VCDNP Vienna Center for Disarmament and Non-Proliferation

April 2025

Nuclear Security in a Changing World

Nuclear Security and the Nuclear Supply Chain in the Age of Artificial Intelligence

Dr. Sarah Case Lackner
Mara Zarka

Vienna Center for Disarmament and Non-Proliferation | vcdnp.org

Authors

Dr. Sarah Case Lackner is a Senior Fellow at the VCDNP. Her work focuses on nuclear security and its interactions with AI and other emerging and disruptive technologies. Among other positions, she was a Senior Nuclear Security Officer at the International Atomic Energy Agency (IAEA), serving as the Scientific Secretary for the Nuclear Security Guidance Committee and the Director General's Advisory Committee on Nuclear Security. She also served as Co-Scientific Secretary of the 2022 Conference of Parties to the A/CPPNM.

Mara Zarka is a Research Associate and Project Manager at the VCDNP, where her research addresses the intersection of emerging and disruptive technologies with nuclear, the security of nuclear and radiological materials against malicious non-State actors, and the non-proliferation regime and nuclear governance. Her work has also included projects on nuclear safeguards and peaceful uses of nuclear technologies, among others.

About the VCDNP

The Vienna Center for Disarmament and Non-Proliferation (VCDNP) promotes international peace and security by conducting research, facilitating dialogue, and building capacity on nuclear non-proliferation and disarmament.

The VCDNP is an international non-governmental organisation, established in 2010 by the Federal Ministry for European and International Affairs of Austria and the James Martin Center for Nonproliferation Studies at the Middlebury Institute of International Studies at Monterey.

Our research and analysis provide policy recommendations for decision-makers. We host public events and facilitate constructive, results-oriented dialogue among governments, multilateral institutions, and civil society. Through in-person courses and online resources on nuclear non-proliferation and disarmament, we train diplomats and practitioners working in Vienna and around the world.

Acknowledgements

The VCDNP thanks **Global Affairs Canada** for supporting this research project, including the expert workshop and this publication.

VCDNP
Vienna Center for Disarmament and Non-Proliferation

Andromeda Tower, 13/1
Donau-City-Strasse 6
1220 Vienna
Austria

www.vcdnp.org
info@vcdnp.org

Sponsored by Canada

Contents

Recent leaps in the capabilities of AI systems present risks and opportunities for the nuclear sector, including the nuclear supply chain.

Executive Summary

The recent revolution in artificial intelligence (AI) technology will present new challenges for nuclear security. While AI technology can create new opportunities to strengthen nuclear security, it will also provide new tools and methods for malicious actors to undertake cyberattacks and target AI systems integrated into nuclear facilities. It may also provide malicious actors with new opportunities to target the nuclear supply chain, with significant consequences for nuclear security.

On 14 and 15 January 2025, the Vienna Center for Disarmament and Non-Proliferation (VCDNP) convened a workshop entitled “Nuclear Security in a Changing World: Exploring Evolving Supply Chain Risks related to Artificial Intelligence”. This workshop, funded by Global Affairs Canada, brought together experts in AI technology, supply chain risks, cyber security, nuclear operations, and nuclear security for two days of intensive discussions. The current report, which draws on the discussions in the workshop as well as expert research, provides an overview of risks and opportunities related to AI technology, nuclear security, and the nuclear supply chain, focusing on three aspects: AI systems integrated into facilities; AI models used by malicious actors; and data security challenges associated with AI. Several conclusions on this topic are also provided, aimed at States and international organisations.

AI systems are already being used in some applications in nuclear facilities, and a range of further applications are being considered. Their continued integration, including in operational technology, will pose new challenges for nuclear security as well as security of the nuclear supply chain. This is due to aspects of AI systems that may be associated with novel supply chain vulnerabilities, such as the importance of their training data and the inscrutable nature of their calculations.

To prevent a successful attack, operators and regulators will need to have confidence that AI systems installed into facilities will reliably function as expected and clarify the ultimate responsibility for actions resulting from AI outputs. Remaining security risks associated with their supply chain will need to be balanced against the benefits these systems provide to nuclear facilities.

Increasingly powerful AI systems are also already in the hands of malicious actors. Currently available and future commercial and open-source AI models may enhance the capabilities of criminals, terrorists, or others to threaten nuclear security. While the purpose of this report is not to sketch out detailed scenarios, such AI models could assist a malicious actor, for example, to improve the generation of counterfeit certifications, to undertake sophisticated social engineering, or to quickly and powerfully process data from images. These examples are enhancements of known capabilities of malicious actors, but innovative modes of supply chain attacks may also develop alongside AI technology advancements.

If relevant data is not properly secured, it can provide a pathway for an attack on nuclear facilities, including via the nuclear supply chain. As AI systems are increasingly used in nuclear facilities, even if only on the business side, the ever larger quantities of potentially sensitive and even export-controlled data processed by these systems will need to be monitored and managed. If data is inadvertently released, including data that might reveal helpful details about the nuclear supply chain, it could be misused by a malicious actor whose own capability to process large quantities of data and draw conclusions has been enhanced by AI systems and models.

To mitigate this risk and ensure continuing security of the nuclear supply chain in the age of AI, there are several broad actions that policy-makers, national regulators, international organisations, and the nuclear security community as a whole should consider taking, as follows.

Guidance, policies, and regulations to secure the nuclear supply chain need to proactively account for the risks and benefits of AI technology. Nuclear security implications of AI systems and models need to be considered in national threat assessments and national regulations. Further, international organisations, non-governmental organisations (NGOs), industry organisations, and other stakeholders can provide guidance and information on mitigating the impact of AI systems and models on nuclear security, including for security of the nuclear supply chain. The international sharