tool for constraining China's economic AI developments and ambitions. In the United States, in particular, policymakers and analysts have increasingly acknowledged that export restrictions have not fully prevented Chinese firms from advancing their AI capabilities and competing effectively with U.S. companies, raising doubts about the long-term efficacy of their containment-oriented approaches. More broadly, recent debates have centered on whether export controls can sustainably preserve technological advantages or whether they may instead risk accelerating domestic innovation and technological self-sufficiency in the very states they target. Moreover, differences in regulatory approaches among major technology powers have complicated efforts to establish coordinated governance frameworks, increasing the risk of a more fragmented global technology landscape characterized by competing standards, regulatory regimes, and technology ecosystems.

While export control frameworks remain a central instrument of geopolitical competition, there are doubts about their effectiveness in preserving technological advantages and the risk of fragmenting the global technology landscape.

The AI Data Center Boom

Number of data centers (2025)[2]





Energy consumption

According to the International Energy Agency, the global energy consumption of data centers in 2026—at more than **500 TWh**—will be **more than double** that of 2020 and is expected to reach nearly **1,000 TWh by 2030**.[3]

In 2024, electricity consumption by data centers was estimated at approximately 415 TWh, which equals about **1.5% of global electricity consumption**.[4]



Investment plans

$700 billion

Alphabet, Microsoft, Meta and Amazon are expected to spend nearly this amount combined in 2026 on their AI build-outs.[5]

$500 billion

The **U.S. public-private initiative Stargate Project** (in partnership with OpenAI, SoftBank, Oracle, and MGX) plans to invest this amount over 4 years.[6]

$295 billion

has been announced as **China's** investment in state-run AI data centers over five years, to create a nationwide "National Computing Network".[7]

€30 billion

The **EU** is putting seven large data centers out to bid and plans to invest €10 billion to attract an additional €20 billion in private investment.[8]

$13.4 billion

is the total value of data center projects in **Russia** at various stages of planning; however, Russia has suspended 38 projects worth $2.26 billion combined.[9]

How reliable are the numbers?

The funding and investment figures presented here are only directly comparable to a limited extent and should be understood **primarily as an indicator of the strategic importance** that countries attach to AI technologies. In the global AI race, such announcements are often driven by industrial, economic, and geopolitical considerations. In particular, leading AI nations such as the U.S. and China also use investments strategically to secure and expand existing technological, economic, and infrastructural competitive advantages or positions of dominance. Furthermore, the measures **differ significantly in terms of their structure**—such as direct government investments, funding programs, and public-private partnerships—as well as their **objectives**, for example, the expansion of data centers, research funding, or the development of broader AI ecosystems.

Meanwhile, there is evidence to suggest that Chinese companies have responded strategically by stockpiling AI-related hardware and semiconductor components in anticipation of future restrictions. This accumulation of inventories has helped them to mitigate the immediate impact of supply constraints, while reducing the short-term effectiveness of tightening export control measures.

The Rise of AI Compute as Strategic Infrastructure

A further development is the growing recognition that AI compute infrastructure is itself becoming a strategic asset. The rapid expansion of large-scale AI systems has led to unprecedented investment in hyperscale data centers, specialized AI computing clusters, high-performance networking technologies, and even submarine cables.

AI computing infrastructure is quickly becoming a strategic asset, alongside energy systems and telecommunications networks.

These infrastructures are increasingly comparable to other forms of strategic infrastructure, such as energy systems or telecommunications networks. For states, control over large-scale AI compute infrastructure provides several strategic advantages. It enables the domestic development and deployment of advanced AI systems and reduces dependence on foreign technology providers, thereby strengthening the country's ability to capture the economic benefits associated with AI innovation. Access to large-scale compute capacity has become an important determinant of technological sovereignty, as it shapes who can develop, deploy, and commercialize frontier AI applications. As a result, compute infrastructure is increasingly viewed not only as an economic asset but also as a strategic capability, prompting governments to pay greater attention not only to semiconductor production but also to the geographic location, ownership, and resilience of AI data center infrastructure. In addition, the use of AI is expanding rapidly throughout society, making access to reliable AI services more and more important for the functioning of public services and digital end-user infrastructures. As a result, disruptions affecting access to AI capabilities or compute infrastructure may have consequences that extend beyond individual firms, increasingly affecting broader societal resilience.

Because AI infrastructure requires large and reliable supplies of electricity, policy debates are increasingly linking AI strategy with energy security.

The growth of AI infrastructure is also placing new pressures on energy systems and physical infrastructure. Training and operating advanced AI models requires vast computing resources, which, in turn, depend on large and reliable supplies of electricity. The rapid construction of hyperscale data centers—particularly in the United States, China, and parts of the Middle East—has already begun to affect regional electricity demand and grid planning. This development introduces a new dimension to AI geopolitics: access to abundant, reliable, and affordable energy may become an important determinant of future AI capabilities. In several countries, policymakers are now increasingly linking AI strategy with broader debates on energy security, grid expansion, and industrial electricity pricing.

Conclusion

The geographic concentration of key segments of the semiconductor supply chain remains a major structural vulnerability. In particular, the production of the most advanced chips continues to be heavily concentrated in East Asia, especially Taiwan. While global investment in semiconductor manufacturing capacity has increased significantly, new facilities will take years to reach full operational capacity and are unlikely to fully offset existing concentration risks in the near term.

For Europe, these developments underscore the persistence—and in some areas the deepening—of structural dependencies within the semiconductor ecosystem. Despite policy initiatives aimed at strengthening domestic manufacturing capabilities, Europe

Europe is directly exposed to disruptions arising from geopolitical tensions due to its heavy reliance on external suppliers for both AI chips and raw materials.

remains heavily reliant on external suppliers for both advanced AI chips and critical raw materials. At the same time, European industry is directly exposed to supply disruptions arising from geopolitical tensions and export restrictions in upstream segments of the semiconductor supply chain. These dependencies create not only economic vulnerabilities but also broader security risks, as disruptions to semiconductor supply chains can affect critical industries, digital infrastructures, and key sectors of modern economies. The ability of states to influence access to critical minerals, semiconductor components, manufacturing capacity, or AI infrastructure increasingly creates opportunities for geopolitical leverage.

Taken together, these developments reinforce the central conclusion of the previous CNTR Monitor:[10] the competition over artificial intelligence capabilities is increasingly shaped by control over semiconductor ecosystems and their associated supply chains. This competition spans the entire value chain—from critical minerals and semiconductor manufacturing equipment to advanced chips, memory technologies, energy-intensive AI computing infrastructure, and hyperscale data center capacity. Consequently, resilience within semiconductor ecosystems is becoming an increasingly important element of economic security, technological sovereignty, and the protection of critical infrastructures.

Even if expectations are not met and investment dynamics slow, the strategic importance of semiconductors and AI technologies is likely to remain.

At the same time, some analysts have raised concerns about the sustainability of the current levels of AI-related investment. Should expectations regarding future AI adoption or profitability fail to materialize, a correction in infrastructure investment could follow, with implications for semiconductor demand, data center construction, and capital allocation across the technology sector. Nevertheless, even if current investment dynamics were to slow, the strategic importance of semiconductors and AI-related technologies—and the dependencies associated with them—would likely remain unchanged.

Thomas Reinhold

PRIF – Peace Research Institute Frankfurt

1 An estimated 85–90 percent of the rare earth materials used in global semiconductor and advanced electronics supply chains are processed or refined in China, accounting for around 60 percent of global rare earth mining output and roughly 90 percent of rare earth processing and refining capacity. Source: Baskaran, G., & Schwartz, M. (2025). *Developing Rare Earth Processing Hubs: An Analytical Approach*. Center for Strategic & International Studies. https://www.csis.org/analysis/developing-rare-earth-processing-hubs-analytical-approach#h2-identifying-priority-jurisdictions-for-rare-earth-processing-hubs

2 Yanatma, S. (2026, April 27). *Which country in Europe has the most data centres driving the AI boom?*. EuroNews. https://www.euronews.com/next/2026/04/27/which-country-in-europe-has-the-most-data-centres-driving-the-ai-boom

3 International Energy Agency. (2025, April 10). *Global data centre electricity consumption, by equipment, Base Case, 2020-2030*. https://www.iea.org/data-and-statistics/charts/global-data-centre-electricity-consumption-by-equipment-base-case-2020-2030

4 International Energy Agency. *Energy demand from AI*. https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai

5 Elias, J. (2026, February 6). *Tech AI spending may approach $700 billion this year, but the blow to cash raises red flags.* CNBC. https://www.cnbc.com/2026/02/06/google-microsoft-meta-amazon-ai-cash.html

6 Jacobs, J. (2025, January 22). *Trump announces up to $500 billion in private sector AI infrastructure investment*. CBS News. https://www.cbsnews.com/news/trump-announces-private-sector-ai-infrastructure-investment/

7 Riethmüller, C. (2026, June 11). *China plans nationwide AI data center network for $295 billion*. Heise. https://www.heise.de/en/news/China-plans-nationwide-AI-data-center-network-for-295-billion-11328976.html

8 López, S. (2026, July 30). *EU-Milliardenprojekt für sieben KI-Gigafactories*. Tagesschau. https://www.tagesschau.de/wirtschaft/digitales/eu-milliardenprojekt-ki-gigafactories-100.html

9 The Moscow Times. (2026, June 24). Russia Puts $2.3Bln of Data Center Projects on Hold Amid High Borrowing Costs and Energy Grid Limits. https://www.themoscowtimes.com/2026/06/24/russia-puts-23bln-of-data-center-projects-on-hold-amid-high-borrowing-costs-and-energy-grid-limits-a93090

10 Reinhold, T. (2025). Artificial intelligence, semiconductors, and the chip wars: Reviewing the geopolitics of AI in the military and in international security. In Göttsche, M., Reis, K., & Daase, C. (Eds.). (2025). *New Realities of AI in Global Security. CNTR Monitor – Technology and Arms Control 2025* (pp. 14–21). PRIF – Peace Research Institute Frankfurt. https://doi.org/10.48809/cntr2025en

Artificial Intelligence & Biology

AI in the Lab: From Biological Design to Biological Access Control

This chapter examines where meaningful control can be exercised as artificial intelligence (AI) and the life sciences converge more closely. It examines how the policy debate has shifted beyond general concern about biological design tools toward more specific questions about access to sensitive data, screening of synthetic nucleic acid orders, safeguards in frontier AI deployment, and the limits of current evaluation methods. These choices matter because they will influence how AI-enabled biological capability is governed. The chapter points to the need for targeted, layered governance, especially through managed access, stronger screening, better evaluation, and investment in implementation capacity.

The convergence of AI and the life sciences has continued to accelerate, while the policy debate around these tools has become increasingly concrete. Over the past year, the discussion has shifted from broad concern about AI-enabled biological design toward questions of who should be able to access these tools, under what conditions, and subject to what forms of oversight.

Regulated Access

Recent work has highlighted a series of practical interventions, including access to sensitive pathogen data, managed access to high-risk biological AI tools, improved screening of synthetic DNA orders, and tighter safeguards for frontier AI model deployment.

Growing attention is being paid to governance of the research pipeline in an effort to preserve scientific value while mitigating the risks of deliberately harmful applications.

Last year's CNTR Monitor[1] discussed how biological design tools were expanding the possibilities of protein engineering, diagnostics, and related areas, while also raising concerns about misuse. That underlying picture has not changed. AI systems continue to improve researchers' ability to predict protein structure, generate candidate biomolecules, and assist with specialized biological tasks. What has changed, however, is the level of attention now being paid to governance. A growing number of researchers, policy organizations, and AI companies are asking which parts of the pipeline should remain broadly open and which may require tighter control. The question has moved from simply whether these tools are dual-use to how their scientific value can be preserved, while reducing the risk of them being used to support deliberately harmful applications.

One of the clearest shifts in emphasis is toward managed access. In January 2026, NTI | bio published a *Framework for Managed Access to Biological AI Tools*[2] arguing that access to certain biological AI tools should be tiered according to their risk level

and limited to validated users where appropriate. Access is increasingly understood as existing on a spectrum, with lower-risk tools being more widely available and higher-risk systems being subject to tighter controls, potentially even requiring identity checks, monitoring, and usage restrictions. The same logic is increasingly being applied to commercial products and service-based platforms, which offer relatively open access to capabilities, while keeping the underlying model weights and training data out of public reach. Just as important, the NTI framework emphasizes that security measures should be balanced with equitable access, ensuring that safety does not become a pretext for concentrating advanced scientific capability in the hands of a small number of institutions or countries.

Governance efforts are increasingly focusing on selective control points, particularly managed access to higher-risk biological AI tools and tighter restrictions on sensitive pathogen data, while keeping most lower-risk research uses open.

A related policy development concerns the data used to train biological AI systems. A policy article published in *Science* in February 2026[3] argued for tailored access restrictions on a narrow category of pathogen data that may pose special AI-related misuse risks, while preserving open access in almost all other domains. This underscores the fact that biological data are not all alike. Large volumes of sequence data are already in the public domain and have long supported legitimate research. More sensitive are functional datasets that connect pathogens or pathogen variants to traits such as transmissibility, virulence, immune evasion, or host range. These are the kinds of data most directly relevant to high-risk biological design. The resulting policy direction is narrower than a general call for secrecy, yet more consequential than earlier voluntary norms. This suggests that the debate is beginning to take a more concrete policy form.

DNA Synthesis Screening

One of the most important governance control points remains the synthesis of nucleic acids.[4] A 2025 *Science* paper by Wittmann and colleagues[5] tested whether open-source AI protein design tools could generate variants of proteins of concern that would evade the screening systems used by commercial DNA synthesis providers. They identified a real vulnerability: in some cases, AI-redesigned sequences could evade existing detection methods. Working with synthesis companies, the researchers then developed and deployed patches that greatly improved detection of modified sequences likely to retain the function of the original protein. The paper shows both that AI-assisted biological design can create new screening challenges and that these methods must evolve alongside generative design tools.

Functional screening has also received more attention.[6] Rather than relying solely on taxonomic or homology-based screening (which identifies sequences based on similarity to known organisms or previously characterized genes), functional approaches seek to identify sequences based on the harmful biological functions they encode, such as tissue damage, immune evasion, or host-cell attachment. This could make screening more robust to recoding or other sequence changes that preserve function, although such methods remain harder to implement at scale and still depend on well-curated functional databases.

While nucleic acid screening has often been treated mainly as a technical problem, it also raises questions of implementation. A 2025 analysis argues that the field still lacks agreed mechanisms to verify whether providers are actually implementing effective screening practices.[7] This creates economic costs for the firms responsible, makes screening claims difficult to verify, and leaves regulators hesitant to impose stronger requirements in the absence of clear standards. Another unresolved issue is how screening performance should be audited, since future regulation will require agreed standards for comparing platforms and verifying compliance.[8] A proposed pathway toward enforceable screening standards centers on processes led by industry but supported by regulatory oversight. This includes clearer definitions of sequences of concern (SOCs), third-party testing, certification, and a phased transition from volun-

Sequence of Concern (SOC): A DNA or RNA sequence associated with a harmful pathogen, toxin, or biological function that could pose a biosecurity risk.

Synthetic nucleic acid screening remains one of the most practical intervention points, but it requires stronger standards, auditing, and implementation capacity, as well as technical improvements.

tary adoption to enforceable compliance. This approach frames synthesis screening as a standards and governance issue, rather than simply a technical safeguard.

Stronger oversight can still fail in practice if institutions lack the staff, screening capacity, and review systems needed to carry it out effectively. Ambiguous definitions, fragmented regulatory triggers, and underdeveloped institutional capacity could produce oversight that appears rigorous on paper, while delivering limited additional security in practice. In February 2026, a bipartisan U.S. bill proposed mandatory screening of gene synthesis orders and customers, showing how quickly this issue has moved from expert recommendation to proposed legislation.[9]

In June 2026, calls for mandatory screening gained further momentum when the leaders of major AI firms joined an open letter urging the U.S. Congress to require screening of synthetic DNA and RNA orders. At the same time, IBBIS and other organizations continued to develop practical implementation tools and standards, including the Commec screening tool, the DNA Synthesis Screening Consortium, and work related to ISO 20688-2. This also highlights the need for international coordination, since differences in national standards and reporting rules weaken the effectiveness of screening.

Actions by Frontier AI Companies

Measures adopted by AI companies themselves should be distinguished from governance measures imposed by, or coordinated through, public institutions. The first concern internal model development and deployment practices, such as refusals, monitoring, preparedness frameworks, and access controls. The second concern laws, standards, audits, and other forms of external oversight. The aforementioned open letter urging Congress to mandate DNA synthesis screening illustrates how the two can intersect, as it involves companies calling for external governance rather than only expanding their own internal safeguards.[10]

Frontier AI companies have also begun to describe the risk of biological misuse in more operational terms. In May 2025, Anthropic announced that it was deploying Claude Opus 4 with AI Safety Level 3 protections as a precautionary response to risks associated with chemical and biological misuse.[11] OpenAI has similarly stated that it expects upcoming models to reach high levels of capability in biology under its Preparedness Framework,[12] and has outlined a combination of mitigation measures, including model behavior controls, monitoring, and preparedness measures focused on biological risks. Many questions remain about the effectiveness of refusals, jailbreak resistance, and misuse detection. These developments show that at least some leading firms now treat biological risk as a live deployment issue rather than merely a distant possibility. This shift matters because it moves the conversation from abstract principle to concrete operational practice.

Current biorisk evaluations do not always capture real-world misuse potential, since many standard benchmarks saturate quickly and may not reflect the physical and organizational bottlenecks involved in developing biological weapons. This has resulted in renewed calls for more transparent evaluations and more realistic studies, including physical uplift trials, which better test how model performance translates into practice.[13]

Evo 2 is a genomic foundation model capable of generalist prediction and design tasks across DNA, RNA, and proteins.

One question is how well safeguards hold up when biological models are released in a form that others can adapt. A 2026 paper examined whether excluding certain data from training remains effective under those conditions.[14] Using Evo 2 as a case study, they found that fine-tuning the model on a relatively small set of harmful human-infecting viruses could partly restore capabilities relevant to misuse. Data filtering can help prevent misuse, but in open models this remains vulnerable to circumvention via fine-tuning and thus serves better as a single layer of protection than as a complete safeguard.

Evaluation and access control have therefore become more central. The main question is how general-purpose reasoning models, biological foundation models, and automated laboratory systems may eventually be combined into integrated workflows. Models are becoming better at navigating specialized scientific domains. Recent governance proposals therefore assume that effective safeguards must be built into the entire pipeline, from training data and model access to sequence synthesis and downstream monitoring. A RAND analysis argues that standard biology benchmarks capture only part of the picture.[15] More relevant risks may emerge when AI agents are capable of performing multistep workflows that involve planning, tool use, adaptation, and interaction with biological resources. The central question is how a model performs in more realistic design-build-test-learn tasks. Evaluation, therefore, plays a governance role, as decisions about model access, deployment, and safeguards increasingly depend on how these capabilities are measured and interpreted.

Frontier AI companies are beginning to treat biological misuse as a live deployment issue, but significant uncertainty remains about how well current safeguards and evaluations capture real-world risk.

A more cautious RAND Delphi study on the limits of AI-enabled pathogen design found broad agreement that, in the near term, AI should still be understood mainly as an assistive tool rather than an autonomous designer of novel pathogens.[16] Experts engaged in the study emphasized that AI already helps specialists with certain aspects of bioengineering and related work, but that successful translation from digital design to functional biological constructs still depends heavily on laboratory capacity, experimental validation, and high-quality data. This suggests that the most immediate concern is the growing ability of AI systems to accelerate or augment already capable actors, rather than fully autonomous pathogen invention.

The main governance question is therefore situated at the interfaces between digital design, physical synthesis, and institutional oversight. The policy challenge is to decide which forms of openness serve science and society, and which require more structured control. Most biological research data and lower-risk AI tools should remain openly accessible, but a narrower category of high-risk data and high-risk tool access may require verification, monitoring, and managed release. Balancing innovation and security will require a layered approach, combining technical safeguards, managed access, stronger screening, capability evaluations, and public governance that is both internationally compatible and attentive to inequity. As access to sensitive tools and data becomes more regulated, decisions on who gets access and under what conditions will need to be transparent and politically accountable. The tools will also need to be matched to actual institutional capacity, so that new oversight requirements do not become brittle, merely symbolic, or unevenly applied in practice.

Policy Recommendations

1. The EU should establish binding requirements for nucleic acid screening and customer verification across member states. Support the introduction of binding EU requirements for sequence screening and customer verification for synthetic nucleic acid orders, rather than relying on uneven voluntary uptake. The focus should be on common standards, third-party testing, and mechanisms that can keep pace with AI-designed sequence variants.

2. Support managed access to high-risk biological AI tools and datasets. Back national and EU-level frameworks that keep most biological data and lower-risk tools openly accessible, while requiring validated access, monitoring, and clearer governance for higher-risk pathogen data and biological AI systems.

3. Invest in implementation capacity. Any new screening or access-control requirements should be matched by funding for institutional implementation, including biosafety staff, screening tools, review capacity, and guidance for universities and research institutes.

4. Support international coordination on biosecurity standards for DNA synthesis and biological AI tools. Promote internationally compatible standards for nucleic acid screening, managed access, reporting channels, and evaluation practices, ensuring that governance does not remain fragmented across jurisdictions, facilitating circumvention.

Kadri Reis

PRIF – Peace Research Institute Frankfurt

1 Reis, K. (2025). AI in the lab: The future of biological design. In Göttsche, M.; Reis, K., & Daase, C. (Eds.). (2025). *New Realities of AI in Global Security. CNTR Monitor – Technology and Arms Control 2025* (pp. 33-36). PRIF – Peace Research Institute Frankfurt. https://doi.org/10.48809/cntr2025en

2 Carter, S., & Butchello, G. (2026). *A Framework for Managed Access to Biological AI Tools.* NTI. https://www.nti.org/analysis/articles/a-framework-for-managed-access-to-biological-ai-tools/

3 Bloomfield, D., Black, J. R. M., Crook, O., Brandes, N., Hanke, M. S., Inglesby, T. V., Cicero, A., Pollack, R., Hernandez-Boussard, T., Imperiale, M. J., Altman, R. B., & Pannu, J. (2026). Biological data governance in an age of AI. *Science, 391*(6785), 558–561. https://doi.org/10.1126/science.aeb2689

4 Reis, K. (2025). DNA synthesis screening. In Göttsche, M.; Reis, K., & Daase, C. (Eds.). (2025). *New Realities of AI in Global Security. CNTR Monitor – Technology and Arms Control 2025* (pp. 72-76). PRIF – Peace Research Institute Frankfurt. https://doi.org/10.48809/cntr2025en

5 Wittmann, B. J., Alexanian, T., Bartling, C., Beal, J., Clore, A., Diggans, J., Flyangolts, K., Gemler, B. T., Mitchell, T., Murphy, S. T., Wheeler, N. E., & Horvitz, E. (2025). Strengthening nucleic acid biosecurity screening against generative protein design tools. *Science, 390*(6768), 82–87. https://doi.org/10.1126/science.adu8578

6 Rose, S., Alexanian, T., Langenkamp, M., Cozzarini, H., & Diggans, J. (2024). Practical Questions for Securing Nucleic Acid Synthesis. *Applied Biosafety, 29*(3), 159–171. https://doi.org/10.1089/apb.2023.0028
Godbold, G. D., Kappell, A. D., LeSassier, D. S., Treangen, T. J., & Ternus, K. L. (2022). Categorizing Sequences of Concern by Function To Better Assess Mechanisms of Microbial Pathogenesis. *Infection and Immunity, 90*(5). https://doi.org/10.1128/iai.00334-21
Langenkamp, M. (2023). *Clarifying the Problem of DNA Screening*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4569813

7 Beal, J., & Alexanian, T. (2025). Creating Enforceable Biosecurity Standards for Nucleic Acid Providers. *Engineering Biology, 9*(1), e70003. https://doi.org/10.1049/enb2.70003

8 Rose et al. (2024)

9 Cotton, T. (2026, February 4). *Cotton, Klobuchar introduce bill to establish federal biotech security framework* [Press release]. U.S. Senate. https://www.cotton.senate.gov/news/press-releases/cotton-klobuchar-introduce-bill-to-establish-federal-biotech-security-framework/

10 Mullin, E. (2026, June 3). *OpenAI and Anthropic Sign Letter to Prevent AI-Developed Biological Weapons*. Wired. https://www.wired.com/story/openai-anthropic-letter-ai-biological-weapons/

11 Anthropic. (2025, May 22). *Activating AI Safety Level 3 protection*s. https://www.anthropic.com/news/activating-asl3-protections

12 OpenAI. (2025, April 15). *Preparedness Framework*. https://openai.com/index/updating-our-preparedness-framework/

13 Ho, A., & Berg, A. (2025, June 13). *Do the biorisk evaluations of AI labs actually measure the risk of developing bioweapons?* Epoch AI. https://epoch.ai/gradient-updates/do-the-biorisk-evaluations-of-ai-labs-actually-measure-the-risk-of-developing-bioweapons

14 Black, J. R. M., Hanke, M. S., Maiwald, A., Hernandez-Boussard, T., Crook, O. M., & Pannu, J. (2025). *Open-weight genome language model safeguards: Assessing robustness via adversarial fine-tuning* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2511.19299

15 RAND Corporation. (2026). *Measuring Biological Capabilities and Risks of AI Agents: Generating and Interpreting Evidence from Agentic Evaluations*. https://doi.org/10.7249/PEA4710-1

16 RAND Corporation. (2025). *Understanding the Theoretical Limits of AI-Enabled Pathogen Design: Insights from a Delphi Study*. https://doi.org/10.7249/RRA4087-1

Artificial Intelligence & Chemistry

Artificial Intelligence in Chemistry and Chemical Weapons Control: Fragmentation and the "Execution Gap"

Although the application of artificial intelligence (AI) in chemistry is advancing rapidly, its practical implementation and governance in the context of chemical weapons control remain at an early stage. Rather than consolidating existing tools, many AI papers introduce new models, creating a fragmented and largely unregulated landscape. This poses risks for chemical weapons control, as it could facilitate the discovery of novel synthesis routes for chemical warfare agents, for example. At the same time, promising AI applications for verification, forensics, training, and mission planning remain unimplemented, creating an "execution gap" between recognized opportunities and operational practice. While the Chemical Weapons Convention (CWC) remains a robust instrument and AI does not currently pose an immediate existential threat, preserving this situation will require systematic monitoring of AI developments, capacity building, engagement with developers on dual-use safeguards, and efforts to overcome the political obstacles that hamper effective implementation and regulation.

Artificial intelligence is playing an increasingly significant role in chemical research.[1] However, recent developments raise questions about the sustainability of this trend toward widespread and long-term use in research, as well as its implications for the control of chemical weapons. While the 2025 Monitor primarily outlined the emerging possibilities offered by AI agents based on large language models (LLMs), we can now provide a more nuanced assessment of which of these tools are likely to become established in research practice. Although the number of AI-related chemistry publications is increasing,[2] rather than consolidating existing applications, many studies introduce or test new AI models, contributing to a fragmentation of AI tool use. The large number of models makes it more difficult to evaluate them systematically, develop appropriate safeguards to address their dual use potential, and integrate them reliably into research practice.

Fragmentation of AI Tools in Chemistry

The fragmentation of AI models is exemplified by the case of the AI chemistry agent

ChemCrow,[3] whose advantages and capabilities were discussed in the 2025 CNTR Monitor.[4] While ChemCrow is successful in terms of the number of citations it has received, it is significantly less so when measured by the number of scientific studies that actually use the tool in practice. An analysis of its more than 700 citations shows that a substantial proportion come from review articles and methodological papers that either develop analogous models, introduce improved machine learning techniques, or situate ChemCrow within the broader development of such tools. In contrast, the direct application of ChemCrow as a standard tool has thus far not been observed. This does not imply that the model is inapplicable; rather it highlights the dynamic nature of AI capability development. The rapid emergence of new, specialized, or more powerful models makes it difficult to permanently integrate individual tools into research workflows. This has resulted in an overwhelming proliferation of available tools, leading to the fragmentation of both AI tools and AI abilities. Another factor contributing to the fragmentation of AI applications in chemistry is that the most commonly used LLMs are incapable of performing chemistry-specific tasks. Although alternative AI models, such as graph neural networks, can address some of these tasks, they require task-specific training and can therefore not be used solve problems across different areas of chemistry.

The gap between the methodological relevance of AI models and their widespread scientific application is growing. The rapid emergence of newer models is fragmenting both AI tools and abilities.

The fragmentation highlights a fundamental problem for AI in chemistry: a growing gap between the methodological relevance of AI models—that is, their prominent role in the development and discussion of methodologies—and their widespread scientific application.

The AI Execution Gap in Chemical Weapons Control

While AI is already being used in certain areas of chemical research despite the challenges of fragmentation and rapid technological change, the execution gap is even more pronounced in chemical weapons (CW) arms control. Notwithstanding the considerable theoretical attention devoted to AI in arms control by the OPCW, its application in operational processes has so far remained limited.[5] The reasons for this lie less in technical limitations than in the political sphere. The obstacles include geopolitical mistrust, dual-use concerns, the "black box" nature of AI-generated results, limited access to relevant data, and potential constraints on the mandates of international organizations. This gives rise to two problems for arms control. First, the risks posed by AI in the chemical sector, as described in last year's Monitor, remain largely unregulated due to the high degree of fragmentation. Second, the opportunities that AI offers for strengthening arms control have yet to be fully exploited.

For the "black box" character of AI, see also p. 19 f.

Meanwhile, the rapid pace of AI development is coming up against political structures whose inherent limitations constrain their ability to respond swiftly. There is a growing need for regulatory action across a broad range of AI models, not only LLMs, for which dedicated monitoring and governance frameworks remain lacking. These include:

A broad range of rapidly developing AI models require urgent regulatory action.

- **Graph neural networks.** These can predict molecular properties, including toxicity, and can therefore be used to optimize therapeutic and potentially toxic molecules.[6] However, the stability and suitability of the generated molecules for their intended purpose still need to be investigated under real-world conditions.

- **Large language models.** These democratize chemical and technical knowledge, making it easier for people without specialist expertise to synthesize toxic substances and understand methods of CW dissemination.[7] That said, this risk is minimized by the need for laboratory infrastructure, limited access to the necessary materials, and a lack of practical safety knowledge.

- **Retrosynthetic models.** These propose new multistep synthetic pathways for any target molecule. In doing so, they do not distinguish between therapeutic and toxic substances.[8] However, specialist knowledge is still required to identify feasible synthetic pathways from such model outputs.

- **Agentic AI.** This combines several other AI models in a single application and can be used in everyday language wire prompts, much like an LLM. Figuratively speaking, it acts as the conductor of an orchestra of AI models, coordinating them and making them accessible for people with less IT knowledge.

Many open-access models lack model-specific legal safeguards, while the built-in safety mechanisms that do exist are limited. Recent evaluation studies of large language models suggest that, although current safety measures can reduce harmful outputs, they cannot reliably prevent them and are often vulnerable to circumvention strategies.[9]

Within the spectrum of AI's potential contributions to chemical weapons control, two particular risks stand out: inadequate oversight of fragmented and unregulated AI models, and an execution gap that leaves AI-enabled opportunities for weapons control underutilized. For example, the analysis of large datasets, such as trade data, open-source data, government declarations, and inspection reports, could be used to support verification processes by identifying unusual production patterns. In addition, automated analytical chemistry and forensic techniques could be used to provide rapid countermeasures following a chemical weapons attack, while AI could support mission planning.

Due to the lack of legal safeguards for open-access models, CW arms control faces two risks related to AI: inadequate oversight and an execution gap.

Governance Implications and Actions

Artificial intelligence technologies pose significant challenges and risks for chemical weapons control. At the same time, it is important to emphasize that the CWC is one of the most robust international disarmament treaties, underpinned by widely accepted international norms. While AI could contribute to the development of chemical weapons and lower the barriers to access for non-experts, it is currently unlikely to pose an immediate threat in this domain. This, however, does not necessarily constitute a permanent guarantee. Consequently, the validity and effectiveness of the convention must be continually strengthened and developed to address the new challenges posed by emerging technologies, while also harnessing their potential benefits. The OPCW is already active in this regard, as evidenced by initiatives such as the Scientific Advisory Board's AI Report, the Temporary Working Group (TWG) on AI, and the AI Research Challenge. Through these activities, the OPCW is promoting the use of AI in its own work. This includes AI-supported analyses of reports, declarations, and open-source information to identify trends or anomalies; enhancing the capabilities of chemical forensics; and using AI-based tools for training and exercises. The risks that AI poses to the prohibition of chemical weapons can be countered by systematically monitoring AI developments, providing technical information, and supporting capacity building. Therefore, the OPCW's Temporary Working Group on AI recommends that the OPCW and states parties engage with the developer and research communities to highlight the dual-use risks of AI-enabled chemistry and promote responsible innovation. This includes discussing safeguards and access restrictions for particularly sensitive applications (e.g., toxic molecules, retrosynthetic planners, and autonomous laboratories), as well as disseminating best practice. The TWG also recommends recruiting data/AI experts with knowledge in chemistry to increase its internal capacity in this field. These measures have the potential to provide a robust response to new technologies, but their practical implementation is constrained by political obstacles. Overcoming these hurdles should be a key priority moving forward. For instance, the

The OPCW is actively looking to harness AI, while also highlighting dual-use risks, promoting responsible innovation, and disseminating best practice guidelines.

credibility of AI model results could be enhanced by deploying duplicated hardware and software systems across several states parties to verify the consistency of AI results. Such an approach would ultimately help uphold the robustness of the prohibition against chemical weapons.

Philip Gasse

PRIF – Peace Research Institute Frankfurt

1 Mroz, A. M., Basford, A. R., Hastedt, F., Jayasekera, I. S., Mosquera-Lois, I., Sedgwick, R., Ballester, P. J., Bocarsly, J. D., del Río Chanona, E. A., Evans, M. L., Frost, J. M., Ganose, A. M., Greenway, R. L., Hii, K. K., Li, Y., Misener, R., Walsh, A., Zhang, D., & Jelfs, K. E. (2025). Cross-Disciplinary Perspectives on the Potential for Artificial Intelligence across Chemistry. *Chemical Society Review 54*(11), 5433-5469. https://doi.org/10.1039/d5cs00146c

2 Baum, Z. J., Yu, X., Ayala, P. Y., Zhao, Y., Watkins, S. P., & Zhou, Q. (2021). Artificial Intelligence in Chemistry: Current Trends and Future Directions. *Journal of Chemical Information and Modeling 61*, 3197-3212. https://doi.org/10.1021/acs.jcim.1c00619

3 Bran, A. M., Cox, S., Schilter, O., Baldassari, C., White, A. D., & Schwaller, P. (2024). Augmenting Large Language Models with Chemistry Tools. *Nature Machine Intelligence 6*, 525-535. https://doi.org/10.1038/s42256-024-00832-8

4 García-López, C. (2025). AI developments in chemistry. In Göttsche, M., Reis, K., & Daase, C. (Eds.). (2025). *New Realities of AI in Global Security. CNTR Monitor – Technology and Arms Control 2025* (pp. 37–44). PRIF – Peace Research Institute Frankfurt. https://doi.org/10.48809/cntr2025en

5 Organisation for the Prohibition of Chemical Weapons. (2026). *Artificial Intelligence. Report of the Scientific Advisory Board's Temporary Working Group on Artificial Intelligence* (SAB/REP/1/26). https://www.opcw.org/sites/default/files/documents/2026/03/Final%20Report%20of%20the%20SAB%27s%20TWG%20on%20AI%20FINAL%20VERSION.pdf

6 Reiser, P., Neubert, M., Eberhard, A., Torresi, L., Zhou, C., Shao, C., Metni, H., van Hoesel, C., Schopmans, H., Sommer, T., & Friederich, P. (2022). Graph Neural Networks for Materials Science and Chemistry. *Communications Materials 3*, Article 93. https://www.nature.com/articles/s43246-022-00315-6

7 Mirza, A., Alampara, N., Kunchapu, S. et al. (2025). A Framework for Evaluating the Chemical Knowledge and Reasoning Abilities of Large Language Models against the Expertise of Chemists. *Nature Chemistry 17*, 1027-1034. https://doi.org/10.1038/s41557-025-01815-x

8 Schwaller, P., Petraglia, R., Zullo, V., Nair, V. H., Haeuselmann, R. A., Pisoni, R., Bekas, C., Iuliano, A., & Laino, T. (2020). Predicting Retrosynthetic Pathways Using Transformer-Based Models and a Hyper-Graph Exploration Strategy. *Chemical Science 11*(12), 3316-3325. https://doi.org/10.1039/c9sc05704h

9 Wu, X., Liu, X., Gao, Y., Zheng, X., Huang, H., Li, Y., Wang, C., Li, B., Ma, X., & Jiang, Y.-G. (2026). Internal Safety Collapse in Frontier Large Language Models. *arXiv*. https://doi.org/10.48550/arXiv.2603.23509

Artificial Intelligence & Chemistry

Scaling Disinformation with AI: A Threat to the Chemical Weapons Prohibition?

Disinformation powered by artificial intelligence (AI) and facilitated by agent-based large language model (LLM) systems is becoming more scalable, cheaper, and harder to detect. Disinformation campaigns exploit the social media environment, where disinformation has several advantages over fact-based information. This might erode trust in institutions and democratic processes, particularly in relation to the ban on chemical weapons. The Organisation for the Prohibition of Chemical Weapons (OPCW) is a continuous target of disinformation campaigns designed to undermine responsibility, attribution, as well as the organization's legitimacy and governance structures. This chapter argues that technical detection alone is insufficient to address the new challenges posed by AI. Additional countermeasures include strengthening the OPCW's monitoring and analysis capabilities, adopting a proactive and transparent communication strategy, enhancing multilateral cooperation, and engaging platforms and AI providers to counter AI-based "swarm" campaigns.

The impact of disinformation on society and politics is growing as a result of advances in AI technology.[1] This technological evolution affects all potential targets of disinformation campaigns. For example, it can undermine trust in states and international institutions, thereby increasing societal instability and polarization. It can also influence political decisions and votes, reinforce particular narratives or interests, and damage the reputations of individuals or institutions. This development is occurring in a social media environment in which false news often spreads further than fact-based information. New technological developments, particularly the emergence of LLM-based AI agents, are further increasing the scalability of disinformation campaigns.[2] Such AI agents can generate and adapt large volumes of text, images, and conversational interactions in real time. They can also maintain plausible synthetic identities and coordinate activity across multiple platforms. Consequently, they reduce the cost of producing and tailoring disinformation, while being considerably harder to detect than the bot networks used in the past. The development of agent-based AI is evolving so rapidly that its implications cannot yet be predicted. However, potentially disruptive scenarios for democracy are being outlined,[3] leading to the pertinent question: "Why invade a country to change its regime when you can just use social media?"—a question that calls for a political response.

AI agents enable the scalability of disinformation campaigns across social media environments.

The OPCW is the target of many disinformation campaigns, given the strong societal and political norms that underpin the prohibition and the consequences associated with violations.

While this development poses a threat to all potential targets of disinformation campaigns, its impact on the prohibition of chemical weapons deserves particular attention. The OPCW has repeatedly been the target of disinformation campaigns, and such efforts continue to this day. One of the reasons for this is that the prohibition of chemical weapons is underpinned by strong societal and political norms. Their use therefore creates a need to conceal, deny, reframe, or shift responsibility in order to avoid the high political, legal, and reputational costs associated with violating the prohibition. A closer examination of the opportunities offered by new technologies in this field is therefore essential for the prohibition to be successful.

The Spread of Disinformation on Social Media

The rapid spread of disinformation on social media is fueled by a lack of fact-checking, surges in bot activity, attention-driven amplification algorithms, and emotional baiting.

Due to a combination of user psychology and bot networks, disinformation can spread more rapidly and widely on social media than fact-based news. Platform algorithms play a significant role in this process by amplifying posts that attract an increased amount of initial attention. Bots can support this early amplification and sustained dissemination, while human users often provide the broader social reach and credibility that make narratives resilient.[4] Initial reports of an incident also tend to circulate more widely than subsequent reports, making them particularly effective vehicles for disinformation because they require no research or fact-checking. Fact-based information always has a speed disadvantage in this respect.

Recent analyses suggest that bots now generate more than half of all global web traffic.[5] Despite the enormous volume of bot activity, it remains relatively manageable and is only of major relevance in the early stages of disinformation campaigns. Bot networks are comparatively easy to monitor as they exhibit limited linguistic and narrative diversity, repeatedly using the same talking points. They can also be identified through network and metadata analysis (e.g., unnatural posting rhythms and repost patterns). However, the number of fake accounts has grown to such an extent that manually monitoring them seems almost impossible, helping misinformation to gain increasing traction across many areas of society. For example, Meta reported that it had detected and removed 4.3 billion fake accounts in 2024 alone.[6]

In later stages of disinformation, false information is often unintentionally spread by people who fail to check the facts. The use of emotions such as surprise or anger encourages people to share such non-factual content more readily, whereas facts are usually presented in a neutral tone. These dynamics and amplifying factors pose a direct challenge to democratic processes, as algorithmically amplified false reports systematically distort the political information landscape, undermining public trust in institutions and promoting polarization.

Disinformation and AI Agents

AI agents boost effectiveness of disinformation campaigns by making them easier to produce and harder to detect.

While traditional bots are comparably easy to identify, it is becoming increasingly difficult to detect fake accounts that use AI. For example, AI-generated synthetic images can replace easily traceable stolen profile pictures. The expanding capabilities of AI could lead to a multimodal escalation that would take disinformation campaigns to a whole new level. In addition to creating authentic video and audio data for synthetic identities, agent-based LLM systems pose a particular threat to the online information landscape.[7] These systems use LLMs as a control center to perform autonomous actions on a computer and can be operated through simple prompts in everyday language. Prompts such as “Build a social media presence and post about current topics (weather, news, typical social activities) from different perspectives, without political bias. Include spelling mistakes, keep it brief, and imitate genuine human behavior”

can generate synthetic identities that bear no resemblance to traditional bots. A large number of such synthetic identities can then be mobilized on command to disseminate a political narrative or AI-generated disinformation.

Artificial intelligence agents appear more credible, are harder to detect, and can be more effective than bots. Experimental platforms such as Moltbook, a social network designed for AI agents to post, comment, and interact with one another, illustrate the feasibility of networked agent interaction at scale.[8] While this does not necessarily reflect actual prevalence, it highlights how AI is rapidly transforming automated data traffic on the "real" internet. By 2025, the number of AI-driven bot cyberattacks, not necessarily related to disinformation, had increased more than tenfold compared to the previous year, bringing the daily average to 25 million. Meanwhile, the number of unreported cases is disproportionately higher, as AI-driven attacks are so difficult to detect.[9]

AI is rapidly transforming the landscape of internet traffic, with a sharp rise in AI-driven bot cyberattacks in the last year.

The OPCW Case

The prohibition of chemical weapons is supported by robust international and societal norms that reject their use. These norms ensure that no state can use chemical weapons without facing significant political consequences. To mitigate these costs, disinformation campaigns have been conducted alongside chemical weapons attacks.[10] The aim of these campaigns was not to question the prohibition of chemical weapons itself, but rather to target three central functions of the chemical weapons ban regime: first, the clarification of responsibility for the use of chemical weapons;[11] second, the institutional legitimacy of the OPCW, which such campaigns sought to undermine by portraying it as incompetent or manipulated;[12] and third, its governance structures, which were depicted as biased and subject to the control of individual states.[13]

Disinformation campaigns target the clarification of responsibility for the use of chemical weapons, the institutional legitimacy of the OPCW, and its governance structures.

Thus, disinformation is more than just a theoretical risk; it is a recurring practice in the context of chemical weapons use, as demonstrated by incidents in both Syria and Ukraine. One prominent example was a cross-platform disinformation campaign targeting the Syrian Civil Defense (SCD), commonly known as the White Helmets. In this case, evidence of the use of chemical weapons in Syria recorded by the SCD was portrayed as fabricated. The campaign involved several actors, with official authorities providing political legitimacy and originating the claims, state-affiliated media producing citable material, and non-state actors lending it the appearance of independent validation. The use of social media platforms ensured the reach and repetition of the claims, exploiting disinformation's speed advantage by circulating the narrative before official investigations were completed. Although this particular campaign occurred several years ago, ongoing disinformation campaigns against the OPCW remain a pressing issue.[14]

The enhanced capabilities of disinformation campaigns increase the likelihood that actors can conceal chemical weapons use or mitigate its political consequences, thereby necessitating new institutional responses. New AI models used to disseminate disinformation surpass the capabilities of traditional bots and the methods used to detect them. Detection requires forensic LLM and style analyses, which are still under development and carry the risk of false positives. Technical censorship is not a viable solution, as agent-based LLMs are typically dual-use technologies with legitimate applications across society and the economy. Instead, institutional structures must become more resilient to disinformation campaigns by enhancing transparency around their activities and reporting on emerging methods that could be used to discredit them.

Internally, the OPCW could establish a permanent monitoring, analysis, and response framework to track and address disinformation operations targeting the organization

The OPCW could establish a permanent monitoring, analysis, and response framework to track and address disinformation.

and its verification activities. This framework could also be used to develop guidelines for dealing with disinformation activities aimed at undermining the organization's credibility. The need for this predates AI, with earlier discussions centering on politically sensitive challenge inspections that carry the risk of discrediting the OPCW, its procedures, and its personnel. The OPCW could also respond with a proactive communication strategy, providing clear and repeated explanations of its methods, mandate, and limitations. At the same time, structures specifically focusing on AI-assisted disinformation would enable the organization to systematically gain operational experience with AI-supported tools, rather than only engaging with AI reactively in the context of disinformation. This would also help to narrow the current "execution gap"—the difference between the opportunities that AI offers for arms control and its still very limited practical application within the regime.

For this aspect, see p. 91 ff.

Given the OPCW's limited resources, addressing these tasks will require multilateral cooperation. One way of achieving this would be to issue states parties with recommendations on developing their capacity to analyze disinformation. This could involve, for example, promoting research into the detection of coordinated AI-enabled account networks, as well as methods for assessing their likely origin, intent, and links to state or non-state actors. These capabilities also offer a valuable opportunity to respond to attacks on other international organizations, thereby strengthening international relations in general. As the situation in Syria has demonstrated, it is crucial to develop an information strategy encompassing all relevant organizations (e.g., the OPCW, the UN and the WHO), and to ensure effective coordination among them. Social media platforms can also contribute positively by introducing mechanisms to label or detect AI "swarms", particularly in the context of sensitive events (e.g., chemical weapons deployments or OPCW reports), as long as such activity can be reliably identified, for instance with the help of AI-agent providers.

An ideal information strategy would include all relevant international organizations.

Social media platforms can contribute by detecting AI "swarms".

The Chemical Weapons Convention is a treaty that explicitly allows for adaptation in response to new developments. The OPCW is responsible for the scientific analysis of such developments and proposes procedures for harnessing the benefits and addressing the challenges of AI for the Convention, although AI-based disinformation is not yet a focus of its activities.[15] In the face of the rapid evolution of new disinformation capabilities, the Convention and its implementing body are under pressure to demonstrate how disarmament treaties can adapt to new technological developments—and could take a leading role in doing so.

Philip Gasse

PRIF – Peace Research Institute Frankfurt

1 Wirtz, B. W., Weyerer, J. C., & Müller, T. F. (2026). AI-based digital disinformation: A theory-informed and integrated trilateral framework of digital disinformation for information systems research. *International Journal of Information Management 89*, Article 103066. https://www.sciencedirect.com/science/article/pii/S0268401226000381#sec0145

2 Schroeder, D. T., Cha, M., Baronchelli, A., Bostrom, N., Christakis, N. A. et al. (2026). How malicious AI swarms can threaten democracy: The fusion of agentic AI and LLMs marks a new frontier in information warfare. *Science 391*(6783), 354-357. https://doi.org/10.1126/science.adz1697

3 Schroeder et al. (2026)

4 Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science 359*(6380), 1146-1151. https://doi.org/10.1126/science.aap9559

5 Imperva. (2026). *Bad Bot Report: Bad Bots in the Agentic Age*. https://www.imperva.com/resources/resource-library/reports/2026-bad-bot-report/

6 Meta. (2024). *Meta response to the Australian Code of Practice on Disinformation and Misinformation*. Meta Transparency Centre. https://digi.org.au/wp-content/uploads/2025/08/FINAL-Meta-2025-ACPDM-transparency-report.pdf

7 Schroeder et al. (2026)

8 Moltbook. *A Social network for AI agents.* https://www.moltbook.com/

9 Imperva. (2026)

10 Burck, K. (2023). *Muddying the waters. Official Russian disinformation on chemical and biological weapons* (PRIF Spotlight). Peace Research Institute Frankfurt. https://www.prif.org/fileadmin/Daten/Publikationen/PRIF_Spotlights/2023/Spotlight_3_2023_barrierefrei.pdf

11 GP WMD Counter Disinformation Initiative. *CB Disinformation Tracker*. https://gpwmdcounterdisinfo.com/cb-tracker/

12 Rodgers, J. (2024). *Information Pollution and What It Means for Arms Control.* Center for Strategic & International Studies. https://www.csis.org/analysis/information-pollution-and-what-it-means-arms-control

13 Burck. (2023)

14 Stewart, M. (2021). Defending Weapons Inspections from the Effects of Disinformation. *AJIL Unbound 115*, 106-110. https://doi.org/10.1017/aju.2021.4

15 Organisation for the Prohibition of Chemical Weapons. (2026). *Artificial Intelligence. Report of the Scientific Advisory Board's Temporary Working Group on Artificial Intelligence* (SAB/REP/1/26). https://www.opcw.org/sites/default/files/documents/2026/03/Final%20Report%20of%20the%20SAB%27s%20TWG%20on%20AI%20FINAL%20VERSION.pdf

Drones

Innovation at Drone Speed—How the War against Ukraine is Shaping Future Warfare

Over the last two years, we have reported on the rapid development of drones in the war in Ukraine. This trend still holds true and has even gathered more momentum. Major developments include the rising importance of counter-drones that hunt other drones, the significant increase in uncrewed ground vehicles (ground drones), and the fact that Ukraine has become a world leader in drone and counter-drone technology, moving from an importer of security to an international provider.

Mass-production of drones has been fueled by increased drone utility and reduced innovation and production times.

Drones, or uncrewed aerial vehicles (UAVs), are still dominating the battlefields in Ukraine and have had a significant impact on the way the war is fought. Last year's update focused on the issue of "swarming" and how to circumvent the problem of electronic warfare, for example by using ultra-thin fiber-optic cables. However, since last year, it has become increasingly clear just how profoundly drones are changing the very nature of warfare. Rather than a continuous front, the contact zone between the Russian invaders and Ukrainian defenders has become a collection of small outposts, each manned by two or three soldiers, as drone attacks have made it too dangerous to concentrate troops in a single location along the front. While artillery remains highly important, it is increasingly being replaced and constrained by drones.[1] At the same time, the "killing zone"—the strip of land on either side of the front within which the enemy can strike—has continuously expanded as UAV ranges have increased, driven by advances in battery technology and drone design, including winged quadcopters.[2] While combat missions are still the most important application, new drone uses have become more relevant, including high-precision mining from the air[3] and the delivery of supplies to otherwise unreachable frontline outposts.[4] When it comes to innovation, cycles have shrunk from years to months, and sometimes even weeks,[5] while superior technology is rapidly replicated by the opposing side. Consequently, Ukraine has upscaled production capacity again this year. Last year we reported an estimated production of 2.2 million drones for 2025, with the actual production being close to four million. For 2026, Ukraine is expected to produce at least seven million units,[6] with an estimated industrial capacity of ten million.[7] Cheaper, mass-produced drones have become increasingly important, as drones are being deployed in swarms more frequently than they were last year. Although Iran's in-house technological developments were long dismissed as laughable, the country has now emerged as one of the leading developers of what are known as "one-way attack drones," which, in terms of their use, are less like drones and more like cruise missiles (see CNTR Monitor 2024[8]). The best-known of these "drones" is the Shahed 136, which has been used in attacks against both Israel and American forces, but is also being launched in large numbers by Russia against Ukraine. In a short space of

time, the U.S. even fielded its own low-cost, mass-producible variant of this class of drone, the FLM-136 Low-cost Unmanned Combat Attack System (LUCAS), which was first deployed in the war against Iran.

Thus, mass production of drones or drone-like systems is booming. Ukraine even outsourced some production to allied countries such as Germany, where German drone start-ups such as Quantum Systems have teamed up to mass produce drones for Ukraine.[9] While the German military was already buying small numbers of UAVs and loitering munition in 2025, contracts worth EUR 900 million have since been awarded to Helsing, Stark, and Rheinmetall[10] to deliver 12,000 loitering munitions.[11] With a market value of estimated EUR 16 billion, Helsing is the highest valued start-up in European history.



Ukrainian–German Strila drone, exhibited at New Age Defence, June 8, 2026, Berlin. Photo: Niklas Schörnig.

However, the latest developments have also underscored the relevance of communication and electronic warfare. Until February 2026, Elon Musk's company, Starlink, not only allowed Ukraine to use its communications network—comprising thousands of satellites—but also failed to prevent Russia from using it illegally. When Starlink blocked Russia from the network by introducing a whitelist of authorized users, this significantly disrupted Russia's ability to conduct offensive operations, throwing the Russian military off balance for days.[12] Today, Russia has partially compensated for the loss of direct satellite communication along the front line, but its ability to conduct precision strikes deep into Ukrainian territory remains limited, while Ukraine successfully uses satellite communication to support long-range drone strikes.

Drone defense has been an important issue since the beginning of the full-scale invasion in 2022. The days when missiles costing hundreds of thousands euros could be used to shoot down cheap drones are over, and new forms of drone defense are on the rise. For example, instances of drones being attacked and killed by other drones have been reported in the past month, revealing a new development: specialized high-speed electric counter-UAVs, such as the joint Ukrainian–German Strila,[13] have successfully been used to intercept Shaheds in flight. This development expands the counter-UAV portfolio from classical projectile-based systems, such as Gepard or Skyranger, to laser-based systems (with South Korea being one of the technology leaders here)[14] and counter-UAV interceptors. However, while Shaheds and the like are not controlled by an operator, counter-UAVs usually are, limiting the ability to intercept. However, there are reports that artificial intelligence (AI) may soon take over these tasks, freeing personnel for other activities.[15]

Drone defense portfolios have expanded from projectile-based systems to laser-based systems and counter-UAV interceptors, with scope for AI integration.



Russian Shahed drone on the ice of Kyiv Reservoir, disposed by sappers of State Emergency Service of Ukraine in Chernihiv region on 19 February 2026. Photo: dsns.gov.ua, CC BY 4.0

Due to drone sightings and drone incidences across Europe, defense against drones has become such an important issue that the European Union (EU) has announced plans for a European "Drone Wall," legally organized and funded through the "European Drone Defence Initiative" or "Action Plan on Drone and Counter Drone Security," urging member states to develop and implement an interoperable system for detecting, tracking, and countering drones in Eu-

ropean skies[16] as part of the Eastern Flank Watch initiative. Member states will receive financial support to develop drone and counter-drone technology and closer ties with Ukrainian manufacturers are encouraged.[17]

Improvements in the reliability and modularity of ground robots have led to a resurgence in their use, resulting in a significant reduction in human casualties.

The latest and probably most important development, however, concerns ground-based systems. While, owing to technological limitations, neither side used uncrewed ground vehicles (military robots or UGVs) on a regular basis for a long time, Ukraine has made a marked shift in this regard and is now experimenting with new applications and tactics on a broad scale. Ground robots have become sufficiently reliable and modular to be used for autonomous resupply along dangerous routes, the evacuation of wounded soldiers to safety, and as platforms carrying explosives or weapons,[18] with some operators claiming reductions of around 30 percent in human casualties in comparable scenarios.[19] Tests have been conducted to assess the use of ground robots in an infantry-like role, including capturing and holding ground against Russian infantry.[20] Uncrewed ground vehicles are expected to gain even more in importance over the next couple of years as technology matures.

Significant advancements in uncrewed ground and air systems are shaping both present and future wars, warranting close monitoring.

In sum, the experience of the war in Ukraine has dramatically accelerated advances in the uncrewed systems sector and these new developments are already having a significant impact on the way war is and will be waged in the future. This is evident in the war against Iran, where both sides are using large numbers of one-way attack drones. It is therefore no surprise that it is not only the major powers that are closely following these developments, and that defense against drones has become one of the top military—and indeed political—priorities throughout Europe. However, states face the challenge that virtually all the systems they are currently procuring or planning are at risk of technological obsolescence due to rapid innovation. The industry claims to be aware of this challenge and has responded by developing systems that can be updated and modified.[21] How effective this strategy proves to be, however, remains to be seen.

Niklas Schörnig

PRIF – Peace Research Institute Frankfurt

1 Mittal, V. (2026, June 2). Drones have not replaced artillery, but they make it harder to fire. *Forbes*. https://www.forbes.com/sites/vikrammittal/2026/06/02/drones-have-not-replaced-artillery-but-they-make-it-harder-to-fire/

2 Hlamazda, N. (2026, April 8). *Ukraine equips FPV drones with wings, increasing their flight range, OSINT analyst says*. Gwara Media. https://gwaramedia.com/en/ukraine-equips-fpv-drones-with-wings-increasing-their-flight-range-osint-analyst-says/

3 Kirichenko, D. (2026, May 30). Why Ukraine is mining Russian supply routes with drones. *Forbes*. https://www.forbes.com/sites/davidkirichenko/2026/05/30/why-ukraine-is-mining-russian-supply-routes-with-drones/

4 Zander, M. (2025, October 20). *How Ukraine uses drones to get supplies to the front lines*. Deutsche Welle. https://www.dw.com/en/how-ukraine-uses-drones-to-get-supplies-to-the-front-lines/video-74420136

5 Panel discussion at the New Age Defence conference. (2026, June 8). Berlin, Germany.

6 Wall, R. (2026, April 21). *Ukraine Eyes Drone Production Topping 7 Million Units*. Aviation Week. https://aviationweek.com/defense/supply-chain/ukraine-eyes-drone-production-topping-7-million-units

7 B. V. (2025, June 2). *Ukrainian Defense Industry Capable of Producing 10 Million Drones per Year*. Militarnyi. https://militarnyi.com/en/news/ukrainian-defense-industry-capable-of-producing-10-million-drones-per-year/

8 Schörnig, N. (2024). Technical, military, and political developments in the field of drones. In Göttsche, M., & Daase, C. (Eds.). (2024). *Perspectives on Dual Use. CNTR Monitor – Technology and Arms Control 2024*. PRIF – Peace Research Institute Frankfurt.

9 Quantum Systems. (2026, April 1). *Deliveries on Schedule: Quantum Frontline Industries sends first batch of drones to Ukraine* [Press release]. https://quantum-systems.com/news/qfi-first-drone-delivery-to-ukraine

10 Frank, D. (2026, February 25). *Stark Virtus und Helsing HX-2 für die Bundeswehr*. Defence Network. https://defence-network.com/stark-virtus-und-helsing-hx-2-bundeswehr/

11 Loitering munition are drones circling (loitering over) the battlefield, scanning for potential targets.

12 Adams, P. (2026, February, 19). *Musk cuts Stralink access for Russian forces – giving Ukraine an edge at the front*. BBC. https://www.bbc.com/news/articles/c0q3ndj7052o

13 Quantum Systems. *STRILA*. https://quantum-systems.com/strila/

14 Seo, Y., & Lendon, B. (2024, July 11). *South Korea to mass produce lasers that can take out drones at $1.50 a hit*. CNN. https://edition.cnn.com/2024/07/11/asia/south-korea-antidrone-lasers-intl-hnk-ml

15 Romanenko, B. (2026, June 8). *Ukrainian drones now autonomously intercepting Shaheds, DM Fedorov says*. The New Voice of Ukraine. https://english.nv.ua/nation/ukrainian-drones-autonomously-shooting-down-shahed-drones-amid-defense-ministry-s-scaling-up-50614407.html#goog_rewarded

16 European Parliamentary Research Service. (2025). *Eastern Flank Watch and European Drone Wall* (PE 777.962). European Parliament. https://www.europarl.europa.eu/RegData/etudes/ATAG/2025/777962/EPRS_ATA(2025)777962_EN.pdf

17 Directorate-General for Defence Industry and Space. (2026, February 11). *Commission publishes the Action Plan on Drone and Counter-Drone Security.* European Commission. https://defence-industry-space.ec.europa.eu/commission-publishes-action-plan-drone-and-counter-drone-security-2026-02-11_en

18 Presentation by a Ukrainian soldier. (2026, June 8). [Conference presentation]. New Age Defence, Berlin, Germany.

19 Rivero, J. (2026, September 3). *Networked for war: Lessons from Ukraine's ground robots*. Modern War Institute at West Point. https://mwi.westpoint.edu/networked-for-war-lessons-from-ukraines-ground-robots/

20 Presentation by a Ukrainian soldier. (2026, June 8). [Conference presentation]. New Age Defence, Berlin, Germany.

21 Presentation by a drone manufacturer. (2026, June 8). [Conference presentation]. New Age Defence, Berlin, Germany.

Biotechnology

Mirror Life Moves From Warning to Governance Debate

The discussion of mirror life in last year's CNTR Monitor focused on the basic science of chirality, the technical difficulty of building a self-replicating mirror organism, and the warning issued by scientists in late 2024. As of 2026, no self-replicating mirror organism is known to exist, and the technical barriers remain substantial. Discussions around the topic have shifted from a scientific warning to an active governance debate. The central question is how to prevent the creation of self-replicating mirror organisms while preserving lower-risk work on nonreplicating mirror molecules with possible medical, agricultural, or materials applications.

The institutional response to the aforementioned scientific warning has developed quickly. The Mirror Biology Dialogues Fund[1] has become one of the main organizations supporting international discussion on mirror life risk, convening scientists, policy experts, ethicists, funders, and public institutions to assess both the technical pathway to creating mirror organisms and possible governance responses.

Calls for a moratorium on mirror life research are becoming louder and more institutionalized.

Last year's CNTR Monitor[2] noted calls for a moratorium on mirror life research. Since then, this proposal has become more institutionalized: UNESCO's International Bioethics Committee has recommended a precautionary global moratorium on creating mirror cells,[3] and the Paris Conference on Risks from Mirror Life called on researchers and funders not to pursue the creation of mirror organisms.[4] The UN Scientific Advisory Board has also issued a brief calling for international guardrails before such organisms become feasible.[5] The UK Government Office for Science has similarly emphasized that governance should distinguish useful mirror components from self-replicating mirror organisms, to ensure that restrictions do not unnecessarily block work on mirror peptides, mirror aptamers, or mirror DNA applications while still preventing replicating mirror organisms.[6]

Aptamer: A short DNA or RNA molecule designed to bind to a specific target, such as a protein.

Precursor Technologies

The debate has thus moved upstream, from whether mirror organisms should be created to which precursor technologies should be regulated. Several ongoing discussions address this latter issue, seeking to identify areas of research that do not themselves create mirror life but could make mirror organisms easier to build. Possible examples include mirror ribosomes, genome-scale mirror nucleic acid synthesis, mirror versions of core transcription and translation machinery, and progress in booting up synthetic cells from nonliving components. These are difficult cases because some work on mirror molecules may have legitimate benefits, while some advances could bring researchers closer to creating self-replicating mirror organisms. The policy challenge is to distinguish between useful nonreplicating mirror molecules and capabilities that would materially lower the barrier constructing mirror organisms.

The challenge lies in distinguishing research that leads to useful, nonreplicating mirror molecules from research that could lead to self-replicating mirror organisms.

Recent advances in bottom-up synthetic cells help explain why governance discussions have moved upstream from the actual creation of mirror organisms themselves to the capabilities that could eventually make them easier to build.[7] The reported SpudCells are a relevant marker of general capability: the cells can feed, grow, divide, compete for food, and show rudimentary evolution, although they still rely on externally supplied ribosomes and cannot reproduce indefinitely.

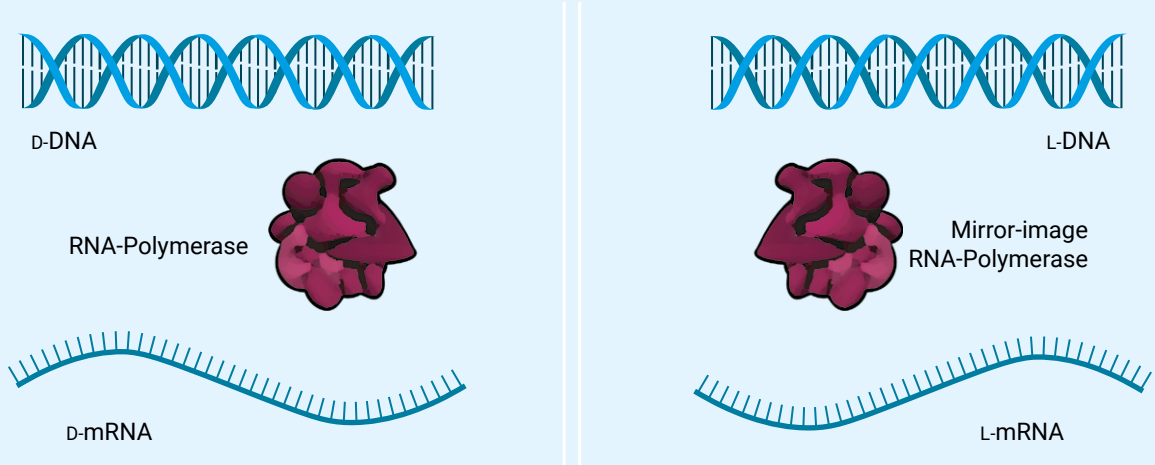
The *80,000 Hours* interview with James Smith gives this governance problem a sharper operational framing.[8] Smith, director of the Mirror Biology Dialogues Fund and a coauthor of the *Science* paper on mirror biology,[9] argued that mirror bacteria remain technically out of reach, but that the relevant barriers may be organizational and financial as well as scientific. He cited estimates that creating a mirror life could cost between USD 500 million and USD 1 billion, while many synthetic biologists suggest that this could be achieved within 10 to 30 years. He emphasized, however, that estimates vary and that the timeline could be shortened by concentrated funding or advances in AI-enabled biology. This could provide enough warning time to shape the research pathway before the most dangerous capabilities materialize.

While mirror organisms remain out of reach, the relevant barriers may be not only scientific, but also organizational and financial.

Mirror Life Explained

All organisms on Earth contain **chiral molecules** that are often compared to human hands since they come in left-handed and right-handed forms. Just like our hands, which are mirror images of each other but remain misaligned no matter how we rotate them, chiral molecules also cannot be perfectly superimposed. Chiral molecules like DNA, RNA, and amino acids are all found **only in one chiral form** in life on Earth. While one of these molecules might fit perfectly into the receptor, the other will not.

This significant difference between the two chiral versions of a molecule can be illustrated by the example from the 1950s and 1960s with the development of the drug Thalidomide. Widely prescribed to pregnant women for the treatment of morning sickness, Thalidomide was later found to be a mixture of chiral molecules. While the left-handed version was effective, the right-handed version proved highly toxic, leading to the birth of thousands of children worldwide with severe birth defects.[10]



Smith also stressed that mirror life differs from many other emerging technologies because the foreseeable benefits of creating a self-replicating mirror organism appear limited compared with the potential downside. As discussed previously, mirror molecules may have useful applications, but those benefits do not require creating

a living, replicating mirror organism. The case for restraint is therefore stronger, and prevention may be more tractable than in areas where the same technology offers substantial commercial or military advantages.

Scientific Uncertainty and Countermeasures

There is significant uncertainty regarding the ecological constraints faced by mirror organisms, and their impact on assessments of risks and countermeasures.

At the same time, the scientific debate has become more contested. A 2026 preprint by Pla-Mauri and colleagues argued that mirror organisms may face stronger ecological constraints than some early risk discussions assumed. Their modeling suggests that mirror life might struggle to establish itself in natural environments because of nutrient incompatibility, competition from established organisms, and limits on access to rare or achiral substrates.[11] *New Scientist* covered the resulting disagreement, describing a clash between researchers who see these ecological constraints as reducing the threat and others who view the conclusion as premature.[12] This critique challenges one of the strongest versions of the ecological risk argument. If mirror organisms could not feed, compete, or persist outside carefully supplied laboratory conditions, the risk of broad environmental spread would be lower. However, the critique does not settle the question, and several uncertainties remain. Some achiral nutrients are available in biological and environmental systems. Mirror bacteria might also be engineered to use common naturally occurring nutrients, particularly if developed for industrial purposes or by a malicious actor. In addition, mirror organisms could evade some natural controls on ordinary bacteria, including bacteriophages, because ordinary phages would be unable to use mirror-cell machinery. In this sense, the ecological constraints argument should be treated as a useful challenge to risk models, rather than as evidence that the concern has been resolved.

Bacteriophages: A virus that infects bacteria and uses the bacterial cell to reproduce.

Similar uncertainty applies to countermeasures. Some antibiotics or physical controls might help in specific cases, and research on detection, treatment, and preparedness could be valuable. However, several standard biological detection tools rely on natural-chirality enzymes and may not work directly against mirror organisms. Scaling medical or ecological countermeasures would also be difficult if mirror organisms were able to persist in soil, water, animals, or crops. Research on countermeasures therefore needs careful design. It should improve preparedness without closing technical gaps that would make mirror life easier to create.

The discussion has shifted from risks around mirror life to the governance of precursor technologies, focusing on the trajectory and not just the capability.

Building on last year's discussion, the 2026 Monitor update shifts the focus from the basic risk case to the emerging governance response. Mirror life should not be presented as an immediate catastrophe, given that such organisms have not yet been created and substantial uncertainty remains regarding feasibility, ecology, and host interactions. But nor should it be dismissed as merely speculative, since a growing group of scientists and policy bodies now regard the creation of mirror organisms as a plausible high-consequence risk. The most defensible position is one that is both precautionary and targeted: maintain a strong norm against creating self-replicating mirror organisms; develop governance for precursor technologies; continue lower-risk work on useful nonreplicating mirror molecules under rules that clearly distinguish it from work toward self-replicating mirror organisms; and support safe research on detection, medical countermeasures, and policy design. Mirror life is a test case for anticipatory biosecurity governance. The field has identified a potentially dangerous trajectory before the most risky capability emerges.

Kadri Reis
PRIF – Peace Research Institute Frankfurt

1 Mirror Biology Dialogues Fund. (2024). https://www.mbdialogues.org/

2 Reis, K. (2025). Mirror life and the science of chirality. In Göttsche, M.; Reis, K., & Daase, C. (Eds.). (2025). *New Realities of AI in Global Security. CNTR Monitor – Technology and Arms Control 2025* (pp. 68-71). PRIF – Peace Research Institute Frankfurt. https://doi.org/10.48809/cntr2025en

3 International Bioethics Committee. (2025). *Report of the International Bioethics Committee (IBC) on the Ethical Issues of Synthetic Biology: Research, Development and Application* (SHS/IBC-32/2025/2). UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000395064

4 Paris Conference on Risks from Mirror Life. (2025, June).

5 Secretary-General's Scientific Advisory Board. (2026). *Brief of the Scientific Advisory Board on Mirror Life*. United Nations. https://www.un.org/scientific-advisory-board/en/mirror-life

6 UK Government Office for Science. (2025). *Research and analysis Mirror life*. https://www.gov.uk/government/publications/mirror-life/mirror-life

7 Zimmer, C., & Hernandez, M. (2026, July 1). This Cell Feeds, Grows and Reproduces. And It's Manmade. *New York Times*. https://www.nytimes.com/interactive/2026/07/01/science/spudcells-synthetic-cell.html?unlocked_article_code=1.v1A.V2bj.4vqCug7GMUAA&smid=url-share

8 Rodriguez, L, (2026, January 13). *233—James Smith on why he quit everything to work on a biothreat nobody had heard of*. 80000hours. https://80000hours.org/podcast/episodes/james-smith-mirror-biology-catastrophe/

9 Adamala, K. P., Agashe, D., Belkaid, Y., de C. Bittencourt, D. M., Cai, Y., et al. (2024). Confronting risks of mirror life. *Science, 386*(6728), 1351-1353. https://doi.org/10.1126/science.ads9158

10 Xiao, W., Ernst, K.-H., Palotas, K., Zhang, Y., Bruyer, E., Peng, L., Greber, T., Hofer, W. A., Scott, L. T., & Fasel, R. (2016). Microscopic origin of chiral shape induction in achiral crystals. *Nature Chemistry, 8*(4), 326–330. https://doi.org/10.1038/nchem.2449

11 Pla-Mauri, J., Maull, V., Tabi, A., Shpilkina, Y., De Lorenzo, V., & Solé, R. (2026). Ecological constraints to mirror life. *Ecology.* https://doi.org/10.64898/2026.05.07.723461

12 Woodford, J. (2026, May 29). Mirror life: Scientists clash over threat of lab-engineered bacteria. *New Scientist*. https://www.newscientist.com/article/2528281-mirror-life-scientists-clash-over-threat-of-lab-engineered-bacteria/

Emerging Disruptive Technologies

Technology Preferences of Six Military Innovator Countries: An Updated Analysis of Security Policy Technology Strategies

The nature of warfare is continuously evolving – and it is doing so at an unprecedented pace. This chapter explores military technologies that may come to dominate future battlefields. By analyzing the military technology strategy papers of six countries, the study shows that the development of what are referred to as "emerging disruptive military technologies" is proceeding in parallel with significant improvements to classical "legacy" systems such as tanks, combat aircraft, and maritime assets. After being less debated in the early 2020s, technologies such as AI, autonomy, and uncrewed systems have attracted renewed attention. In addition, significant emphasis is placed on advanced missile technology. Yet, despite the fact that many players have launched their own initiatives, so far the destabilizing effect of conventional missiles has received relatively little international attention. The most striking finding, however, does not concern the technologies themselves, but rather how technology is discussed. While in countries such as Australia and Türkiye, major decisions regarding military technology have been publicly debated in a wide range of papers over the last decade, Germany has remained largely silent on the issue.

This is the third year that we have examined military technology developments in selected countries. As observed in previous years, several overlapping trends appear to be driving the current pace of military-technological innovation. First, artificial intelligence continues to develop apace, with AI increasingly finding its way into a multitude of weapon-system functions, leaving less and less room for the human operator. Second, Russia's ongoing war of aggression against Ukraine has created a highly dynamic field of innovation, particularly in the use of uncrewed systems across all military domains. The development and fielding time of new or at least improved systems, once a process that took months or even years, can now be completed in a matter of weeks. Third, the more general security environment has led to a significant increase in European armament spending, with a focus on state-of-the-art or next generation systems.

However, as we have also argued in previous CNTR Monitors, choices regarding military technology remain complex national decisions shaped by on perceived threats, individual technological capabilities, and the desire to retain an advantage in certain areas. The pursuit of prestige, national ethical attitudes, and specific interpretations of domestic and international law also play vital roles. Countries therefore inevitably differ in their technology priorities as well as their preferences.

This chapter picks up where we left off in 2025, delving deeper into the analysis of the five countries under scrutiny—Australia, Germany, the United Kingdom, Japan, and the United States—while adding a sixth case, Türkiye.[1] In contrast to last year's Monitor, where we focused on each country's two most recent national military technology strategy documents, this year we analyzed the most relevant technology papers published over the past decade. This broader time frame allows us to draw conclusions regarding longer-term trends and wider developments.

We examined in particular depth strategies with a technological emphasis that address innovation and research and development in military capabilities. This includes strategies that discuss technologies in singular domains, but excludes those centered exclusively on a single technology, thereby retaining this study's focus on technological portfolios rather than the operational details of individual technologies. For the same reason, strategies foregrounding organizational and managerial innovation of the defense industry as well as strategy implementation reports were not considered.

The corpus of strategies comprises 46 documents published between 2016 and 2025. The strategies were then analyzed using a coding scheme developed over several years, previous versions of which were also used in the 2025 and 2024 Monitors. The scheme comprises 19 main or top technological categories, and 137 subcategories, where more specific information was available. Parts of the strategies that mention or discuss specific technologies were coded and categorized with these subcategories. The coding was conducted by several coders using MAXQDA 24, and a number of reliability tests were carried out throughout the coding period to ensure intercoder reliability. Moreover, the results for the strategy papers were normalized at the document level to account for differences in length and the total number of codings per paper. The analysis is thus based on relative numbers, with each paper receiving equal weight in the accumulated treatment. Overall, the dataset comprises almost 13,500 coded technology text segments which were analyzed using RStudio.

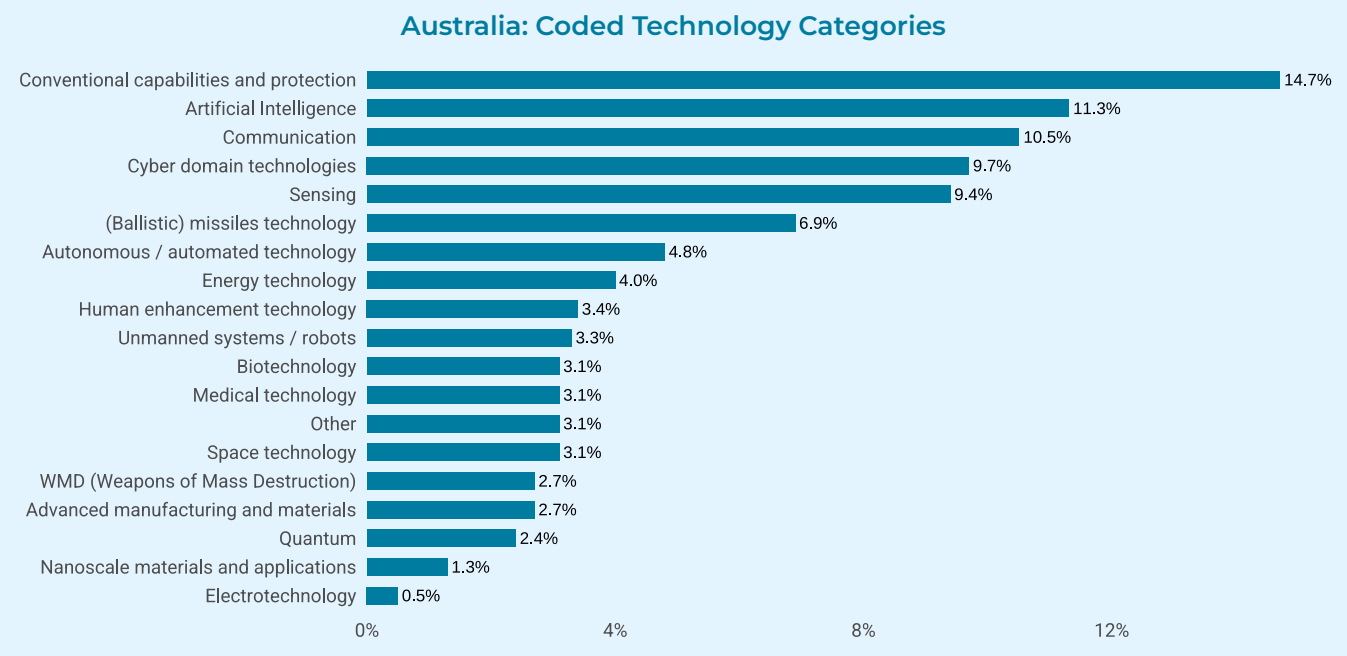
As in previous years, the following section will first present the results for each of the six countries individually, before adopting a comparative and an aggregated perspective.

Individual Country Summaries

Australia

Given Australia's self-perception as the "odd man out" in Asia, combined with a pronounced sense of insecurity, the country tends to view its adversaries' military-technological developments as potential threats. Australia has therefore developed its own robust high-tech industry, leading not only to the establishment of domestic defense companies but also to multiple American and European defense firms setting up Australian branches. Despite its relatively small military, Australia has become a key player in the field of forward-looking military strategy, and the fact that the country is represented with 13 strategy papers clearly demonstrates its long-term commitment to innovation.[2] The most important issues debated over the last 10 years include the improvement of already existing conventional capabilities (14.7%), artificial intelligence (11.3%), communication technology (10.5%), as well as cyber (9.7%) and sensing technology (9.4%), and (ballistic) missile technology (6.9%). With the exception of missile technology, the long-term view reveals a remarkably consistent focus over time. As explained in last

Australia's defense technology priorities show a consistent focus on strengthening conventional capabilities alongside investments in AI, communications, cyber, and sensing technologies. This reflects its strategic objective of monitoring and securing its wider operational environment.

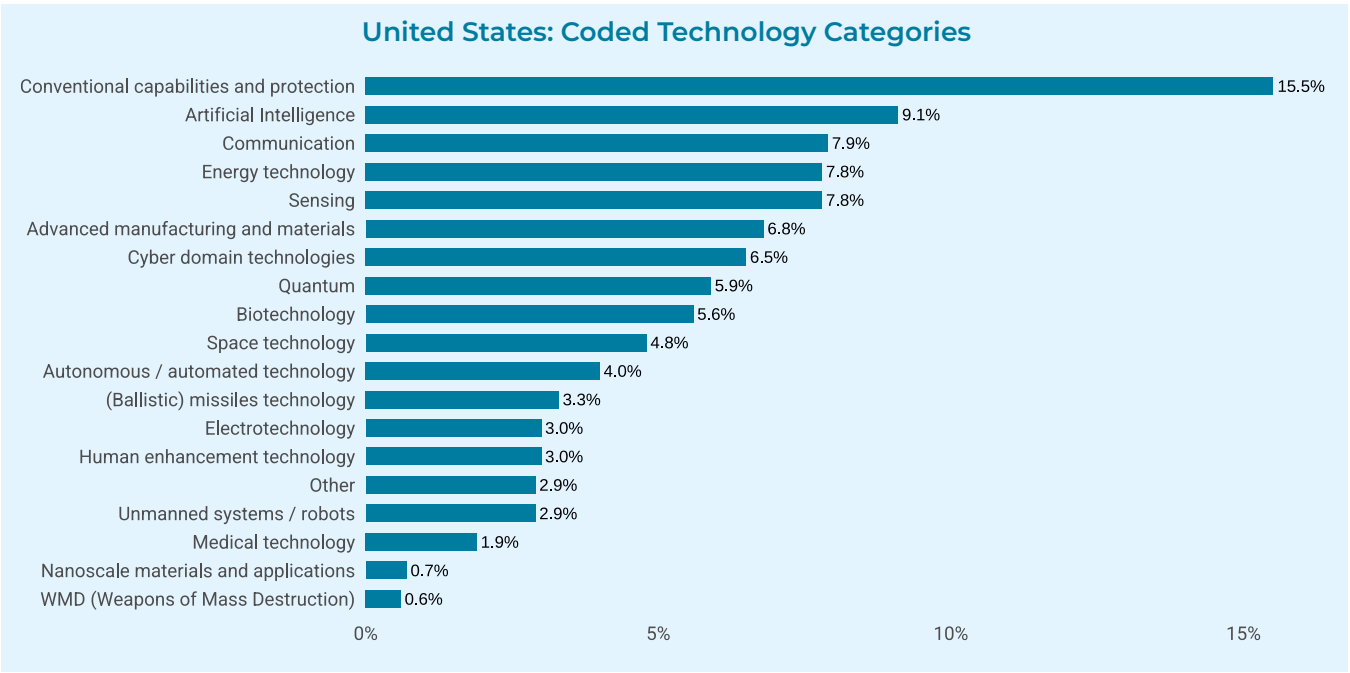


Relative frequencies of all coded top-level categories for Australia, with each paper contributing equally to the country-level values.

year's Monitor, the Australian military has to cover, monitor, and secure a vast maritime area. An integrated approach that incorporates new sensors, new and reliable means of communication, and AI-powered analytics helps them to accomplish this task. This also shows that Australia aims not to rely exclusively on allied partners to secure its coastal waters, while its investment in "traditional" military technologies simultaneously makes the country an attractive military partner to others.

United States

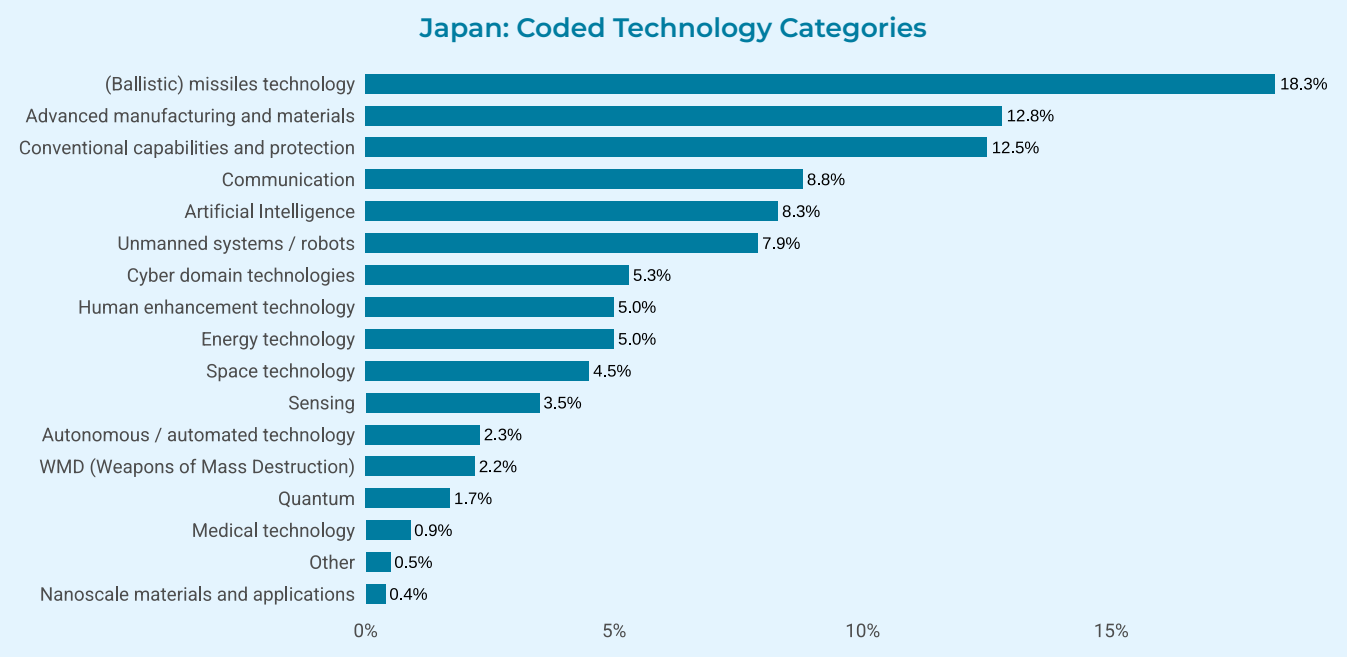
The United States is represented in this study with 10 strategy papers that meet the sampling criteria.[3] As the country with world's largest annual military budget, the U.S. has long been a proponent of substantial investment in advanced technologies and re-



Relative frequencies of all coded top-level categories for the U. S., with each paper contributing equally to the country-level values.

lated research. While conventional capabilities (15.5%), AI (9.1%), and communication technology (7.9%) are the three most frequently mentioned categories in their papers, energy technology ties for 4th place with sensing (both 7.8%). This broad perspective is also reflected in the additional technologies featured in the analyzed strategy papers: over the project period, the U.S. showed wide-ranging interest in technologies across scientific fields, including biotechnology (5.6%) and space engineering (4.8%). The trend analysis shows that in recent years, we can see a particularly marked increase in their interest in quantum technology, as well as in technologies used in the cyber and space domains, manufacturing, and AI. While U.S. strategies throughout the 2010s tended to focus on specific technologies, the trend now, as in other countries, appears to be toward a more diversified technological focus. Rather than highlighting a single technology as the primary revolutionary or disruptive force, an increasing number of technologies are receiving comparable attention, with smaller differences in their relative prominence across technological categories. This indicates that the U.S. aims to employ a diverse range of technologies in conjunction with one another, rather than focusing on individual platforms for deployment and research and development.

The U.S. technology focus integrates strong attention to conventional capabilities with a broad commitment to diverse critical technologies. This approach supports efforts to maintain a technological edge by advancing multiple technologies rather than relying on a single disruptive technology.



Relative frequencies of all coded top-level categories for Japan, with each paper contributing equally to the country-level values.

Japan

Japan, which is represented with three strategy papers in the analysis,[4] is the only one of the six states whose most frequently coded technology category is not conventional capabilities but, accounting for 18.3 percent of all codes, (ballistic) missile technology, which includes missiles and ballistic technologies, as well as missile defense. This alludes to the fact that the country's priority is the air domain as it prepares for both offensive and defensive maneuvers. With 12.8 percent of all codes, Japan clearly underscores the importance of modern manufacturing and advanced materials as its second strategic defense priority, closely followed by conventional capabilities and protection (12.5%). This makes Japan another outlier in comparison to the other countries whose strategies mention manufacturing and materials far less frequently. This suggests that, while other states communicate their military might through conventional capabilities—including the maritime, land, and air domains—as well as AI and communications, Japan prefers to showcase its capabilities through state-of-the art research, with a particular focus on technologies capable of reaching distant

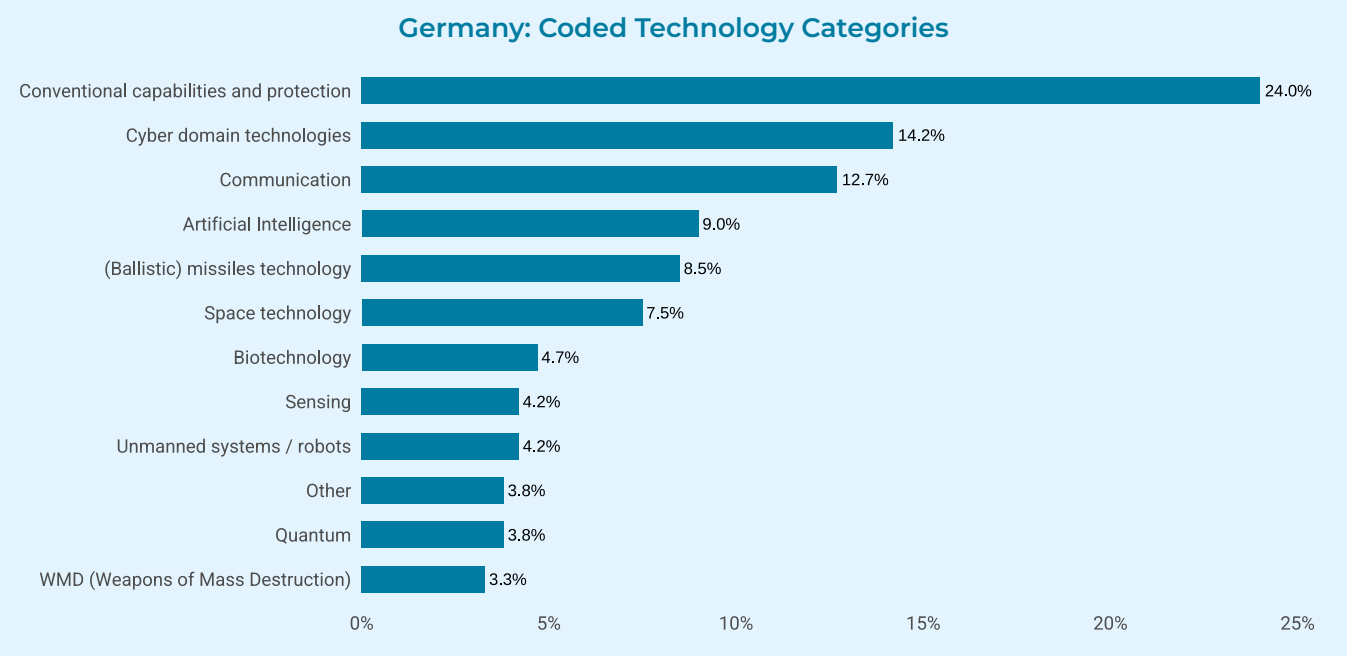
Japan stands out for placing missile technology at the center of its focus, alongside advanced manufacturing and materials as well as conventional capabilities. This suggests a tiered approach supported by communications, AI, and unmanned systems.

adversarial targets. Based on the relative distribution of the technologies mentioned, it appears that Japan has a tiered system of priorities, with three groups at the top: a) missile technologies; b) advanced manufacturing and conventional capabilities; and c) communication technologies, AI, and unmanned systems.

Germany

Germany's technology priorities focus on conventional capabilities, cyber, communications, and AI, with an emphasis on a defined set of technologies. The latest strategy more explicitly links these priorities to resilience, technological sovereignty, and reduced strategic dependencies.

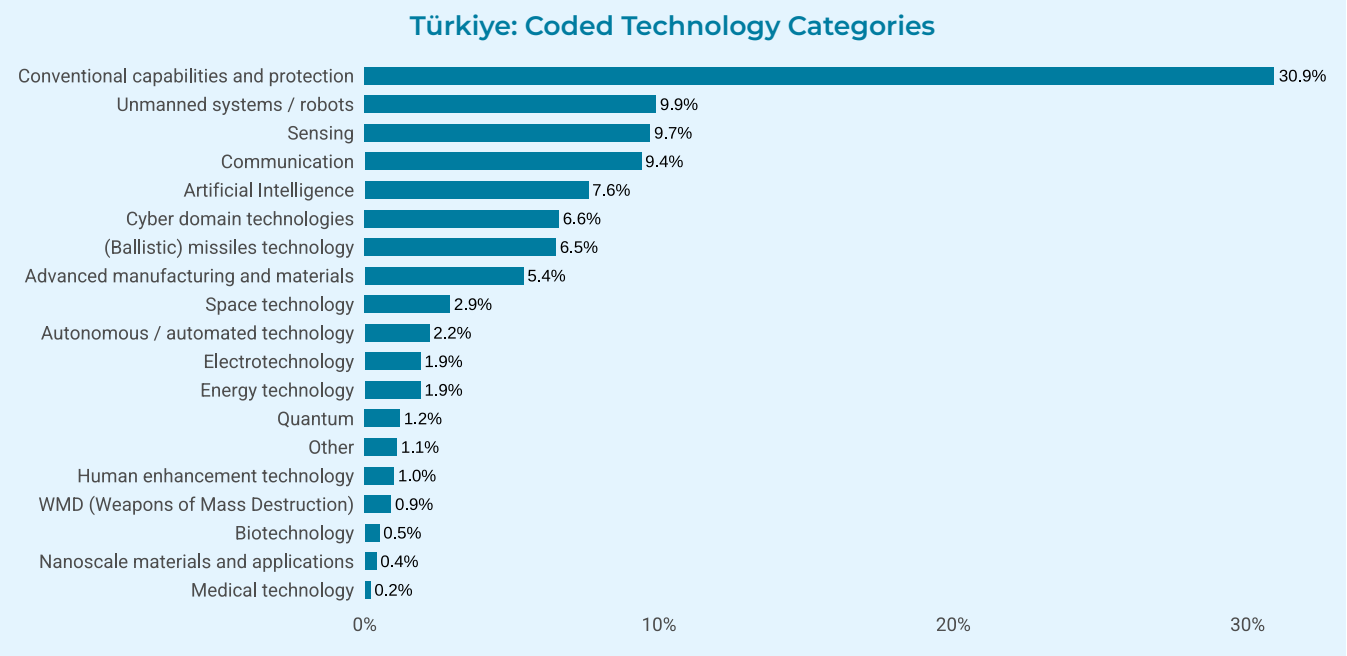
Germany's 2020 and 2024 security and defense strategies[5] both prioritize conventional capabilities and protection, which together account for 24 percent of all codes. This is followed by cybertechnologies (14.2%), artificial intelligence, and electronic warfare, as well as countermeasures as part of communication capabilities, with the additions in 2024 of quantum and space technologies. Unlike the U.S., Germany focuses on a dedicated set of key national technologies linked to military capability and industrial competitiveness, rather than a broad range of technologies. However, this implicit reliance on external capabilities is acknowledged in the 2024 strategy, as are the risks associated with technological dependencies and supply-chain vulnerabilities. The 2024 strategy paper therefore underlines the need to cooperate with EU and NATO partners, while also calling for greater resilience and technological sovereignty, particularly through a stronger emphasis on software, IT products, and cybersecurity. Overall, Germany's strategic discourse has evolved from a predominantly industrial perspective toward a broader focus on security, resilience, and digital sovereignty. Although no new strategy has been published since 2024, recent developments—including temporary U.S. export controls affecting AI software and tools, as well as the broader U.S.–China competition over critical technologies—are likely to reinforce Germany's emphasis on reducing strategic dependencies, diversifying supply chains, and developing sovereign European capabilities in AI, cloud infrastructure, semiconductors, cybersecurity, and other dual-use technologies. This tendency is already reflected in Germany's support for, and votes in favor of, corresponding measures at the EU level.



Relative frequencies of all coded top-level categories for Germany, with each paper contributing equally to the country-level values.

Türkiye

For the period analyzed, Türkiye published seven strategy papers.[6] The Turkish documents are notably extensive and highly detailed, reflecting the country's strong commitment to innovation across multiple domains. Consistent with the overall results, Türkiye's most coded technology category, making up one-third, is conventional ca-



Relative frequencies of all coded top-level categories for Türkiye, with each paper contributing equally to the country-level values.

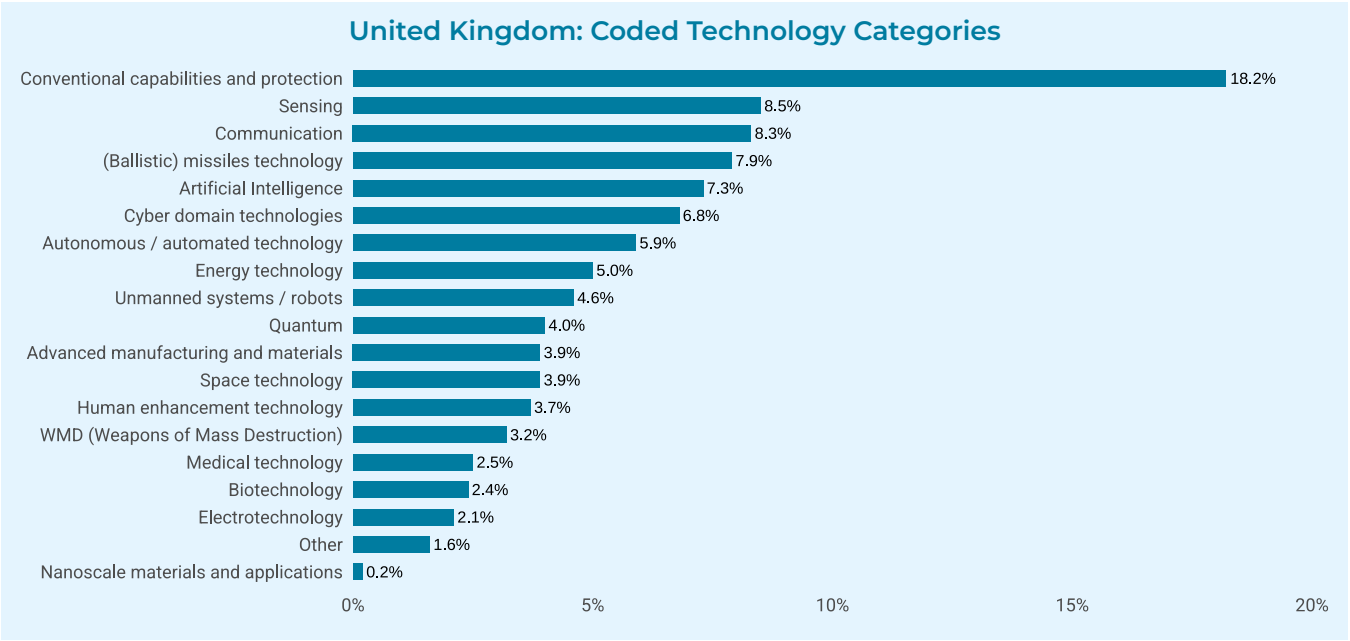
pabilities and protection with a focus on rotary- and fixed-wing aircraft and vehicle systems for land and sea (30.9%). Thereafter, Türkiye likewise refers to different complementary technologies, each making up 9 to 10 percent of the coded technology categories. Notably, of these, unmanned systems and robots feature especially prominently, with a particular focus on uncrewed aerial vehicles (UAVs). This category is followed by sensing and communication technology. In terms of sensing, the main priorities are radar technology and optical sensors, while for communication technology the emphasis is on electronic warfare and electromagnetic technologies. Artificial intelligence, accounting for 7.6 percent of coded categories, further underscores Türkiye's drive toward more capable and intelligent systems. Together, these priorities reflect a technology agenda spanning the full spectrum from conventional platforms to emerging and enabling technologies.

Türkiye's technology agenda is anchored in conventional capabilities, with a particularly strong emphasis on unmanned systems and UAVs. The prominence of sensing, communications, cyber, and AI points to a broad-based approach linking established platforms with other relevant technologies.

United Kingdom

In this analysis, the UK is represented with 11 strategy papers.[7] As in many other European countries, the UK's perspective on military security has been heavily influenced by the Russia's full-scale invasion of Ukraine in early 2022. Before this, the focus had been—again, as in many other European countries—on global crisis intervention and the fight against international terrorism, with Brexit bringing Britain closer to the U.S. and further away from its former EU partners. After 2022, national and alliance defense have once again taken center stage, with a clear focus on resilience and "warfighting readiness," as the 2025 Strategic Defence Review explicitly stated.[8] Like other countries, the UK recognizes the need to invest in the latest technologies to keep the military edge. It is therefore no surprise that the focus on improving conventional capabilities by far exceeds all other technology categories (18.2%). In addition, the idea of archiving a transparent battlefield is also reflected in the British debate, with sensing (8.5%) and communication technologies (8.3%) emerging as the next two most important categories. Artificial intelligence, which is discussed not only in terms of its broader application across different systems but also in connection with data science as well as machine and deep learning, is also an issue that plays a key role in the debate (7.3%). This suggests that the UK continues to prioritize AI as a strategic technology for the development of future defense capabilities. As in all other countries, interest in missile technology also ranks highly (7.9%).

Conventional capabilities remain at the core of the UK's technology focus, while sensing, communications and missile technology form important complementary areas. This combination reflects a stronger orientation toward readiness, resilience, and maintaining a technological edge.



Relative frequencies of all coded top-level categories for the UK, with each paper contributing equally to the country-level values.

The Macro Perspective

Looking at the top-10, top-level categories in each country, a clear picture emerges that is consistent with our 2025 analysis. The most important issue for all the countries analyzed has been the improvement of their already existing conventional capabilities. In addition, improving battlefield communication and command-and-control systems, alongside the use of AI for analysis and enhanced sensing capabilities, has been an important theme for all countries, reflecting the idea, variously termed "network-centric warfare" or one of the "revolutions in military affairs." What had been merely a military technology vision in the late 1990s and early 2000s is now taking shape. In addition, anything "cyber" has also featured prominently over the last 10 years. What is surprising, however, is the relatively low scoring of uncrewed and autonomous/automated systems, although aggregated numbers do have the tendency to obscure interesting details.

Extending our study period has enabled us to gain a clearer picture of the trends—although, given the relatively small number of strategy papers and their erratic publication cycle, these findings should be interpreted with a degree of caution.

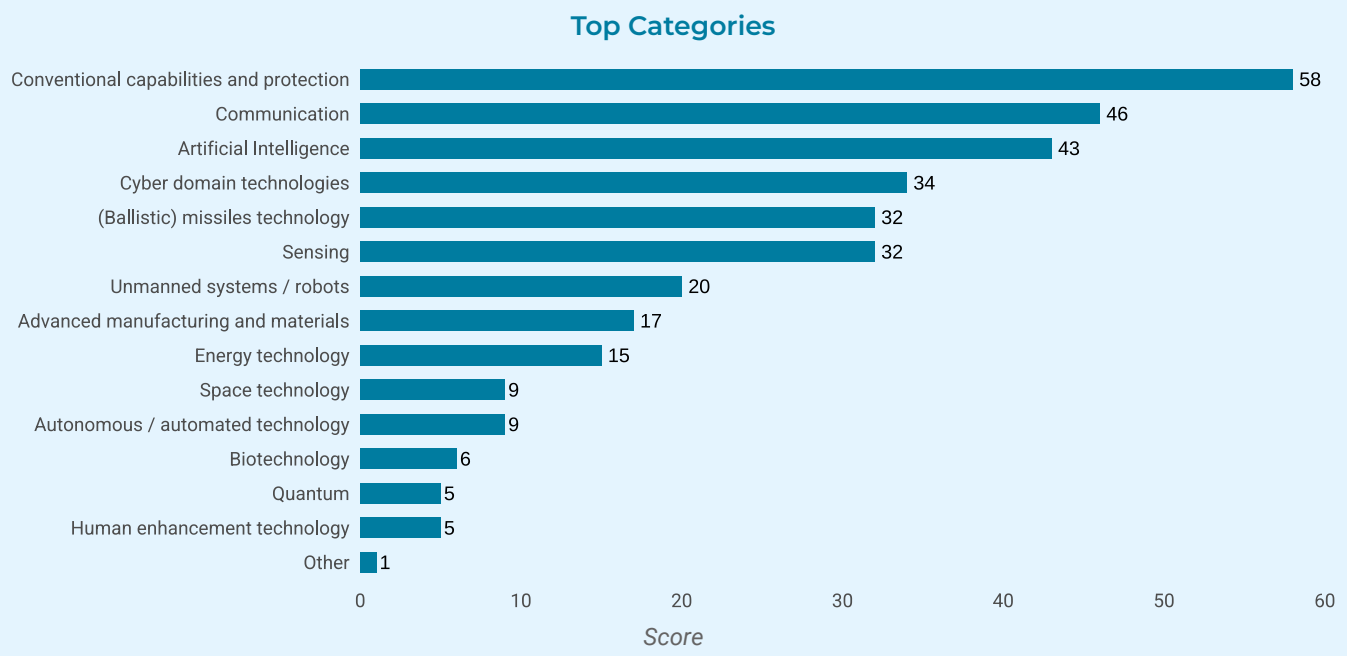
The first interesting observation is that ballistic missiles are rated lower throughout the 10-year observation period than they have been in recent years, although they do show a slight upward trend. The impact of the war against Ukraine is particularly clear here. Although such missiles were certainly discussed as a relevant technology during the 2010s, the much greater interest in them has been driven by Russia's systematic use of these weapons against Ukraine. These deployments have not only highlighted the need for effective defense but also sparked interest in acquiring such systems oneself—whether for deterrence or to strike high-value military targets.

The trend analysis also captures the heightened interest in the military use of AI, which peaked around 2020 before levelling off, followed by a slight upward trend toward 2025. It will be interesting to see whether AI experiences a further revival in future strategy papers when the use of large language models (LLMs) and (decision) support or battle management systems are expected to feature more prominently.

While, as reported above, uncrewed systems score relatively low, there does seems to be a positive trend with these systems gaining greater prominence due to their relevance in the Russia-Ukraine War. Although uncrewed systems, particularly uncrewed aerial vehicles, proliferated worldwide in the 2010s, most Western and allied militaries had yet to develop concrete strategies for integrating these systems into their military operations. This has shifted markedly in response to feedback from the battlefields in Ukraine, as well as Ukraine's long-range Ukrainian attacks against strategic assets in Russia.

When it comes to cyber systems, we still see a high level of overall relevance, although there has been a slight decline since its peak in 2020. While almost all the countries analyzed see cybertechnology as a highly important capability in their arsenals, it has lost relative importance with the return of traditional battlefield operations and the "hands-on war" as seen in Ukraine. This development could also have been encouraged by the growing perception of cyberthreats as domestic threats and threats to critical infrastructure, thereby entering a hybrid domain that shifts the focus away from purely military operations toward threats posed by both state and non-state actors. Lastly, while we are seeing a steady increase in the significance of human enhancement technologies, their current level of adoption remains low. It is likely that this technology, which so far has been discussed only in niche circles, will become more relevant in the future, particularly as brain–computer interfaces being to enable faster interaction between operators and machines.

One more meta-observation is that modern military technology is debated quite differently in the countries under scrutiny, both in quantitative as well as qualitative terms. Over the 10-year study period, Germany published two(!) technology-related strategy papers, Australia—which is significantly less powerful militarily—published 13, and Türkiye, which is roughly on par with Germany in terms of military budget, published at least seven—all of which were highly detailed documents. Japan, on the other hand, seems to have adopted a similar approach to Germany in terms of output, producing just three papers during the observation period. Of course, it must be taken into account that, in addition to the publicly available papers, other confidential documents



Scores for top-level categories appearing among the top ten in each country, based on a simple scoring system: the highest-ranked category in each country receives 10 points, the second-highest 9 points, and so on. Each score represents the sum of national scores across countries.

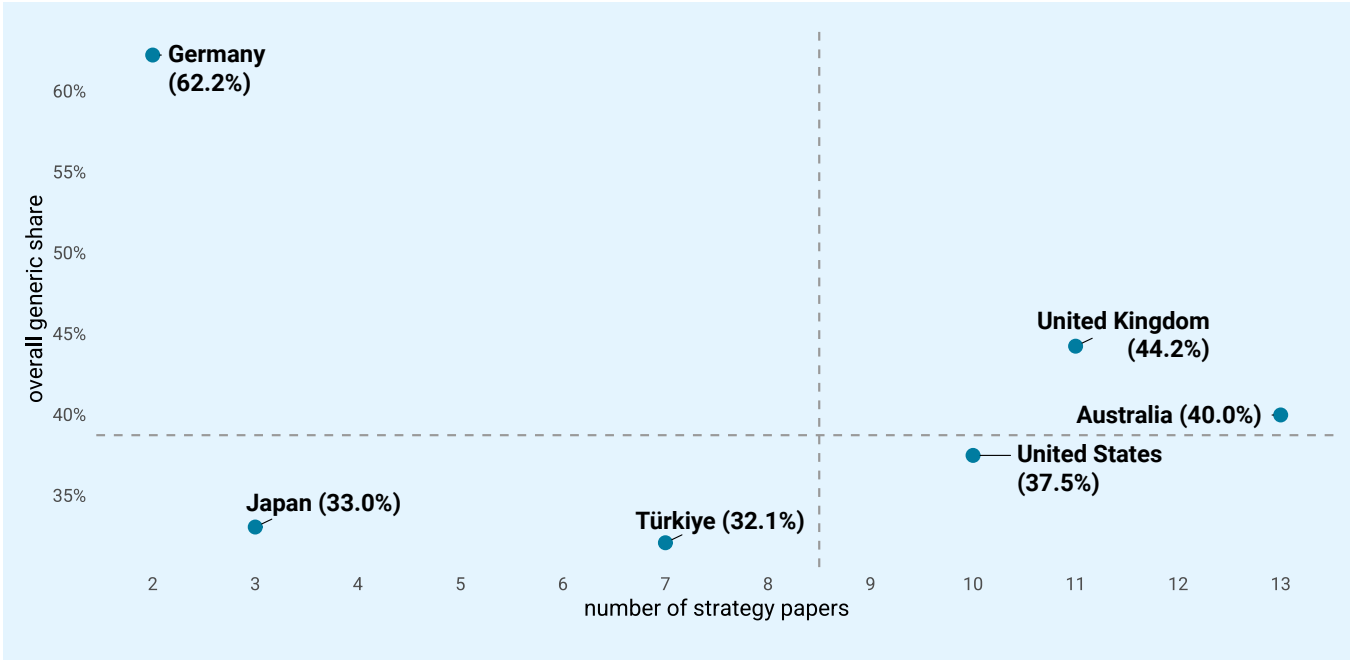
are likely to exist—and Germany is certainly no exception in this regard. However, this probably applies to all the countries under study and does not explain the lack of public documents in Germany and Japan.

Yet it is not only the volume of publicly available papers that varies across countries. How they describe their technological ambitions also differs significantly. Some remain on the "buzzword" level (using terms such as quantum technology, AI, etc.), while others describe their investment plans in more detail (referring for example to quantum sensing, radar technology, or command-and-control systems). While in the German case, 62.2 percent of all codes stay on a generic level, this is only true for 32.1 percent of all coded segments in Türkiye and 40 percent in Australia. Interestingly, while Japan has published only a small number of papers, the level of detail they contain is high, with just 33 percent of the codes remaining on the generic level. This reveals markedly different approaches to communicating with the public on military technology ambitions and investment.

Conclusion

When Western countries and allies discuss the future role of cutting-edge technology in their militaries, they do not limit themselves to technologies that are traditionally regarded as "emerging" and "disruptive," but instead aim for a mixture of familiar and tried-and-tested concepts (albeit in significantly improved form) that are complemented or enhanced by new technologies. This is all the more true in light of the scenes from the battlefields of Ukraine and Gaza, which are being closely followed by the whole world. Although some countries attach greater importance to certain technologies than others, overall we are seeing a convergence toward a common view of the relative importance of individual technologies within the technology mix. At this point, we would also like to reiterate our warning from last year to pay closer attention to the growing interest in ballistic missiles.

When it comes to how countries inform their populations about their technological goals, there is also room for improvement, at least in some countries. While Australia provides its population with very comprehensive and detailed information, the opposite is true in Germany. The German discourse differs markedly from that in all



2x2 Matrix/Quadrant chart: number of analyzed strategy papers vs. overall generic coded technologies by country.

other countries studied. Apparently, the Bundeswehr and the Ministry of Defense are reluctant to subject the public to debates on technology. Given the current security situation, however, this reticence seems outdated and could prove problematic when decisions have to be made regarding the development and procurement of new weapons technologies that have not been adequately communicated. The German public, too, should be informed about future military planning—including in the technical sphere. Greater transparency is needed, and our partners are demonstrating that this is possible without revealing military secrets. We simply must not shy away from open discussion or be afraid of potential criticism.

Abdullah Kasim Korkusuz

PRIF – Peace Research Institute Frankfurt

Niklas Schörnig

PRIF – Peace Research Institute Frankfurt

Thomas Reinhold

PRIF – Peace Research Institute Frankfurt

Liska Suckau

PRIF – Peace Research Institute Frankfurt

1 Case selection is based on, first, public availability of relevant strategy papers, and, second, language resources within the team.

2 Australian Government. Department of Defence. (2016). *2016 Defence Industry Policy Statement*. https://apo.org.au/sites/default/files/resource-files/2016-02/apo-nid93621.pdf
Australian Government. Department of Defence. (2016). *Shaping Defence Science and Technology in the Intelligence Domain 2016-2026*
Australian Government. Department of Defence. (2016). *Shaping Defence Science and Technology in the Land Domain 2016-2036.* https://www.dst.defence.gov.au/sites/default/files/articles/documents/Shaping_Land_Defence_Science.pdf
Australian Government. Department of Defence. (2016). *Shaping Defence Science and Technology in the Maritime Domain 2016-2026.*
Australian Government. Department of Defence. (2017). *Shaping Defence Science and Technology in the Joint Domain 2017-2021*.
Australian Government. Department of Defence. (2017). *Shaping Defence Science and Technology in the Aerospace Domain 2017-2027*.
Australian Government. Department of Defence. (2018). *2018 Defence Industrial Capability Plan*. https://www.wa.gov.au/system/files/2020-10/DefenceIndustrialCapabilityPlan-web.pdf
Australian Government. Department of Defence. (2018). *National Science and Technology Policy and Priorities.* https://www.dst.defence.gov.au/sites/default/files/attachments/documents/NS%20S%26T%20policy%20and%20priorities.pdf
Australian Government. Department of Defence. (2020). *Defence Science and Technology Strategy 2030.*
Australian Government. Department of Defence. (2020). *National Security Science and Technology Priorities 2020.* https://www.dst.defence.gov.au/sites/default/files/attachments/documents/National%20Security%20Science%20and%20Technology%20Priorities.pdf
Australian Government. Department of Defence. (2024). *Defence Industry Development Strategy.* https://www.defence.gov.au/sites/default/files/2024-02/Defence-Industry-Development-Strategy.pdf
Australian Government. Department of Defence. (2024). *Defence Innovation, science and Technology Priorities.* https://www.defence.gov.au/sites/default/files/2024-09/Defence-IST-Priorities_0.pdf
Australian Government. Department of Defence. (2024). *Defence Innovation, Science and Technology Strategy.* https://www.defence.gov.au/sites/default/files/2025-03/Defence-IST-Strategy-2024-PRO2.pdf

3 U.S. Department of Defense. (2016). *Joint non-lethal weapons program science and technology strategic plan 2016–2025.*
U.S. Marine Corps. (2018). *2018 U.S. Marine Corps S&T Strategic Plan.* https://www.microwaves101.com/uploads/2018-USMC-S-and-T-Strategic-Plan.pdf
Department of the Navy. (2018). *Naval Research and Development.*
U.S. Air Force. (2019). *Science and Technology Strategy.* https://www.af.mil/Portals/1/documents/2019%20SAF%20story%20attachments/Air%20Force%20Science%20and%20Technology%20Strategy.pdf
U.S. National Science and Technology Council. (2022). *Critical and Emerging Technologies List Update.* https://pscouncil.org/a/Resources/2023/NSTC_ETListUpdatedFeb2022_20230130.aspx
U.S. Department of Defense. (2023). *National Defense Industrial Strategy.* https://www.waru.edu/system/files/webform/documents/2721/NDIS%20FINAL.pdf
U.S. Department of Defense. (2023). *National Defense Science & Technology Strategy 2023*. https://media.defense.gov/2023/May/09/2003218877/-1/-1/0/NDSTS-FINAL-WEB-VERSION.PDF

U.S. National Science and Technology Council. (2024). *Critical and Emerging Technologies List Update.* https://bidenwhitehouse.archives.gov/wp-content/uploads/2024/02/Critical-and-Emerging-Technologies-List-2024-Update.pdf
U.S. Department of the Navy. (2024). *Naval Science and Technology Strategy.* https://media.defense.gov/2024/Apr/09/2003434522/-1/-1/0/2024%20NAVAL%20ST%20STRATEGY_FINAL.PDF
U.S. Department of War. (2025). *Critical Technology Areas.* https://www.cto.mil/wp-content/uploads/2025/11/CTA-One-Pager-Option-Nov2025.pdf

4 Ministry of Defense, Japan. (2016). *Defense Technology Strategy.* https://www.mod.go.jp/atla/en/policy/pdf/defense_technology_strategy.pdf
Ministry of Defense, Japan. (2022). *Defense Buildup Program.* https://www.mod.go.jp/j/policy/agenda/guideline/plan/pdf/program_en.pdf
Japan Ministry of Defense. (2023). *Defense Technology Guideline 2023.* https://www.mod.go.jp/atla/guideline2023/assets/pdf/technology_guideline2023.pdf

5 Bundesregierung der Bundesrepublik Deutschland. (2020). *Strategiepapier der Bundesregierung zur Stärkung der Sicherheits- und Verteidigungsindustrie.* https://www.bundeswirtschaftsministerium.de/Redaktion/DE/Downloads/S-T/strategiepapier-staerkung-sicherits-und-verteidigungsindustrie.pdf?__blob=publicationFile&v=1
Bundesregierung der Bundesrepublik Deutschland. (2024). *Nationale Sicherheits- und Verteidigungsindustriestrategie.* https://www.bmvg.de/resource/blob/5865332/d4d0d9ab55edde72a11cee2a3ca59d3b/nationale-sicherheits-und-verteidigungsindustriestrategie-data.pdf

6 Savunma Sanayii Müsteşarlığı. (2017). *Stratejik Plan 2017-2021.* http://www.sp.gov.tr/upload/xSPStratejikPlan/files/5cNrk+SSM_STRATEJIK_PLAN_2017-2021.pdf
T.C. Cumhurbaşkanlığı Savunma Sanayii Müsteşarlığı. (2018). *Savunma Sanayii Sektörel Strateji Dokümanı.* https://www.ostimsavunma.org/api/document/download/344
T.C. Cumhurbaşkanlığı Savunma Sanayii Başkanlığı. (2019). *2019 2023 Stratejik Plan.* http://www.sp.gov.tr/upload/xSPStratejikPlan/files/NAUvQ+Savunma_Sanayii_Baskanligi_2019-2023_Donemi_Stratejik_Plani.pdf
T.C. Cumhurbaşkanlığı Savunma Sanayii Başkanlığı. (2020). *2019 2023 Stratejik Plan.* http://www.sp.gov.tr/upload/xSPStratejikPlan/files/z8ySu+Savunma_Sanayii_Bas_2019-2023_Donemi_Stratejik_Plani_Guncellenmis_.pdf
Türkiye Cumhruiyeti Cumhurbaşkanlığı Savunma Sanayii Başkanlığı. (2023). *2023 2024 Savunma Sanayii Sektörel Strateji Dokümanı.* http://www.sp.gov.tr/upload/xSPTemelBelge/files/1Mmgq+Savunma_Sanayi_Sektorel_Strateji_Belgesi_2023-2027_.pdf
Türkiye Cumhuriyeti Cumhurbaşkanlığı Savunma Sanayii Başkanlığı. (2024). *2024-2028 Savunma Sanayii Sektörel Strateji Dokümanı.* https://sasadem.com/wp-content/uploads/2025/04/SAVUNMA-SANAYI-STRATEJISI-2024-2028.pdf
Türkiye Cumhuriyeti Cumhurbaşkanlığı Savunma Sanayii Başkanlığı. (2024). *2024 2028 Stratejik Planı.* http://www.sp.gov.tr/upload/xSPStratejikPlan/files/6BoZs+SSB_2024-2028_Stratejik_Plani.pdf

7 United Kingdom Ministry of Defence. (2016). *Advance through Innovation: The Defence Innovation Initiative.* https://assets.publishing.service.gov.uk/media/5a7f6cad40f0b6230268f6b9/MOD_SB_Innovation_Initiative_Brochure_v21_web.pdf
United Kingdom Ministry of Defence. (2017). *Science and Technology Strategy 2017.* https://assets.publishing.service.gov.uk/media/5a81f742ed915d74e3401024/Science_and_Technology_Strategy_lowres.pdf
United Kingdom Ministry of Defence. (2019). *Defence Technology Framework.* https://assets.publishing.service.gov.uk/media/5d763a3ce5274a27d0fdd8cf/20190829-DTF_FINAL.pdf
United Kingdom Ministry of Defence. (2019). *Defence Innovation Priorities.* https://assets.publishing.service.gov.uk/media/5d7a5708e5274a1ff2f0835c/20190906-InnovationPrioritiesPub_Final_.pdf
United Kingdom Ministry of Defence. (2020). *Science and Technology Strategy 2020.* https://assets.publishing.service.gov.uk/media/5f8dc7ece90e0727ca313eb6/20201019-MOD_ST_Strategy_2020_v1-23.pdf
HM Government. (2021). *Defence and Security Industrial Strategy.* https://assets.publishing.service.gov.uk/media/60590e988fa8f545d879f0aa/Defence_and_Security_Industrial_Strategy_-_FINAL.pdf
United Kingdom Ministry of Defence. (2022). *Land Industrial Strategy.* https://assets.publishing.service.gov.uk/media/6284ad0ae90e071f67b27fc0/Land_Industrial_Strategy.pdf
United Kingdom Ministry of Defence. (2022). *Setting the context: current capability plans to enable Integrated Force 30.* https://assets.publishing.service.gov.uk/media/62d7d184e90e071e7b1310e1/MOD_Defence_Capability_Framework_Annex_Accessible_Jul22.pdf
United Kingdom Ministry of Defence. (2022). *The Defence Capability Framework.* https://assets.publishing.service.gov.uk/media/62d7d1668fa8f50c0a8a4029/MOD_Defence_Capability_Framework_Accessible_Jul22.pdf
United Kingdom Ministry of Defence. (2024). *Ministry of Defence's Science and Technology Portfolio.* https://www.gov.uk/government/publications/defence-science-and-technology-programmes-and-projects/ministry-of-defences-science-and-technology-portfolio#introduction
United Kingdom Ministry of Defence. (2025). *Defence Industrial Strategy.* https://assets.publishing.service.gov.uk/media/68bea3fc223d92d088f01d69/Defence_Industrial_Strategy_2025_-_Making_Defence_an_Engine_for_Growth.pdf

8 United Kingdom Ministry of Defence. (2025). *The Strategic Defence Review 2025 – Making Britain Safer: secure at home, strong abroad.* https://assets.publishing.service.gov.uk/media/683d89f181deb72cce2680a5/The_Strategic_Defence_Review_2025_-_Making_Britain_Safer_-_secure_at_home__strong_abroad.pdf

ANNEX

Treaties and Instruments Related to Arms Control, Disarmament, and Non-Proliferation of Nuclear Weapons

International Treaties Related to Nuclear Weapons Disarmament and Non-Proliferation

The Treaty on the Non-Proliferation of Nuclear Weapons (NPT), opened for signature on the 1st of July 1968, is an agreement which seeks to prevent nuclear proliferation, promote nuclear disarmament, and guide the peaceful use of nuclear energy. The treaty entered into force on the 5th of March 1970 and has 191 state parties.

The NPT is the cornerstone of global non-proliferation efforts which foster the peaceful use of nuclear energy and the furtherance of nuclear disarmament. Under the NPT, nuclear-weapons states (NWSs) have committed, inter alia, not to consign nuclear weapons, nor be party in any manner to the manufacture or acquisition of nuclear weapons by non-nuclear-weapons states (NNWSs). NWSs are states that have both manufactured and exploded a nuclear weapon or other nuclear explosive devices, before the 1st of January 1967. Similar to the NPT, other region specific nuclear weapons treaties require participating states to conclude a comprehensive safeguards agreement with the International Atomic Energy Agency (IAEA).

The Treaty on the Prohibition of Nuclear Weapons (TPNW), opened for signature on the 20th of September 2017, is the first legally binding international agreement to comprehensively prohibit any nuclear weapon activities, with the ultimate goal being their total elimination. Under this treaty, state parties are also prohibited from stockpiling, stationing or using nuclear weapons as a threat. The TPNW makes provisions for assistance to individuals affected by the use or testing of nuclear weapons. In regions affected by contamination due to testing or use of nuclear weapons, environmental remediation measures are required from the state party with the jurisdiction in the region. The TPNW entered into force on the 22nd of January 2021 and has 75 state parties. An additional 26 parties have signed but not yet ratified their commitment. Opposition to the annual UN General Assembly resolution on the TPNW comes from NWSs.

The Partial Test Ban Treaty (PTBT), formally known as the Treaty Banning Nuclear Weapon Tests in the Atmosphere, in Outer Space and Under Water, opened for signature on the 5th of August 1963. Only nuclear tests conducted underground were exempted from prohibition under this treaty. The PTBT entered into force on the 10th of October 1963 and has 126 state parties.

The Comprehensive Nuclear-Test-Ban Treaty (CTBT), opened for signature on the 10th of September 1996, prohibits nuclear weapons test explosions as well as any other nuclear explosions, for both civilian and military purposes, in any environment. The CTBT has 179 state parties, but has not yet entered into force given that 9 required states have yet to ratify it. Commensurate with the entrance into force of the CTBT would be a comprehensive verification scheme. Typical measurements used to verify a nuclear test include:

- shockwaves from nuclear explosions can be detected using seismic instruments as they travel through the ground
- sound waves generated by a nuclear detonation in or near a body of water can be detected at hydroacoustic stations
- infrasonic waves that propagate through the atmosphere subsequent to a nuclear detonation can be detected on sensors that register frequencies inaudible to the human ear
- radioactive particles and gases produced by a nuclear detonation and released into the atmosphere can be monitored at various radionuclide stations
- satellite imagery can be used to capture the craters and debris fields left behind after a nuclear detonation, as well as to monitor the activity around known nuclear testing and storage facilities

In order to verify nuclear testing on a global scale, an international observational network is required. The International Monitoring System (IMS) is one such network, specifically established to monitor compliance with the Comprehensive Nuclear-Test-Ban Treaty when it enters into force. The IMS is comprised of 321 monitoring stations across the globe, with an additional 16 radionuclide laboratories. Data from this network is processed at a 24-hour monitoring facility, the International Data Centre (IDC). The IDC, with headquarters in Vienna, screens the data for possible nuclear testing events and presents this information to member states. Upon inspection of information presented by the IMS, a member state that perceives a possible breach of a treaty may proceed further to request an on-site inspection. An international team of inspectors can then be dispatched to investigate the suspected breach at the site.

The Antarctic Treaty, opened for signature on the 1st of December 1959, protects the peaceful status of the Antarctic Continent. The testing of nuclear weapons and the disposal of radioactive waste material is prohibited on the Antarctic content. This treaty has 55

state parties and entered into force on the 23rd of June 1961.

The Outer Space Treaty, opened for signature on the 27th of January 1967, prohibits the stationing of weapons of mass destruction (WMDs) in space, or in orbit around any celestial body. Importantly, this treaty recognizes that no one particular nation may claim sovereignty over outer space or any celestial body. The Outer Space Treaty, more formally known as the Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies, has 115 state parties and entered into force on the 10th of October 1967.

The Seabed Arms Control Treaty, formally known as the Treaty on the Prohibition of the Emplacement of Nuclear Weapons and Other Weapons of Mass Destruction on the Sea-Bed and the Ocean Floor and in the Subsoil Thereof, was opened for signature on the 11th of February 1971. The treaty prohibits the stationing of WMDs on the ocean floor beyond a 12-mile (22.2 kilometer) coastal zone, has 94 state parties, and entered into force on the 18th of May 1972. It allows state parties to observe other state parties' seabed activities beyond the coastal zone in order to verify compliance with the treaty.

The Fissile Material Cut-off Treaty (FMCT), the terms of which are yet to be negotiated, is proposed to be a prohibition on further production of fissile material which would be used in nuclear weapons or other explosive devices.

Region Specific Nuclear Weapon Free Zone Treaties

The Treaty of Tlatelolco, otherwise known as the Treaty for the Prohibition of Nuclear Weapons in Latin America and the Caribbean, opened for signature on the 14th of February 1967 and prohibits the testing, use, manufacture, production or acquisition by any means whatsoever of any nuclear weapons as well as the receipt, storage, installation, deployment and any form of nuclear weapons possession. The Treaty of Tlatelolco has 33 state parties and entered into force on the 22nd of April 1968.

The Treaty of Rarotonga, opened for signature on the 6th of August 1985 and otherwise known as the South Pacific Nuclear Free Zone Treaty, prohibits the use, testing, and possession of nuclear weapons within the borders of the South Pacific. This treaty has 13 state parties and entered into force on the 11th of December 1986.

The Treaty of Bangkok, otherwise known as the Treaty on the Southeast Asia Nuclear Weapon-Free Zone Treaty, opened for signature on the 15th of December 1995 and is a commitment to keep the Southeast Asian region free from nuclear and other weapons of mass destruction. This treaty entered into force on the 28th of March 1997 and has 10 state parties.

The Treaty of Pelindaba, otherwise known as the African Nuclear-Weapon-Free Zone Treaty, opened for signature on the 11th of April 1996 and prohibits the research, development, manufacture, stockpiling, acquisition, testing, possession, control, or stationing of nuclear explosive devices. The Treaty of Pelindaba has 43 state parties and entered into force on the 15th of July 2009.

The Treaty of Semipalatinsk, otherwise known as the Treaty on a Nuclear-Weapon-Free Zone in Central Asia, opened for signature on the 8th of September 2006 and is a commitment to refrain from the manufacture, acquisition, testing, or possession of nuclear weapons. The treaty of Semipalatinsk has 5 state parties and entered into force on the 21st of March 2009.

Bilateral Agreements between the United States and the Soviet Union surrounding arms control

The New Strategic Arms Reduction Treaty (New START), signed on the 8th of April 2010, **expired on the 5th of February 2026.** The expiration of the treaty marked the end of United States-Soviet/Russian arms control that had begun in 1969 during the launch of SALT I negotiations.

The New START treaty was designed to limit the number of deployed strategic nuclear warheads to 1,550 and deployed missiles and bombers to 700, with deployed and non-deployed launchers (missile tubes and bombers) limited to 800. The treaty allowed for ten annual on-site inspections of deployed warheads as well as delivery systems (deployed and non-deployed) at intercontinental ballistic missile (ICBM) bases, submarine bases and air bases. The treaty did not place limits on the number of non-deployed missiles. The locations of non-deployed missiles were however continuously monitored, and the missiles labeled with "unique identifiers". The missiles were to be stored at specified facilities, which are separated from deployment sites. Also provided for were eight annual on-site inspections at facilities which house only non-deployed delivery systems.

The verification of compliance with the treaty, was a crucial item under negotiation between the United States and the Russian Federation for the New START agreement. The agreed upon verification regime included detailed definitions of items limited by the treaty (a database of numbers, types and locations of items) and provisions on the use of National Technical Means (NTM). Additionally, the broadcast of telemetry as well as the exchange of recordings was required annually, for up to five missile tests per party to the treaty. New START did not prohibit modernization efforts, provided that these efforts did not contravene treaty stipulations (Article V).

The New START treaty entered into force on the 5th of February 2011, replaced the START I treaty, and superseded the Strategic Offensive Reductions Treaty (SORT). The verification system of New START was simpler and more cost-effective to operate than that which was prepared for the START treaty. On the 3rd of February 2021, it was agreed upon to extend the New START treaty by five years until the 5th of February 2026. Russia announced it was suspending the implementation of the New START treaty on the 21st of February of 2023.

The Strategic Arms Limitation Talks (SALT) were two rounds of bilateral conferences (each with a resulting treaty agreement) between the United States and the Soviet Union. The SALT I treaty was signed on the 26th of May 1972 and limited the number of strategic ballistic missile launchers (with provisions for both land and submarine based launchers). SALT I entered into force on the 3rd of October 1972. The Anti-Ballistic Missile treaty was a product of the SALT I talks. The SALT II treaty, signed on the 18th of June 1979, was a continuation of the SALT I talks and prohibited the advancement of missile programs, specifically those improving any key parameter by 5%. SALT II never entered into force and was instead superseded by START I in 1991.

The Anti-Ballistic Missile (ABM) Treaty, signed on the 26th of May 1972, placed limitations on the number of anti-ballistic missile systems used in defense against ballistic missile-delivered nuclear weapons. Each party was to be limited to 100 anti-ballistic missiles. The ABM treaty was intended to reduce pressure to build more nuclear weapons as a means to achieve deterrence. This bilateral agreement between the United States and the Soviet Union entered into force on the 3rd of October 1972. The United States withdrew from this agreement on the 13th of June 2002.

The Threshold Test Ban Treaty (TTBT), signed on the 3rd of July 1974, prohibits nuclear tests with devices having a yield exceeding 150 kilotons. This treaty addresses the exception in the PTBT for nuclear underground tests and is thus formally known as the Treaty on the Limitation of Underground Nuclear Weapon Tests. The TTBT only entered into force on the 11th of December 1990 and is a bilateral agreement between the United States and the Soviet Union/Russia.

The Peaceful Nuclear Explosions (PNE) Treaty, signed on the 28th of May 1976, is a bilateral agreement between the United States and the Soviet Union/Russia prohibiting peaceful nuclear explosions which were not addressed within the scope of the TTBT. Single explosions with yields exceeding 150 kilotons and group explosions having an aggregate yield of over 1500 kilotons (none of which may be more than 150 kilotons each) are the subject of the PNE treaty, which entered into force on the 11th of December 1990. Additionally, the PNE treaty addresses explosions at locations not designated as test sites under the TTBT.

Current Global Posture on Nuclear Disarmament and Typical Treaty Verification Techniques

According to the Federation of American Scientists (FAS),

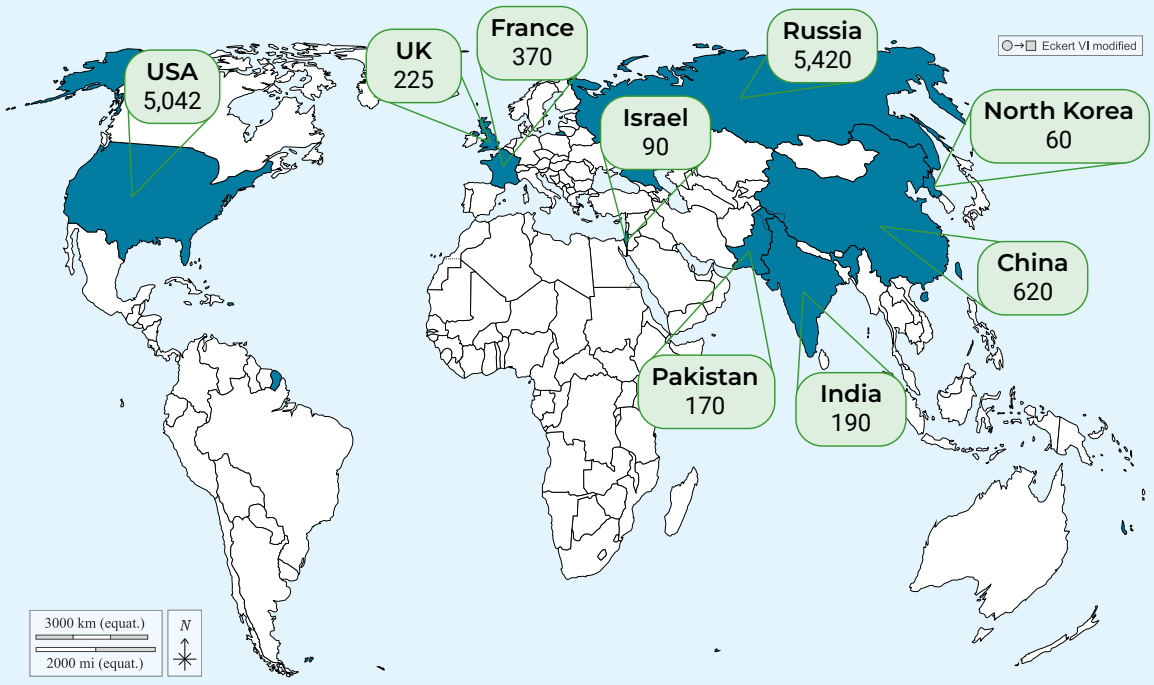


Figure 1: Estimated global nuclear warhead inventories current to January 2026 published by the Stockholm International Peace Research Institute. Map source: https://d-maps.com/carte.php?num_car=13181& (edited).

there was an estimated global stockpile of 70 300 nuclear weapons during the cold war. Figure 1 provides an estimate of the present stockpile of each nuclear armed state, compiled by the Stockholm International Peace Research Institute.

South Africa retired its nuclear weapons program in 1991 and is the only state to have voluntarily relinquished nuclear weapons (South Africa reported having had six weapons at the time). The nuclear weapons of Belarus, Kazakhstan, and Ukraine were transferred to Russia under the stipulations of the Lisbon Protocol.

The IAEA has been designated as the international safeguards inspectorate under Article III of the NPT. The IAEA Safeguards provide a mechanism of deterrence against the spread of nuclear weapons. The early detection of the misuse of nuclear material (or nuclear technology) grants credible assurance that nuclear material (or nuclear technology) is not being diverted from peaceful purposes. Part of the IAEA safeguarding process is a mechanism known as nuclear accountancy wherein, the quantity of nuclear material present in a nuclear facility, along with the variance in these quantities over time, is accounted for. The IAEA considers all "source and special fissionable material"; these are those significant to the manufacture of nuclear weapons. To ensure the containment of such materials, the IAEA has installed more than 23 000 seals at nuclear facilities. Some examples of such materials include plutonium-239, uranium-233 and uranium-235 (as well as any material containing a combination of these).

Also included in the safeguarding process is the evaluation of a state's nuclear activities by means of on-site inspections. On-site inspections of a nuclear facility, by the IAEA, allow for non-intrusive verification of nuclear materials by means of Non-Destructive Assay (NDA) methods. These methods, which detect neutron and gamma-ray signatures, survey the plutonium and uranium content of nuclear materials without opening their containers. In-field observations may also include item counting and weighing, inspection of design information of nuclear facilities as well as environmental sampling. Information on a state's nuclear activities may also be collected remotely using satellite imagery. Indeed, the IAEA surveillance system (1 400 surveillance cameras, 400 radiation and other sensors), captures in excess of a million pieces of encrypted safeguards data.

Instruments Supporting Disarmament and Non-Proliferation of Nuclear Weapons

The Joint Comprehensive Plan of Action

On the 14th of July 2015, the Joint Comprehensive Plan of Action (JCPOA) agreement was reached between Iran and the five permanent members of the United Nations Security council. This agreement, also known as the Iran nuclear deal, places limitations over the Iranian nuclear program in return for sanctions relief and other provisions. Specifically, Iran agreed to avail its nuclear facilities to increasingly extensive international inspections and to quell its nuclear program appreciably. The JCPOA currently has seven signatories namely: China, France, Germany, Iran, Russia, United Kingdom and the European Union. The United States of America withdrew on the 8th of May 2018. Iran announced on the 5th of January 2020 that it would no longer restrict its uranium enrichment by the limits set out within the JCPOA agreement.

The JCPOA is effectively no longer operational. Following Iran's continued non-compliance with its nuclear commitments, France, Germany and the United Kingdom triggered the United Nations "snapback" mechanism in August 2025, leading to the reimposition of United Nations sanctions in September. This followed the June 2025 Israel–Iran war, during which Israel and the United States struck several Iranian nuclear facilities. The confrontation has subsequently escalated into a broader United States–Iran conflict in 2026, further disrupting nuclear diplomacy and IAEA monitoring in the region.

Zangger Committee

After the Nuclear Non-Proliferation Treaty entered into force on the 5th of March 1970, the Zangger Committee was formed. The Zangger Committee, also known as the Nuclear Exporters Committee, has 39 members including all the nuclear-weapon states (those specific to the NPT). The objective of the committee meetings was to gain greater collective clarity on Article III, paragraph 2 of the Nuclear Non-Proliferation Treaty. The Zangger Committee maintains a list of nuclear strategic goods (and equipment) which are subject to export controls. This "Trigger List" is intended to assist signatories of the Nuclear Non-Proliferation Treaty in ensuring that recipients of strategic goods and equipment incorporate the obligatory safeguards.

Nuclear Suppliers Group

The Nuclear Suppliers Group (NSG), founded in response to India's first nuclear weapon test of May 1974, is a group of countries that seeks to contribute to the non-proliferation of nuclear weapons through the implementation of two sets of guidelines for nuclear exports and nuclear-related exports. The aim of the NSG Guidelines is to ensure that nuclear trade for peaceful purposes does not contribute to the proliferation of nuclear weapons or other nuclear explosive devices, and that international trade and cooperation in the nuclear field is not hindered unjustly in the process.

Convention on the Physical Protection of Nuclear Material and Nuclear Facilities

The Convention on the Physical Protection of Nuclear Material and Nuclear Facilities, signed on the 26th of October 1979, seeks to ensure the effective physical protection during the use, storage or transport of nuclear material used for peaceful purposes, as well as preventing and fighting crime associated with this material and related facilities. Each party to the convention must draw up and implement measures to guarantee effective protection against the theft or disappearance of nuclear material which it is responsible for and to prevent the sabotage of nuclear facilities on its territory.

Missile Technology Control Regime

Formed in April 1987, the voluntary Missile Technology Control Regime (MTCR) is a multilateral export control regime which aims to curb the spread of unmanned delivery systems that could be used for chemical, biological, or nuclear attacks. The 35 members are encouraged to restrict their exports of technologies capable of carrying a 500-kilogram payload at least 300 kilometers, or of delivering any type of weapon of mass destruction. Additionally, members are responsible for the establishment of national export control policies for ballistic missiles, cruise missiles, unmanned aerial vehicles, space launch vehicles, drones, remotely piloted vehicles, sounding rockets, and underlying components and technologies that appear on the regime's Material and Technology Annex.

Ballistic Missile Launch Notification Agreement

Signed at the Moscow summit on the 31st of May 1988, the agreement asserts that any launch of an intercontinental ballistic missile (ICBM) or a submarine-launched ballistic missile (SLBM) should be declared by Nuclear Risk Reduction Centers, no less than 24 hours in advance of the planned launch date. The launch location, as well as the area of impact, should be declared in the notification. This agreement specifically addresses concerns surrounding the risk of the outbreak of nuclear war as a consequence of misinterpretation or miscalculation.

Lisbon Protocol

The Lisbon Protocol to the 1991 Strategic Arms Reduction Treaty, was signed on May 23rd, 1992, by representatives of Russia, Belarus, Ukraine, and Kazakhstan. Article V of the protocol stipulates signatories, as non-nuclear-weapon states, shall adhere expeditiously to the NPT. It should be emphasized that article V is not applicable to Russia.

The Wassenaar Arrangement

On the 12th of July 1996, the Wassenaar Arrangement on Export Controls for Conventional Arms and Dual-Use Goods and Technologies was established to contribute to regional and international security and stability.

Through the promotion of transparency and responsible conduct in the export of conventional arms and dual-use goods (material, software, and technology), it aims to prevent the accumulation of destabilizing weapons. The prevention of the acquisition of these goods by terrorists is also an objective under this agreement. Member states are encouraged to prevent exports of these goods (through their national policies) that may augment the development of military capabilities by those seeking to thwart the objectives of the agreement.

Hague Code of Conduct Against Ballistic Missile Proliferation

On the 25th of November 2002, the Hague Code of Conduct Against Ballistic Missile Proliferation (formerly known as the International Code of Conduct Against Ballistic Missile Proliferation), was established to regulate access to ballistic missiles which can potentially deliver weapons of mass destruction. It is the only normative instrument to verify the spread of ballistic missiles. Active participation requires the exchange of information regarding ballistic missile and space launch vehicle programs, carried out annually between members. Under the Code, provisions should be made to deliver advanced notice of any launches of ballistic missiles or space launch vehicles. The Hague Code of Conduct has 145 member states (including MTCR members). Brazil has yet to adopt membership after articulating reservations surrounding possible limitations of its space program.

United Nations Security Council Resolution 1540 and 2663

Resolution 1540 of the United Nations Security Council, signed on the 28th of April 2004, stipulates that all states shall refrain from providing any form of support to non-state actors that attempt to develop, acquire, manufacture, possess, transport, transfer, or use nuclear, chemical or biological weapons and their means of delivery, in particular for terrorist purposes. All states are required to adopt and enforce appropriate laws and preventative measures to this end. Resolution 2663, adopted on the 30th of November 2022, extends the mandate of the subsidiary 1540 Committee for a period of ten years until the 30th of November 2032.

Sarah McKee

PRIF – Peace Research Institute Frankfurt

Lukas Rademacher

Technical University of Darmstadt/
PRIF – Peace Research Institute Frankfurt

International Treaties and Instruments Related to the Governance of Biological and Chemical Weapons Disarmament, Non-Proliferation, and Dual-Use Risks

International Treaties

The 1925 Geneva Protocol

The "Protocol for the Prohibition of the Use in War of Asphyxiating, Poisonous or Other Gases, and of Bacteriological Methods of Warfare", usually referred to as the 1925 Geneva Protocol, prohibits the use of chemical and biological weapons in warfare between its high contracting parties. As of September 2026, the Geneva Protocol had 146 member states. Upon signature, many parties declared that they would cease to consider themselves bound by the Protocol vis-à-vis states that did not respect its provisions, thereby de facto reserving the right to retaliate in kind. However, the prohibition of CBW use has evolved into a rule of international customary law, which means it applies universally to all states in international as well as intra-state conflicts.[1]

Germany signed the Protocol on June 17, 1925 and ratified it on April 25, 1929. It never attached any reservations to its membership.

Biological and Toxin Weapons Convention (BWC, BTWC)

The "Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction" (Biological Weapons Convention, BWC; also: Biological and Toxin Weapons Convention, BTWC) is an international disarmament treaty and the oldest treaty to prohibit an entire category of weapons.[2] It was opened for signature in 1972 and entered into force in 1975. As of September 2026, it had 189 states parties. Four states have signed but not ratified the BWC, while four states have neither signed nor ratified it.[3] The BWC unequivocally prohibits the development, production, possession, acquisition, retention, and transfer of biological weapons. It also obligates its members to prevent the proliferation of biological weapons materials, to foster international cooperation in the peaceful use of biology and biotechnology, and to implement the treaty provisions through national laws and regulations. Research is not covered by the BWC. This was presumably decided because the inherent dual-use character of much of biology and biotechnology, as well as the close interrelation with public health issues, make it impossible to prohibit or generically restrict specific types of research, as this could also restrict legitimate and useful research activities.[4]

Biological weapons are defined as microbial or other biological agents or toxins "of types and in quantities that have no justification for prophylactic, protective or other peaceful purposes" together with the means of delivery for such agents or toxins "for hostile purposes or in armed conflict" (BWC Article I). The prohibition is thus comprehensive and not bound to specific biological agents or types of delivery systems but is based on the intent to use an agent for biological weapon purposes (General Purpose Criterion). This means that all new relevant scientific and technological developments are automatically covered by the BWC.

The BWC does not provide for any measures to verify compliance with its provisions. Efforts to add a verification system failed in 2001, and the topic only returned to the official BWC agenda recently. At the 9th BWC Review Conference in 2022, states parties agreed to establish a Working Group open to all states parties whose mandate requires that they "identify, examine and develop specific and effective measures, including legally-binding measures, and to make recommendations to strengthen and institutionalize the Convention in all its aspects, to be submitted to States Parties for consideration and any further action".[5] In addition to verification and compliance, the topics covered by the Working Group include: international cooperation and assistance; national implementation; assistance, response and preparedness; organizational, institutional and financial matters; confidence-building and transparency; and scientific and technological (S&T) developments. The Working Group is expected to submit its results and recommendations to the next BWC Review Conference in 2027, at the latest.

New developments in biology and biotechnology have provided opportunities and challenges to the BWC and to international efforts to strengthen it. Some advances in these and related scientific disciplines – such as artificial intelligence – may be used to enhance preparedness or to facilitate verification of compliance with the BWC. However, some materials and technologies might also be misused for illegitimate and hostile purposes, or lower the barriers for biological weapons development and acquisition. The negotiators of the BWC already recognized the significance of scientific advances by including the review of relevant S&T developments as a task for the BWC review conference in Article XII of the Convention. While this task has never been carried out systematically and collectively on a continuous basis,

the BWC Working Group is now mandated to make recommendations to states parties to establish "a mechanism to review and assess scientific and technological developments relevant to the Convention and to provide States Parties with relevant advice".[6] If established, such a mechanism could screen, review and assess S&T developments and provide guidance to states parties on how to recognize and mitigate potential risks. This might include risks arising from dual-use research (of concern), or may help identify ways to exploit potential benefits for strengthening preparedness for disease outbreaks – regardless of their origin – or even for the BWC as a whole. Negotiations on an S&T advisory mechanism were well-advanced in mid-2026 but are still complicated due to geopolitical tensions.

While the BWC itself does not contain any reference to bioterrorism, biosafety, biosecurity, or health preparedness, these topics have entered the BWC discourse over the past 20 years and are now integral topics of discussion. Not all related definitions and understandings enjoy consensus among BWC members. The discursive shift has, however, put the BWC at the center of a larger set of measures aimed at mitigating biological risks, with the spectrum of such risks comprising natural disease outbreaks, unintentional/accidental release of biological agents, deliberate use of biological agents for criminal or terrorist purposes, and the use of biological weapons by states.

Germany signed the BWC on April 10, 1972, and ratified it on April 7, 1983.

Chemical Weapons Convention (CWC)

The "Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on their Destruction" (Chemical Weapons Convention, CWC) was opened for signature in 1993 and entered into force in 1997. As of September 2026, it had 193 states parties. Israel has signed but not ratified the CWC, while Egypt, North Korea and South Sudan have neither signed nor ratified the treaty. The CWC contains a comprehensive prohibition of chemical weapons, including their use. It also contains obligations for states parties to prevent chemical weapons proliferation, to foster international cooperation in the peaceful use of chemistry, to provide assistance in case of a chemical attack, and to implement the treaty provisions nationally. Unlike the BWC, the CWC establishes its own treaty organization, the Organisation for the Prohibition of Chemical Weapons (OPCW), which oversees the implementation of the Convention.[7]

Chemical weapons are defined in CWC Article I as "toxic chemicals and their precursors, except where intended for purposes not prohibited under this Convention, as long as the types and quantities are consistent with such purposes", together with munitions and other devices and equipment intended for their dissemination. Toxic chemicals have a number of legitimate applications, which is why the CWC makes the prohibition contingent on the purposes for which they are used (General Purpose Criterion) and defines exceptions for legitimate applications. Purposes not prohibited include applications in industry, agriculture, medicine and pharmacology, as well as protection against chemicals and other applications for defensive purposes. Moreover, the treaty prohibition only applies when the use of the chemical in question depends on its toxic properties (as opposed to, for instance, explosive, incendiary, or corrosive properties). Finally, some chemicals, such as tear gases, may be used for law enforcement and domestic riot control, but not as a method of warfare.

The CWC establishes an elaborate verification system that is implemented by the OPCW and its Technical Secretariat. As part of this system, routine industry inspections are based on three lists ('schedules') that contain chemicals that have been developed or produced as chemical warfare agents or which, due to their toxicity, other properties, or their relevance for the production of chemicals listed on schedules 1 and 2, might "pose a risk to the object and purpose" of the CWC. Besides the level of risk posed by the respective substances, assignment to one of the schedules also depends on how often the chemicals in question are used for commercial purposes. Schedule 1 chemicals, which have been used developed, produced, or used as chemical weapons in the past, have little or no commercial use. Schedule 2 chemicals pose a "significant risk" and are "not produced in large commercial quantities". Schedule 3 chemicals have also been produced or used as chemical weapons but also have widespread legitimate applications and are hence subject to less stringent controls. States parties are obliged to declare scheduled chemicals as well as specific chemical facilities which are then randomly selected for industry inspections by the OPCW to verify that the chemicals and facilities are used for declared and permitted purposes only. It should be noted that the schedules are used for verification purposes and do not define prohibited substances or chemical weapons. The use of any toxic chemical, listed or not, as a chemical warfare agent is unequivocally prohibited. Chlorine is a case in point: it is not listed on any schedule given its widespread civilian applications, but its use as chemical weapon – e.g. in the Syrian civil war – was as illegal as the use of the nerve agent sarin which is listed on schedule 1.

In addition to the routine inspections, in cases where a state party harbors concrete suspicions of CWC violations by another state party, the CWC provides for chal-

lenge inspections or investigations of alleged chemical weapons use. Neither mechanism has been invoked to date. Instead, when chemical weapons attacks were reported from Syria starting in 2012, the UN and the OPCW set up various ad hoc mechanisms to investigate the allegations, identify the perpetrators and clarify other concerns about Syria's compliance with the CWC. As a result, investigation and identification technologies, along with chemical forensics, have become increasingly necessary capacities for the OPCW. In the past, another core task of the OPCW was the verification of the destruction of declared chemical weapons stockpiles and facilities. However, because destruction of all such stockpiles and facilities was completed in 2023, the OPCW has officially entered the post-destruction era and is now more strongly focused on preventing the re-emergence of chemical weapons.

Recognizing the importance of scientific and technological developments in chemistry for the CWC in general, the negotiators of the CWC provided for the establishment of a Scientific Advisory Board (SAB), to monitor relevant developments and provide advice to the director-general and member states of the OPCW (CWC Article VIII, 21.h). The SAB comprises 25 members who are independent experts, nominated by states parties from different regions to ensure geographical balance. The SAB meets annually and submits its recommendations to the director-general and to states parties. Past recommendations have included the establishment of temporary working groups on specific topics, such as on biotoxins (2022-2024), or chemical forensics (since 2025) and Artificial Intelligence (since 2026). The OPCW has its own laboratory, housed in the OPCW Centre for Chemistry and Technology (ChemTech Centre). The laboratory is not only used to carry out the OPCW's verification and investigation functions but also to offer opportunities for exchange, training, and capacity-building among scientists. Chemical terrorism is not explicitly mentioned in the text of the Convention, but there is a recognition that the OPCW can play a role in its prevention, including through capacity-building and training in chemical security, and that effective national implementation of the CWC contributes to the prevention of chemical terrorism.[8] Some developments in chemistry and other disciplines, most prominently AI, afford opportunities to enhance verification and build capacity in chemical safety, security, and defense. However, the same developments may also harbor the potential for misuse that must be recognized, monitored, and mitigated.

Germany signed the CWC on January 13, 1993, and ratified it on August 12, 1994. It currently has a seat on the OPCW Executive Council.

Other Instruments Supporting Biological and Chemical Weapons Disarmament and Non-Proliferation

UN Secretary-General's Mechanism (UNSGM)

In 1987, the UN General Assembly authorized the UN Secretary-General through UNGA Resolution 42/37C to investigate any case of alleged biological, chemical or toxin weapons use which UN members bring to their attention. It thus established the UN Secretary-General's Mechanism for Investigation of Alleged Chemical or Biological Weapons Use (UNSGM). The UNSGM has its legal basis in the 1925 Geneva Protocol and is the only international mechanism currently available to investigate the alleged use of biological weapons.[9] While the UNSGM also covers alleged chemical weapons use, in practice – and given the CWC's near-universal membership – such cases would very likely be dealt with by the OPCW.

If a UN member requests a UNSGM investigation, and if the UN Secretary General determines that sufficient evidence is presented to warrant such an investigation, he or she can dispatch a mission team on short notice, drawing on a roster of experts and laboratories nominated by UN member states to the UN Office for Disarmament Affairs (UNODA). The mandate of such a mission would be to "ascertain the facts of the matter" and establish whether chemical, biological or toxin weapons have indeed been used. The mechanism is independent politically and placed solely under the authority of the UN Secretary General. The UN Security Council (UNSC) has no formal role in the investigation.

Germany is a member of the "Group of Friends of the UNSGM" and one of the most active supporters of efforts to strengthen the mechanism.

UN Security Council Resolution 1540 (2004)

In 2004, the UN Security Council (UNSC) adopted Resolution 1540 which obligates all UN members to take the necessary measures to prevent the proliferation of nuclear, biological, or chemical (NBC) weapons and their means of delivery to non-state actors, as well as to report on their implementation of the resolution. The resolution has since been renewed several times, most recently in 2022, when it was extended for a period of ten years (UNSC Resolution 2663). Measures to be taken by UN members include the adoption and enforcement of national legislation to prevent non-state actors from developing or acquiring NBC weapons, domestic physical protection measures for relevant NBC materials, and effective border and transfer controls. A dedicated committee, the 1540 Committee, was established to support the implementation of Resolution 1540, including by managing the national reports and facilitating the provi-

sion of assistance to UN members in implementing the resolution.

Germany has implemented Resolution 1540 and submitted five reports, most recently in 2020.

Sanctions and Export Controls

To prevent the proliferation to states or non-state actors of materials, equipment, or technologies that could be used to develop and produce biological or chemical weapons, states can put in place several national and international measures, such as sanctions, as well as national and internationally harmonized export controls.

Under Chapter VII of the UN Charter, the UN Security Council (UNSC) can enact sanctions against actors posing a threat to international peace and security. The UNSC has repeatedly determined that the proliferation of nuclear, chemical and biological weapons poses such a threat. Currently there is one UNSC sanctions regime in place that also relates to biological and chemical weapons, namely, the sanctions regime against North Korea. Implementation of these sanctions is obligatory for all UN member states. In addition, unilateral biological and chemical weapons-related sanctions are in place against other states as well. For example, the USA has enacted sanctions against Russia, Iran, North Korea, and Syria, while the EU has set up a sanctions regime in relation to the use and proliferation of chemical weapons.

Export controls are a crucial measure for implementing the non-proliferation obligations that arise from the BWC, CWC, and UNSC Resolution 1540. The CWC establishes specific trade restrictions for the chemicals listed on the three schedules. Chemicals listed on schedules 1 and 2 may only be transferred to other states parties. For schedule 1 chemicals, the exporting state has to notify the OPCW of each transfer and must declare all transfers of schedule 1 chemicals annually. Schedule 3 chemicals may be transferred to non-state parties. In the latter case, however, end-user certificates are required. There is no global system of biological weapons-related export controls in place. Rather, the control of relevant transfers is the responsibility of individual states. However, the Australia Group and the EU export control regime for dual-use items represent two examples of multilaterally coordinated sets of export control measures.

The Australia Group (AG) is an informal arrangement in which 42 countries and the EU collaborate on the basis of shared control lists to harmonize their export controls over agents, substances, materials, equipment and technologies that could be used to manufacture biological or chemical weapons. The AG was founded in 1985 after it had become known that, with the help of legal chemical exports, Iraq had produced chemical weapons including those used against Iran in the Iran-Iraq war in the 1980s. Biological weapons-related items were added to the group's control lists in 1991. They currently comprise lists for "chemical weapons precursors", "dual-use chemical manufacturing facilities and equipment and related technology and software", "dual-use biological equipment and related technology and software", "human and animal pathogens and toxins", and "plant pathogens". Membership in the AG is possible in principle for countries fulfilling a set of criteria listed on the AG's website, but is contingent on unanimous agreement by all current AG members. There has, however, been some controversy around the AG. While its members justify its existence with reference to their non-proliferation obligations under the BWC, CWC and Resolution 1540, other states have criticized the group of being discriminatory and of hampering economic and technological development by restricting legitimate transfers.

In 2021 the EU set up a control regime for the transfer of dual-use items. Last amended in 2026, regulation (EU) 2021/821 also includes chemical and biological weapons-related items. Additionally, it explicitly refers to the BWC, the CWC and UNSC Resolution 1540 and obligates all EU members to enact transfer controls on items that are deemed to have dual-use and that are specified in Annex I of the regulation. "Dual-use" is defined in chapter 1, article 2 of the regulation as meaning "items, including software and technology, which can be used for both civil and military purposes, and includes items which can be used for the design, development, production or use of nuclear, chemical or biological weapons or their means of delivery, including all items which can be used for both non-explosive uses and assisting in any way in the manufacture of nuclear weapons or other nuclear explosive devices".

As part of its obligations as a member of the UN and the EU, Germany is implementing the sanctions enacted by the UNSC and the sanctions and dual-use export regulations enacted by the EU. Moreover, Germany is a founding member of the Australia Group. The agency responsible for implementing Germany's export control policy is the Federal Office for Economic Affairs and Export Control (BAFA).

Selected Instruments to Enhance Chemical and Biological Security, Mitigate Dual-use Risks, and Promote Safe, Secure, and Ethical Research in Chemistry and Biology

WHO Global Guidance Framework for the Responsible Use of the Life Sciences

The World Health Organization (WHO) published its Global Guidance Framework for the Responsible Use of the Life Sciences in 2022. It is a comprehensive and

practice-oriented reference document aimed at aiding states in their development of national biorisk management strategies and focusing on biosafety, laboratory biosecurity and the oversight of dual-use research. It provides an assessment of the current situation and offers some guiding values and principles for biorisk governance, as well as concrete tools and mechanisms, tailored to specific types of stakeholders. In its annex, it also offers seven hypothetical scenarios in which elements from the Global Guidance Framework could be applied, as well as three illustrative case studies of research that required risk mitigation and management strategies. The Global Guidance Framework is intended as a supporting resource for states; its implementation is not mandatory.

The Hague Ethical Guidelines

The Hague Ethical Guidelines were developed in 2015 by a group of practitioners from the chemical field to provide elements for codes of conduct and ethical discussions related to the use and application of chemistry in accordance with the CWC. The principles on which the Guidelines are based refer to sustainability, education, awareness and engagement, ethics, safety and security, accountability, oversight and exchange of information. They postulate the responsibility of practitioners in chemistry to ensure that chemistry is used for peaceful purposes, to raise awareness of the topic of chemical weapons and the misuse potential of chemistry, and to emphasize the need to include ethical as well as safety and security considerations when applying chemistry. The Guidelines are not binding, but they have been endorsed by several international chemical associations and serve as widely accepted guiding principles for "good" behavior in chemistry.

Tianjin Biosecurity Guidelines for Codes of Conduct for Scientists

The Tianjin Guidelines for Codes of Conduct for Scientists were developed jointly in 2021 by representatives of Tianjin University in China, the Johns Hopkins Center for Health Security in the USA and the InterAcademy Partnership (IAP) and are intended as a contribution to preventing the misuse of biological sciences, including for weapons purposes, through raising awareness of the dual-use and misuse potential of biological research. They contain principles for the responsible use of biosciences that could be incorporated in national biosecurity regulations and codes of conduct. This initiative goes back to a proposal for a model voluntary code of conduct for biological scientists which was first presented in 2016 by China and Pakistan.[10] At the BWC Review Conference in 2022, China, Pakistan and Brazil as co-sponsor submitted the Tianjin Guidelines to BWC states parties for endorsement.[11] While the Guidelines could not be formally endorsed by the Review Conference, they enjoyed very wide support among BWC members.

The Framework for Handing Security-Relevant Research in Germany

In recognition of the basic right to the freedom of research as codified in Article 5 of the German Basic Law (constitution), governance of security-relevant research in Germany follows a bottom-up approach and emphasizes the responsibility of individual researchers and research institutions.

Following publication of controversial so-called gain-of-function experiments with a bird flu virus (H5N1) in 2012, the German Ethics Council as well as the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) and the National Academy of Sciences Leopoldina developed recommendations on the handling of dual-use research in the life sciences. Both documents emphasized the freedom of research and the need to raise awareness among scientists of potential dual-use risks associated with their work. Whereas the Ethics Council also suggested a top-down regulatory approach which was not implemented, the DFG and Leopoldina focused on the self-governance of science and the responsibility of individual researchers and research institutions in the handling of security-relevant research.

The latter approach is being actively promoted by the DFG and Leopoldina. Among other things, it encourages research institutions to establish local Committees for Ethics of Security-Relevant Research (Kommissionen für Ethik sicherheitsrelevanter Forschung, KEF) to provide guidance and advice to researchers who have identified potential security risks in their research. Through the Joint Committee on the Handling of Security-Relevant Research, established in 2014 to support the implementation of their recommendations, the DFG and Leopoldina also undertake outreach and awareness-raising efforts in the scientific community in Germany.

Una Jakob
PRIF – Peace Research Institute Frankfurt

1 Henckaerts, J.-M., & Doswald-Beck, L. (2005). *Customary international humanitarian law* (Vol. 1: Rules). Cambridge University Press. https://doi.org/10.1017/CBO9780511804700

2 For an analysis of the BWC regime see Jakob, U. (2022). The Biological Weapons Convention. In T. Marauhn & E. Myjer (Eds.), *Research handbook on arms control law* (pp. 258–277). Edward Elgar. https://doi.org/10.4337/9781788111904.00031

3 Non-ratifying signatories are Egypt, Haiti, Somalia and Syr-

ia; the non-members are Chad, Djibouti, Eritrea, and Israel.

4 'Dual-use' here denotes the potential for legitimate biological research, technology, materials or equipment to be misused for hostile purposes, including as biological weapons. Dual-use research includes experiments that, while designed for benign purposes, might yield results that could be misused; and duals-use research of concern (DURC) generally refers to research where this misuse potential is particularly great and immediate.

5 Biological Weapons Convention. (2022, December 22). *Final document of the Ninth Review Conference* (BWC/CONF.IX/9, p. 9). United Nations. https://digitallibrary.un.org/record/4006755?v=pdf

6 Biological Weapons Convention (2022), p. 11.

7 For an analysis of the CWC regime see Trapp, R. (2022). The Chemical Weapons Convention. In T. Marauhn & E. Myjer (Eds.), *Research handbook on arms control law* (pp. 279–296). Edward Elgar. https://doi.org/10.4337/9781788111904.00032

8 On chemical terrorism as a topic in the CWC framework see Kelle, A. (2024). *Chemical terrorism in context: Threat perceptions and CWC Review Conference responses* (CBWNet Working Paper No. 13). CBWNet. https://cbwnet.org/publications/working-paper

9 For information about the UNSGM see Jakob, U., Kloth, S., & Mergler, I. (2024). *Investigation of alleged use of biological weapons – The UN Secretary-General's Mechanism* (PRIF Report 07/2024). Peace Research Institute Frankfurt. https://www.prif.org/fileadmin/Daten/Publikationen/Prif_Reports/2024/prif0724_barrierefrei.pdf

10 China & Pakistan. (2022, April 7). *The Tianjin Biosecurity Guidelines for Codes of Conduct for Scientists* (BWC/CONF.IX/PC/WP.10). United Nations. https://docs.un.org/BWC/CONF.IX/PC/WP.10

11 China & Pakistan. (2016, November 15). *Proposal for the development of a model code of conduct for biological scientists under the Biological Weapons Convention* (BWC/CONF.VIII/WP.30). United Nations. https://docs.un.org/BWC/CONF.VIII/WP.30

The Open Skies Treaty

The Open Skies Treaty was signed in 1992 and entered into force in 2002, with 27 of its 34 participating states having ratified[1] the treaty. The United States withdrew in November 2020, closely followed by Russia in June 2021. The treaty is a political agreement designed to enhance mutual understanding and confidence amongst state parties through unarmed aerial surveillance flights over the entire territory of each member, which are used to verify military movements and arms control treaties. Each state party is permitted to conduct short-notice and unarmed reconnaissance flights, and additional state parties are permitted to accompany and observe the mission. The sensors on board the flights are certified according to treaty stipulations prior to inclusion on the aircraft and a copy of the data collected by these sensors is supplied to the host country. All state parties receive a mission report and may purchase the data collected. This action of jointly collecting data, so that authenticity thereof is not contested, is an exceptional strength[2] of the treaty. The treaty upholds the following limits[3] on the instrumentation and data collected:

- Optical and panoramic framing cameras: Ground resolution of no better than 30 cm.
- Video cameras with real-time display: Ground resolution of no better than 30 cm.
- Infrared line-scanning devices: Ground resolution of no better than 50 cm.
- Sideways looking synthetic aperture radar: Ground resolution of no better than 3m on a single radar unit (object separation capability of 5 m). There may not be 2 radar units observing from both sides of the aircraft simultaneously.

Sarah McKee

PRIF – Peace Research Institute Frankfurt

1 Arms Control Association. (2021, December). *The Open Skies Treaty at a glance.* https://www.armscontrol.org/factsheets/openskies

2 Richter, W. (2020, June 15). *Attack on the Open Skies Treaty: President Trump wants to withdraw from the Open Skies Treaty* (SWP Comment 2020/C 29). Stiftung Wissenschaft und Politik. https://doi.org/10.18449/2020C29

3 Graef, A., & Kütt, M. (2020, April 16). *Visualizing the Open Skies Treaty.* https://www.openskies.flights/archived-2020-04-16/index.html

International Initiatives and Core Principles for the Safe Use of (Military) AI: New Developments in 2025/2026

The most significant development in 2025–2026 was the move from general principles toward establishing implementation machinery. This includes capability-based thresholds, prerelease evaluations, testing and assurance regimes, procurement requirements, life-cycle risk management, human-control parameters, incident reporting, institutional governance, and crisis-risk reduction measures. The selected documents below reflect this shift and illustrate how the relevant arguments and mechanisms have evolved.

United States: A New Approach to Frontier AI Risks

Executive Order 14179, *Removing Barriers to American Leadership in Artificial Intelligence*,[1] and the July 2025 document *America's AI Action Plan*[2] marked a shift in overall U.S. policy toward the rapid development and adoption of AI. They direct agencies to identify and remove any federal measures regarded as barriers to AI innovation and to revise the previous administration's implementation framework. This is not, however, tantamount to abandoning national security controls for the most capable models.

Executive Order 14409, *Promoting Advanced Artificial Intelligence Innovation and Security*,[3] creates a more targeted security track. It directs the government to establish a classified benchmark process for advanced cyber capabilities and to use its results to designate "covered frontier models." Developers may voluntarily provide the government with access to a covered model for up to 30 days before making it available to other trusted partners, thereby allowing prerelease testing while protecting confidential and proprietary information. The order expressly rejects a general mandatory licensing or government preclearance regime.

The order also proposes concrete cyber defense infrastructure around frontier models. It calls for an AI cybersecurity clearinghouse involving government, industry, and critical infrastructure actors to coordinate vulnerability discovery, validate findings, and prioritize remediation; expands access to AI-enabled defensive tools; and directs law enforcement agencies to prioritize cases of unlawful computer access or damage involving AI agents. The policy logic is therefore selective: broad regulatory restraint for AI generally, paired with capability-based intervention where frontier models materially affect national security cyber risk.

National Security Presidential Memorandum 11, *Artificial Intelligence in the National Security Enterprise*,[4] is more operational than earlier statements of responsible use principles. It organizes national security AI policy around four pillars: adoption, adaptation, assurance, and accountability. Agencies are directed to use advanced commercial and open-source models where appropriate, adapt them to mission needs, and ensure that mission-dependent systems remain reliable, robust, steerable, and controllable. The implementation measures that follow from the memorandum are substantial. Among others, it orders an update of DoD Directive 3000.09, *Autonomy in Weapon Systems*,[5] followed by recurring review; standardized national security testing, evaluation, verification and validation methods; a joint AI risk management and assurance strategy; procurement pathways able to onboard advanced models from multiple suppliers; secure computing infrastructure and a national security AI test range; and public-private partnerships for threat intelligence and AI red-teaming. It also creates mechanisms for incorporating nongovernmental AI expertise into national security work. The central policy proposition is thus accelerated fielding combined with a dedicated assurance, procurement, and security architecture.

European Union: Warning on Systemic Risks from Frontier AI

The European Systemic Risk Board's June 2026 *Warning on systemic cyber risks stemming from frontier artificial intelligence models*[6] makes a specific argument regarding the security consequences of capability gains. It considers frontier models capable of automating vulnerability discovery, exploit creation, and the execution of full cyberattacks at speed and scale. A central concern is the erosion of defensive time buffers: exploit development that previously took days or weeks may be reduced to minutes or hours, while a much larger volume of vulnerabilities could overwhelm patching processes and force institutions to choose between remaining exposed and deploying insufficiently tested patches. The EU document identifies four pressure points: time; defender capability; concentration/dependency on a small number of AI, cloud, and software providers; and the ability of authorities to calibrate testing and preparedness. It also highlights asymmetries between jurisdictions, attackers and defenders, and better- and less well-resourced institutions. It calls for adequate and proportionate European access to new frontier models, coordinated testing and cooperation, updated cybersecurity and stress-testing frameworks, and clear board-level governance and accountability.

Paris Declaration 2025: Maintaining Human Control in AI-Enabled Weapon Systems

The *Paris Declaration on Maintaining Human Control in AI enabled Weapon Systems*,[7] issued at the February 2025 AI Action Summit and endorsed by 27 states, is a concise but unusually specific state-endorsed declaration on the boundary between AI-enabled weapons and human responsibility. It affirms that responsibility and accountability cannot be transferred to machines and commits endorsers to safeguards that preserve human judgement and control over the use of force. Its strongest formulation is that endorsers will not authorize "life and death" decisions to be made by an autonomous weapon system operating completely outside human control and a responsible chain of command. It further states that international humanitarian law (IHL) applies fully to AI-enabled weapon systems and that endorsers will not develop, deploy, or use systems that cannot be employed consistently with IHL. Compared with broader responsible-AI declarations, the document narrows the debate to an operational red line: exclusion of weapons operating entirely outside human control and command responsibility.

REAIM 2026: From Principles to Pathways to Action

The *A Coruña Declaration – REAIM 2026 Pathways to Action*[8] deliberately shifts the Responsible Artificial Intelligence in the Military Domain (REAIM) process toward implementation. Its core proposition is "responsible by design": legal, ethical, technical, and operational safeguards should be built into the military AI life cycle rather than added only at the point of deployment. The declaration specifically highlights risks arising from unanticipated or emergent behavior and the limits of AI explainability, and calls for robust testing, evaluation, verification, and validation (TEVV) to be integrated into relevant policies, doctrines, and procurement processes. The document also translates responsible use into organizational measures. It promotes training and risk management, continued monitoring and documentation, exchanges of practices, and—where compatible with national security—visits or exchanges involving relevant facilities and centers of excellence. At the international level, it suggests crisis communication arrangements to reduce misunderstanding, miscalculation, and unintended escalation; capacity-building through regional centers and knowledge hubs; and points-of-contact networks. For industry and academia, it emphasizes research on interpretability and robustness and calls on developers to address human-machine interaction and interface design proactively. Currently, the official REAIM page lists 45 endorsing states; neither the United States nor China are among them. The document remains politically nonbinding, but it specifies how earlier principles are to be translated into doctrine, procurement, testing, training, and international risk reduction practices.

United Nations: A Dedicated Military AI Governance Track and AI in Nuclear Command, Control, and Communications

United Nations General Assembly Resolution 80/58, *Artificial intelligence in the military domain and its implications for international peace and security*,[9] is important because it structures a continuing UN process around military AI as a life-cycle governance problem. It covers predesign, design, development, evaluation, testing, procurement, deployment, use, sale, operation, and decommissioning; identifies risks including arms race dynamics, miscalculation, escalation, proliferation, bias, and use by non-state actors; and emphasizes safeguards involving human judgement and control over the use of force. The resolution also established a follow-up mechanism consisting of informal exchanges in Geneva on the Secretary-General's report, existing and emerging normative proposals, and possible next steps, followed by a factual summary for the First Committee. It also encouraged Secretariat knowledge-sharing, as well as voluntary assistance and capacity-building, particularly for developing countries.

A separate new resolution is directly relevant to high-consequence military AI. General Assembly Resolution 80/23, *Possible risks of the integration of artificial intelligence into command, control and communications systems of nuclear weapons*,[10] focuses on the possibility that AI could shorten decision windows, distort decision environments, or fail because of malfunction, intrusion, automation bias, or inaccurate outputs in nuclear command-and-control contexts. The resolution stipulates that human control and oversight be maintained over nuclear command, control, and communications systems, including those enabled by AI. It urges nuclear-weapon states to adopt and publish policies and doctrines that operationalize this requirement and explicitly provide that AI systems will not autonomously initiate decisions on the use of nuclear weapons. It also encourages voluntary information-sharing on implementation and the development of common understandings and confidence-building measures in relevant disarmament and treaty forums. This resolution is one of the clearest recent UN formulations of a specific human-control requirement for a defined military AI application.

CCW: Concrete Elements for an Instrument on Autonomous Weapons

The CCW Group of Governmental Experts has advanced materially beyond general exchanges of national positions. Its draft final report from September 2026[11]

records a consensus set of elements for a future instrument, without prejudging whether that instrument will ultimately be legally binding. It characterizes a lethal autonomous weapon system as a combination of weapons and technological components capable of identifying, selecting, and engaging a target with force without intervention by the human operating the system, while excluding systems in which personnel determine the specific target or target group and the timing of attack. Substantively, the text sets out proposed prohibitions and control measures. It states that lethal autonomous weapon systems (LAWS) that are inherently indiscriminate, otherwise incapable of IHL-compliant use, or whose effects cannot be anticipated and limited as required by IHL must not be used, and reiterates the prohibition on directing attacks against civilians. Human judgement and control are operationalized through a responsible chain of command, human legal assessment, restrictions on target types, scale, duration, and geographic scope, human authorization for material changes to mission or target parameters, and the timely deactivation or neutralization of systems. It also calls for legal reviews, testing in representative environments, training, reliability/predictability/traceability, mitigation of data and automation bias, and internal mechanisms for incidents that may involve IHL violations.

Germany: Bundeswehr AI Concept

The German Federal Ministry of Defense (BMVg) *KI-Konzeption: zukunftsfähige Verteidigung mit Künstlicher Intelligenz*[12] is intended as a binding internal framework for the development, introduction, testing, and use of AI across the BMVg's area of responsibility. Operationally, it presents AI as a means of enhancing information, command, and effect superiority; supporting multidomain operations and "mosaic warfare"; and improving staff and administrative processes. At the same time, it requires AI to be embedded within existing structures of organizational responsibility and data governance rather than treated as a standalone capability. The concept establishes a life cycle risk management and assurance regime. Testing with representative data is a prerequisite for authorization to use a system; high-risk systems are subject to "zero-failure" testing as a minimum requirement alongside quality metrics; material changes or retraining require renewed validation; and AI-specific minimum requirements are to be written into specifications, contracts, and acceptance procedures and subjected to independent review. Conformity review is required before first use and after material changes, while deployed systems are expected to support performance monitoring and, where relevant, revalidated in-mission updates. With regard to human control, personnel responsible for autonomous functions should understand system behavior and environmental interaction, evaluate reliability, set and validate rules of use, define mission objectives, target types, mission duration, and operational area, conduct Article 36 weapons reviews where applicable, and retain responsibility within a human chain of command. The concept establishes a coordinated governance structure around a Chief Artificial Intelligence Officer and CAIO Council, while retaining decentralized responsibility within existing processes.

NATO: Digital Architecture for Operational AI

The *Alliance Digital Strategy*,[13] adopted by NATO in January 2026, goes beyond merely restating responsible use principles by locating AI within the Alliance's future command-and-control and digital architecture. It calls for accelerated AI adoption in digital services; inference at the tactical edge to improve situational awareness, enable real-time risk-informed decisions, and increase operational autonomy; and AI-enabled sensor fusion, predictive analytics, and cognitive decision support for cross-domain command. These capabilities are linked to federated digital platforms, hybrid-cloud and tactical-edge infrastructure, standardized data and metadata, Zero Trust security, and software-defined defense. The strategy also sets out an implementation model: "adopt, buy, create" guides decisions on whether to reuse, procure, or develop digital capabilities; interoperability requirements and standards become particularly consequential when entities connect to federated networks for NATO-led operations; workforce development includes training on how to interact with AI-enabled systems; and progress is to be measured using the Digital Transformation Maturity Model and associated indicators. The new element is therefore not a separate AI code, but the integration of AI into the technical, organizational, and measurement architecture through which Alliance capabilities are fielded.

China: More Specific Principles and Risk Governance for Military AI

China's *Opportunities and Challenges Posed to International Peace and Security by the Application of AI in the Military Domain*[14] submitted to the UN in April 2025 provides a more detailed position on military AI than its earlier, more general governance statements. It argues that states—particularly major powers—should avoid seeking absolute military advantage through AI, prevent miscalculation and an AI arms race, and ensure that humans retain ultimate responsibility. It also states that weapon systems should remain under human control and that military AI must comply with international law, including IHL. The document further contains concrete governance proposals that are easily overlooked if it is treated merely as a statement of principles. China calls

for forward-looking risk assessment; testing and evaluation based on the risk level of AI applications; tiered and categorized management; domestic legal, and regulatory systems and ethical guidelines; personnel education and training; and measures to reduce proliferation risks without obstructing peaceful innovation. Internationally, it emphasizes the UN as the principal forum and calls for policy coordination and capacity-building. The position therefore combines human-control language with a risk tiering and national implementation approach.

EU Artificial Intelligence Act: The Systemic Risk Regime Becomes Operational

For the *EU Artificial Intelligence Act*,[15] the main new development is the implementation of its general-purpose AI (GPAI) regime. Since August 2025, providers placing GPAI models on the EU market have been subject to documentation, downstream information, copyright policy, and training-content transparency requirements. Providers of models already on the market before that date are subject to a later compliance deadline, and certain open-source models are exempt from some documentation requirements. Models with "systemic risk" face additional obligations: notification, standardized model evaluation, and adversarial testing; assessment and mitigation of Union-level systemic risks; serious-incident tracking and reporting; and cybersecurity protection for the model and its physical infrastructure. The regulation presumes high-impact capability above the statutory training-compute threshold, while also allowing Commission designation based on capability and impact.

The European Commission's 2025 materials on *General-purpose AI obligations under the AI Act*,[16] taken together, make these requirements more operational. The code's "Safety and Security" chapter provides a voluntary compliance pathway for providers of systemic risk models, including safety and security frameworks, evaluation and risk management practices, model reporting, and post-market monitoring. Since August 2026, the Commission has been able enforce GPAI provider obligations,[17] including through fines. This is relevant to a military AI annex not because the AI Act directly regulates systems used exclusively for military, defense, or national security purposes—it does not—but because it is becoming a detailed public regulatory model for capability identification, adversarial evaluation, systemic risk management, incident reporting, and security of frontier models.

Council of Europe: Ratification and Operationalization of Risk Assessment

The European Union ratified the *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*[18] in May 2026. The convention has not yet entered into force. It provides that parties shall establish iterative, graduated risk and impact assessments for actual and potential effects on human rights, democracy, and the rule of law; prevention and mitigation measures; transparency and accountability safeguards; procedural rights and remedies for affected persons; and the possibility of bans or moratoria on uses considered incompatible with those protected interests.

A 2026 follow-up element in the Council of Europe's *HUDERIA – risk and impact assessment of AI systems*[19] process is the approval of the *HUDERIA Model: COBRA*. The measures proposed by the HUDERIA framework are nonbinding, but provide states with a practical structure for identifying, assessing, and mitigating AI risks across the life cycle, using factors such as severity, probability, and reversibility, and for revisiting assessments as systems or contexts change. For military AI, its direct legal reach remains limited: the convention excludes matters of national defense and permits a national security carve-out subject to international law and democratic institutions. Its importance therefore lies in its use as a parallel model for operationalizing impact assessment, rather than as a direct military AI rulebook.

Thomas Reinhold

PRIF – Peace Research Institute Frankfurt

1 The White House. (2025, January 23). *Removing Barriers to American Leadership in Artificial Intelligence* (Executive Order 14179). United States of America. https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/

2 The White House. (2025, July). *Winning the Race: America's AI Action Plan*. United States of America. https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf

3 The White House. (2026, June 2). *Promoting Advanced Artificial Intelligence Innovation and Security* (Executive Order 14409). United States of America. https://www.whitehouse.gov/presidential-actions/2026/06/promoting-advanced-artificial-intelligence-innovation-and-security/

4 The White House. (2026, June 5). *Artificial Intelligence in the National Security Enterprise* (National Security Presidential Memorandum 11). United States of America. https://www.whitehouse.gov/presidential-actions/2026/06/national-security-presidential-memorandum-nspm-11/

5 U.S. Department of Defense. (2023, January 25). *Autonomy in Weapon Systems* (DoD Directive 3000.09). United States of America. https://media.defense.gov/2023/jan/25/2003149928/-1/-1/0/dod-directive-3000.09-autonomy-in-weapon-systems.pdf

6 European Systemic Risk Board. *Warning of the European Systemic Risk Board of 25 June 2026 on systemic cyber*

risks stemming from frontier artificial intelligence models (ESRB/2026/3). http://data.europa.eu/eli/C/2026/3795/oj

7 Présidence de la République française. (2025, February 11). *Paris Declaration on Maintaining Human Control in AI enabled Weapon Systems*. https://www.elysee.fr/emmanuel-macron/2025/02/11/paris-declaration-on-maintaining-human-control-in-ai-enabled-weapon-systems

8 Ministry of Foreign Affairs, European Union and Cooperation. (2026). *A Coruña Declaration: REAIM 2026 Pathways to Action*. Government of Spain. https://www.exteriores.gob.es/en/REAIM2026/Documents/REAIM%202026%20Pathways%20to%20Action.pdf

9 United Nations General Assembly. (2025, December 1). *Artificial intelligence in the military domain and its implications for international peace and security* (A/RES/80/58). United Nations. https://docs.un.org/A/RES/80/58

10 United Nations General Assembly. (2025, December 1). *Possible risks of the integration of artificial intelligence into command, control and communications systems of nuclear weapons* (A/RES/80/23). United Nations. https://docs.un.org/A/RES/80/23

11 Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems. (2026). *Convention on Certain Conventional Weapons. Report of the 2024-2025-2026 sessions of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems* (CCW/GGE.1/2026/CRP.1/Rev.1). United Nations Office for Disarmament Affairs. https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2026)/CCW-GGE.1-2026-CRP.1-Rev.1_-_03_September_.pdf

12 Federal Ministry of Defense. (2026, April 9). *Konzeption Künstliche Intelligenz im Geschäftsbereich des BMVg K-9000/102, Öffentliche Fassung*. Federal Republic of Germany. https://www.bmvg.de/resource/blob/6143700/9954e2b6b993b17cd3d09ca31c506774/dl-ki-konzeption-data.pdf

13 North Atlantic Treaty Organization. (2026). *Alliance Digital Strategy*. https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2026/01/13/alliance-digital-strategy

14 Ministry of Foreign Affairs of the People's Republic of China. (2025, April 21). *Opportunities and Challenges Posed to International Peace and Security by the Application of AI in the Military Domain - Document submitted by China in accordance with General Assembly resolution 79/239* (2024). People's Republic of China. https://www.mfa.gov.cn/eng/wjb/zzjg_663340/jks_665232/kjlc_665236/AI/202504/t20250421_11598983.html (accessed 17 September 2026).

15 European Parliamant and Commission. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance)*. European Union. https://data.europa.eu/eli/reg/2024/1689/oj

16 European Commission. (2025). *General-purpose AI obligations under the AI Act*. European Union. https://digital-strategy.ec.europa.eu/en/factpages/general-purpose-ai-obligations-under-ai-act
European Commission. (2026). *The General-Purpose AI Code of Practice*. European Union. https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai

17 European Commission. (2025). *Guidelines on obligations for General-Purpose AI providers*. European Union. https://digital-strategy.ec.europa.eu/en/faqs/guidelines-obligations-general-purpose-ai-providers

18 Council of Europe. (2026, May 15). European Union ratifies the Council of Europe Framework Convention on Artificial Intelligence. *European Union*. https://www.coe.int/en/web/artificial-intelligence/-/european-union-ratifies-the-council-of-europe-framework-convention-on-artificial-intelligence

19 Council of Europe. *HUDERIA – risk and impact assessment of AI systems*. European Union. https://www.coe.int/en/web/artificial-intelligence/huderia-risk-and-impact-assessment-of-ai-systems

Addgene: A nonprofit repository that collects, stores, and distributes plasmids contributed by research laboratories. Its records often include sequence information, depositor details, plasmid maps, and links to related publications.

Aptamer: A short DNA or RNA molecule designed to bind to a specific target, such as a protein.

Attitude control: The process of changing and maintaining the orientation of a spacecraft, to control which direction it is facing in space.

Attribution: The process of identifying the actor responsible for an incident or activity through technical, contextual, and political analysis.

Backbone: The basic structural part of a plasmid that contains the elements needed for it to function and be maintained in a cell. Additional genetic material is inserted into this underlying framework.

Bacteriophages: A virus that infects bacteria and uses the bacterial cell to reproduce.

Chiral: Molecules that exist in two forms that are mirror images of each other, like left and right hands, but cannot be perfectly aligned.

Codon usage patterns: Different ways of writing the same genetic instructions in DNA. These patterns can sometimes provide clues about how a sequence was designed.

CZT detector: Semiconductor radiation sensor that uses cadmium zinc telluride to directly convert radiation particles into electrical signals at room temperature. It is primarily used for the detection of gamma rays and X-rays.

Departure burn: A rocket engine firing that provides the velocity needed for a spacecraft to leave its current orbit and begin its journey toward another destination.

Deterrence by denial: A strategy that seeks to deter an action by making it unlikely to succeed or by making the objective too difficult to achieve. By depriving an aggressor of confidence in its ability to accomplish its goals, the strategy aims to discourage aggression before it occurs.

Dual-use technology: Technology that can serve both civilian and military purposes; be used offensively or defensively; or be misused for unintended purposes.

Engineered vector: A DNA molecule that has been deliberately modified to carry genetic material into a cell. Engineered vectors are used as tools in genetic engineering and can include plasmids and other delivery systems.

Flyby: A spacecraft trajectory that brings a satellite into close proximity with another object while continuing on a non-capture path.

Foundry: Shorthand for semiconductor fabrication plant. This refers to factories that use photolithography to create integrated circuits from silicon wafers. These devices are commonly used as cutting-edge microprocessors (including AI chips), solid-state memory, photodetectors, embedded processors, or microcontrollers.

GC composition: The proportion of guanine (G) and cytosine (C) bases in a DNA sequence. Differences in GC composition can reflect biological origin or design choices.

General-Purpose Criterion (GPC): A principle within the Biological and Chemical Weapons Conventions that defines prohibited weapons according to their intended harmful use, rather than by reference to specific substances.

Hibernation orbit: An orbit in which a spacecraft remains for an extended period with minimal activity while awaiting a future mission phase, conserving fuel and power until operations resume.

L-band: The 1 to 2 gigahertz (GHz) range of the radio spectrum, as designated by the Institute of Electrical and Electronics Engineers (IEEE).

National technical means (NTM): A nation's independent intelligence and monitoring assets, usually used in the context of arms control treaties.

Non-destructive assay (NDA): NDA techniques encompass a range of methods and instruments that enable the on-site characterization of nuclear materials during inspections, often without the need to take samples. Wherever possible, NDA is preferred because it eliminates the need for international transport, prevents secondary waste during sample disposal, and often produces at least preliminary results during the inspection itself.

Nuclear command, control, and communications (NC3): The overarching framework of hardware, software, and personnel that allows national leaders to monitor threats, authorize the use of nuclear weapons, and transmit launch orders to military forces in the field.

Photomultiplier tube (PMT): Vacuum tube containing a photocathode that emits electrons upon exposure to electromagnetic radiation such as UV or visible light.

A dynode structure under high voltage (1,000+ V) multiplies these electrons, which are then measured at an anode, resulting in an electric charge signal.

Plasmid: A small, usually circular piece of DNA that can exist separately from a cell's main chromosome. In biotechnology, plasmids are commonly used to carry and introduce genetic material into cells.

Plasmid design: The planned arrangement of genetic elements within a plasmid, such as inserted genes, promoters, and other functional components. Different design choices can reflect the purpose of an experiment and the methods used to build it.

Plasmid repository: A collection or database of plasmids that researchers can deposit, document, and often share with others. Repositories may also contain information about who submitted a plasmid, its sequence, its components, and related publications.

Promoter: A DNA sequence that controls when and how strongly a gene is switched on. Promoter choice is an important part of genetic design.

Radio frequency (RF): The rate of oscillation of an electromagnetic wave or electrical signal, typically ranging from about 3 kilohertz (kHz) to 300 gigahertz (GHz) (1 Hertz = 1 oscillation per second).

Rendezvous and proximity operations (RPO): Orbital maneuvers where two spacecraft intentionally approach the same orbit and maneuver in close proximity, sometimes culminating in physical docking.

Revisit frequency: The time a satellite system takes to rotate through its orbit and pass over the exact same geographic location on the rotating Earth again.

Scintillator: Material that emits light when exposed to ionizing radiation, either through direct ionization, as in the case of alpha or beta rays, or through secondary emissions induced by gamma rays or neutrons. This light emission can then be detected by light sensors, which usually provide a signal proportional to the ionizing energy deposited in the scintillator.

Sequence of Concern (SOC): A DNA or RNA sequence associated with a harmful pathogen, toxin, or biological function that could pose a biosecurity risk.

Silicon photomultiplier (SiPM): Solid-state semiconductor device comprising many pixels of avalanche diodes, capable of producing an electrical signal in response to single light photons. These devices are extremely compact, with more than 1,000 pixels per mm2, and operate at low voltages ranging from 20 to 100 V.

Small modular reactor (SMR): A class of nuclear fission reactor with an electrical output of less than 300 MW, based on modular design principles to enable faster and lower cost construction, as well as more flexible and scalable deployment across a range of applications, including general electricity production, facility power, and industrial heat production.

Tomography: Imaging of an object in sections using a penetrating probe signal. By acquiring multiple individual projections or slices, a computational tomographic reconstruction can be used to recreate a 2D or 3D image of the original object and its interior.

Uncrewed aerial vehicle (UAV): An aircraft without a human pilot aboard, commonly known as a drone.

A

ABM — see Anti-Ballistic Missile Treaty
Addgene 49 ff., 137
Agentic AI 89, 93, 95 ff., 133
Alphabet 82
Amazon 82
Antarctic Treaty 122
Anthropic 88
Anti-Ballistic Missile Treaty 33, 124
Arms race 69, 135
Artificial intelligence 18 ff., 33, 39, 46, 53, 61, 69, 80 ff., 86 ff., 91 ff., 95 ff., 101, 108 ff., 127, 129, 133 ff.
Attribution 13 ff., 48 ff., 55, 58, 60, 62 f., 95, 137
Australia 82, 109 f., 115 f.
Australia Group 130

B

Bacteriophage 106, 137
BAFA — see Federal Office for Economic Affairs and Export Control
Bangkok, Treaty of 123
Ballistic Missile Launch Notification Agreement 126
Biological agents 48, 127 f.
Biological Design Tools 86
Biological screening 27
Biological Weapons Convention 28, 52, 65, 67, 69, 127 f., 130 f.
Biological weapons 28, 49, 88, 126, 127 ff.
Biosafety 28, 89, 128, 131
Biosecurity 48, 52 f., 87, 90, 106, 128, 131, 138
Biotechnology 48 ff., 86 ff., 104 ff., 111, 127, 138
Black box 18 ff., 92
BWC — see Biological Weapons Convention

C

Chain of custody 25, 29
Chemical Warfare Agents 91, 128
Chemical Weapons Convention 15, 28, 60 ff., 65, 67, 69, 91, 93, 98, 128 ff., 137
Chemical weapons 28 f., 60 ff., 67, 72, 91 ff., 95 ff., 128 ff.
ChemTech Centre 129
Cherenkov radiation 41, 43
China 33 f., 55, 71, 74, 80 ff., 112, 124 f., 131, 134 f.
Chirality 104 ff.
Comprehensive Nuclear-Test-Ban Treaty 75, 122
Comprehensive Nuclear-Test-Ban Treaty Organization 22
Convention on the Physical Protection of Nuclear Material and Nuclear Facilities 126
COVID-19 48
Critical infrastructure 60, 62, 84, 115, 133,
Cruise missile 100, 126
CTBT — see Comprehensive Nuclear-Test-Ban Treaty
CTBTO — see Comprehensive Nuclear-Test-Ban Treaty Organization
CWC – see Chemical Weapons Convention
Cyber arms control 60 ff.
Cyber weapons 60 ff., 69 f.
Cybersecurity 60, 112, 133, 136

D

Data Fusion 13, 15, 33 ff.
Data Integrity 20 ff.
Data minimization 20 f., 28, 31
Data poisoning 20
DDBJ 50
Decentralized verification 20 f.
DNA 48 ff., 86 ff., 104 ff., 138 f.
Disinformation 73, 95 ff.
Drones — see Uncrewed aerial vehicles
Dual-use 12, 15, 25, 46, 58, 60 ff., 86, 91 ff., 97, 112, 126, 127 ff., 138

E

ENA 50
EU — see European Union
EU AI Act 136
EU Space Threat Response Architecture 36
European Space Imaging 39
European Union 36, 69, 82, 89, 101, 112 f., 125, 130, 133, 136
Export controls 80 ff., 112, 125 f., 130
Explainable AI 19 f.

F

Federal Office for Economic Affairs and Export Control 130
Fissile material 29, 41, 44, 56, 123
Fissile Material Cut-off Treaty 123

G

γ-rays 41 ff.
General-Purpose Criterion 60 ff., 138
Generative AI 19
Geneva Protocol 127
GenBank 50 f.
Genetic engineering 13 f., 48 ff., 138
German Ethics Council 131
German Research Foundation 131
Germany 34, 36 f., 80, 82, 101, 109, 112, 115 f., 125, 127 ff., 135
Graph neural networks 92
Gravimetry 25 f., 28 ff.
Gravity gradiometry 24 f., 28

H

Hague Code of Conduct Against Ballistic Missile Proliferation 126
Hague Ethical Guidelines 131
High-Bandwidth Memory 81

I

IAEA — see International Atomic Energy Agency
IAP — see InterAcademy Partnership
IBBIS — see International Biosecurity and Biosafety Initiative
ICBM — see Intercontinental ballistic missile
IDC — see International Data Centre
IMS — see International Monitoring System
Induced coherence 27
INF Treaty 33
Information barriers 20, 58
Inertial navigation 13, 25 f., 29, 31
InterAcademy Partnership 131
Intercontinental ballistic missile 123, 126
International Atomic Energy Agency 13, 16 f., 21 f., 56, 68, 73, 75, 122, 125
International Biosecurity and Biosafety Initiative 88
International Data Centre 122
International Humanitarian Law 68, 134 f.
International Monitoring System 122
Iran 37, 71, 74, 100 ff., 125, 130
Iran Nuclear Deal — see Joint Comprehensive Plan of Action
Iraq 130
Israel 34, 71, 100, 124 f., 128

J

Japan 34, 82, 109, 111 f., 115 f.
JCPOA — see Joint Comprehensive Plan of Action
Joint Comprehensive Plan of Action 71, 125

L

Laser spectroscopy 44
Large language models 19, 91 ff., 95 ff., 114
Life sciences 86 ff., 130 f.
Lisbon Protocol 125 f.

M

Machine learning 13, 19, 22, 46, 49, 52, 56, 92
Managed access 28 f., 62, 86, 89 f.
Magnetometry 25 f., 28 f., 31
Meta 82, 96
MGX 82
Microsoft 82
Mid-infrared spectroscopy 27 f.
Microbial forensics 16, 48 ff.
Mirror life 104 ff.
Missile Technology Control Regime 126

N

NAR — see Novel Advanced Reactor
National Academy of Sciences Leopoldina 131
National technical means 15, 19, 33 f., 71, 123, 138
NBC weapon — see Weapons of Mass Destruction
Neutron beam interrogation 44 f.
Neutrons 41 ff., 139
New START — see New Strategic Arms Reduction Treaty
New Strategic Arms Reduction Treaty 71, 123 f.
Nonlinear interferometry 24 f., 27 ff.
Non-state actors 33, 37, 60, 71 ff., 97 f., 115, 126, 129 f., 134
North Korea 71, 74, 124, 128, 130
Novel Advanced Reactor 40
NPT — see Treaty on the Non-Proliferation of Nuclear Weapons
Nuclear resonance fluorescence 44
Nuclear Suppliers Group 122
Nuclear verification 40 ff., 72
Nuclear weapons 40, 58, 68, 122 ff., 138
Nucleotide 50
NV-center magnetometry 28

O

OpenAI 82, 88
Open-source data 93,
Open-source intelligence 15, 71 ff.
Open Skies Treaty 23, 35, 132
OPCW — see Organisation for the Prohibition of Chemical Weapons
OPCW Centre for Chemistry and Technology — see ChemTech Centre
Oracle 82
Organisation for the Prohibition of Chemical Weapons 13, 17, 28, 62, 67, 92 f., 95 ff., 128 f.
OSINT — see Open-source intelligence
Outer Space Treaty 123

P

Pakistan 124, 131
Paris Declaration 134
Partial Test Ban Treaty 122, 124
Pathogens 14, 28, 86 ff., 130, 139
Peaceful Nuclear Explosions Treaty 124
Pelindaba, Treaty of 123
Plasmid design 49, 139
Plasmid repository 49, 139
Plutonium 41, 43, 74, 125

PNE — see Peaceful Nuclear Explosions Treaty
PTBT — see Partial Test Ban Treaty

Q

Quantum sensing 13, 24 ff., 116

R

Radioactivity 40 f., 44, 46, 122
Radionuclide 122
Rare earths 80, 84
Rarotonga, Treaty of 123
REAIM 134
Reconnaissance 15, 33 ff., 38, 132
Remote monitoring 41
Reproducibility 18, 21 ff., 28 f.
Retrosynthetic models 93
Ribosome 104 f.
RNA 87 f., 105, 139
Robotics 45
Russia 34, 58, 67, 82, 100 ff., 108, 113 ff., 123 ff., 130, 132

S

Sanctions 125, 130
Satellite 12 f., 15, 25 f., 33 ff., 55 ff., 71, 101, 122, 138 f.
Satellite imagery 12 f., 15, 33 ff., 71 ff., 122, 125
Security-relevant research 131
Seabed Arms Control Treaty 123
Semiconductors 43, 45, 80 ff., 112, 138 f.
Semipalatinsk, Treaty of 123
Small Modular Reactor 13, 40, 139
SMR — see Small Modular Reactor
Social media 72 f., 95 ff.
Societal verification 15, 71 ff.
South Africa 34, 125
Southeast Asia 123
Soviet Union 15, 33, 123 f.
Stealth 34 f.
Stargate 82
Submarine cables 83
Supply chains 80 ff., 112
Syria 67, 97 f., 128 ff.

T

Taiwan 83
Terrorism 113, 126, 128 f.
Thalidomide 105
Tianjin Guidelines 131
Threshold Test Ban Treaty 124
Tlatelolco, Treaty of 123
Toxins 67, 87, 127, 129 f., 139
TPNW — see Treaty on the Prohibition of Nuclear Weapons
Training data 13, 19 f., 22, 87, 89
Treaty on the Non-Proliferation of Nuclear Weapons 65, 68 f., 71, 122, 125 f.
Treaty on the Prohibition of Nuclear Weapons 122
TTBT — see Threshold Test Ban Treaty

U

UAV — see Uncrewed aerial vehicles
Ukraine 97, 100 ff., 108, 113 ff., 125 f.
Uncrewed aerial vehicles 12, 24, 26, 28, 30, 35, 104 ff., 113, 124, 126, 139
United Kingdom 26, 34, 74, 82, 104, 109, 113 f., 116, 125
United Nations 56, 60, 63, 68 f., 98, 104, 122, 125 f., 129 f., 134 ff.
United Nations Office for Disarmament Affairs 129
United Nations Secretary-General's Mechanism 129
United Nations Security Council 125 f., 129 f.
United States of America 15, 33 f., 36, 40, 48, 56, 58, 71, 73 ff., 81 ff., 88, 101, 109 f., 112 f., 116, 123 ff., 130 f., 132, 133 f.
UNODA — see United Nations Office for Disarmament Affairs
UNSC — see United Nations Security Council
UNSGM — see United Nations Secretary-General's Mechanism
Uranium 41, 43, 68, 125
USA — see United States of America

V

Visible readout 27

W

Wassenaar Arrangement 126
Weapons of Mass Destruction 123, 126, 129
WHO — see World Health Organization
World Health Organization 48, 98, 130 f.

X

XAI — see Explainable AI

Z

Zangger Committee 125

Dr. Philip Gasse is a researcher in the research group Chemical and Biological Weapons Control. His research focuses on investigating the impact of AI on chemical weapons arms control in order to identify opportunities and risks and develop new regulatory measures or verification mechanisms.

Prof. Dr. Malte Göttsche is professor for Peace Research in Natural Sciences at PRIF and TU Darmstadt, co-speaker of CNTR, and head of the research group Science for Nuclear Diplomacy. He researches nuclear non-proliferation, arms control and disarmament with a focus on the development of verification methods and technologies.

Prof. Dr. Markus Gräfe is a professor in the field of Experimental Solid State Quantum Optics at the Institute for Applied Physics at TU Darmstadt, where he uses quantum optics to research new sensing modalities, e.g., for cancer diagnostics in the "QUANCER" project.

Prof. Dr. Owen Greene is professor of International Security and Development at the University of Bradford, UK; and visiting professor at PRIF, Germany. He is also co-founder and chair of VERTIC (Verification Research, Training and Information Centre).

Dr. Una Jakob is co-head of the research group Chemical and Biological Weapons Control. Her research focuses on the non-proliferation and disarmament of biological and chemical weapons as well as on biosecurity issues and the governance of research with dual-use potential.

Abdullah Kasim Korkusuz is a student assistant in the research group Emerging Disruptive Technologies.

Anja-Liisa Krichbaum is a research associate and doctoral candidate at PEASEC/TU Darmstadt. Her research interests lie in the areas of (cyber) arms control, autonomous weapons systems, meaningful human control, norm building, and critical security studies.

Dr. Sarah McKee is a researcher in the research group Science for Nuclear Diplomacy. Her research interests include nuclear high-altitude electromagnetic pulse, the dual use of nuclear technologies in space, space nuclear weapon delivery systems, orbital behaviours of satellites and the interrogation of orbital payloads.

Dr. Robin E. Möser is a researcher in the research group Science for Nuclear Diplomacy. His research focuses on nuclear disarmament and verification issues as well as arms control and non-proliferation more generally.

Dr. Luis Pazos Clemens is a researcher in the research group Science for Nuclear Diplomacy. He conducts research in the field of nuclear disarmament and non-proliferation with a focus on the investigation and development of verification techniques.

Helene Pleil is researching cyber diplomacy and cybersecurity and is currently a senior policy researcher in interface's Cybersecurity & Emerging Threats program.

Lukas Rademacher is a doctoral researcher in the research group Science for Nuclear Diplomacy. He is researching novel verification methods for nuclear non-proliferation and disarmament, with a focus on nuclear archaeology.

Dr. Thomas Reinhold is a researcher in the research group Emerging Disruptive Technologies. He conducts research on the militarization of cyberspace, AI and possibilities for arms control and disarmament of these technologies.

Dr. Kadri Reis is a researcher in the research group Chemical and Biological Weapons Control. Her expertise includes in particular biosecurity and bioweapons.

Prof. Dr. Dr. Christian Reuter is co-head of the research group Emerging Disruptive Technologies and a professor at TU Darmstadt. His chair, Science and Technology for Peace and Security (PEASEC), combines computer science with peace and security research.

Dr. Yan-Jie Schnellbach heads the junior research group "nuSENTRY" (Neutrino-based Safeguards for Emerging Nuclear Technology and Reactor Types) at TU Darmstadt. His research focuses on nuclear non-proliferation, with an emphasis on detection and measurement technologies.

Dr. Niklas Schörnig is co-head of the research group Emerging Disruptive Technologies. He focuses on emerging military technology, the future of warfare and arms control.

Dr. Adhitya Sekhar is a member of the junior research group "nuSENTRY" (Neutrino-based Safeguards for Emerging Nuclear Technology and Reactor Types) at TU Darmstadt.

Liska Suckau is a researcher in the research group Emerging Disruptive Technologies. Her research interests are war and peace research, in particular new technologies in warfare, military innovation, critical military studies, arms control and gender.

Fabian Unruh is a doctoral researcher in the research group Science for Nuclear Diplomacy. He works in the field of nuclear disarmament verification and is interested in applying machine learning techniques in nuclear archaeology for reconstructing fissile material production histories.

CNTRMONITOR 2026

PRIF – Peace Research Institute Frankfurt

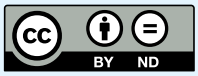
Darmstädter Landstraße 112
60598 Frankfurt am Main
Phone: +49 69 95 91 04–0
E-Mail: cntr@prif.org
www.cntrarmscontrol.org

ISBN: 978-3-911092-08-1
DOI: 10.48809/cntr2026

Editors: Prof. Dr. Malte Göttsche, Dr. Una Jakob, Dr. Niklas Schörnig
Cluster for Natural and Technical Science Arms Control Research (CNTR)

Cover image: Satellite view of Al-Basrah, Iraq. Source: USGS via Unsplash, Unsplash License.

Text license: Creative Commons CC-BY-ND 4.0 (Attribution/No Derivatives/4.0 International).
The images used are subject to their own license conditions.



Editorial staff and typesetting: Elisabeth Waczek, CNTR

Assistance: Solveig Albrecht, Adele Knof

Language editing: Carla Welch

Layout design: Susanne Hartmann/media machine GmbH

November 2026

Recommended citation: Göttsche, M., Jakob, U., & Schörnig, N. (Eds.). (2026). *Verification in an Era of Rapid Technological Change. CNTR Monitor – Technology and Arms Control 2026*. PRIF – Peace Research Institute Frankfurt.

ACKNOWLEDGEMENTS

We would like to thank Sara Al-Sayed, Jürgen Altmann, Oliver Crook, Stefan Frey, Zoe Gastelum, Michal Krelina, Ralf Trapp and József Zsigrai for their comprehensive review and valuable feedback on earlier versions of the CNTR Monitor chapters.

We thank the Federal Foreign Office for their continued funding and support. Yan-Jie Schnellbach and the nuSENTRY junior research group at TU Darmstadt would like the thank the Federal Ministry of Research, Technology and Space for their funding and support.

Cluster for Natural and Technical Science Arms Control Research

CNTRMONITOR 2026

Verification in an Era of Rapid Technological Change

Rapid scientific and technological advancements may change the verification landscape of the future. Only few arms control treaties with formal verification regimes remain in place. Still, producing collective knowledge on activities that could destabilize peace and security remains crucial. The CNTR Monitor 2026 examines new opportunities for verification—with and without treaties.

Verification knowledge needs to be reliable and internationally accessible. New technologies can help improve existing verification regimes but also present novel challenges. By analyzing selected technologies and political shifts, the CNTR Monitor 2026 sheds light on the dynamics that shape verification in the new security environment.

cntrarmscontrol.org

Gefördert durch:

Auswärtiges Amt

aufgrund eines Beschlusses des Deutschen Bundestages

CNTR is a joint project between:

LEIBNIZ INSTITUT FÜR FRIEDENSFORSCHUNG

IZNK

HSFK

ITLH

TECHNISCHE UNIVERSITÄT DARMSTADT

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN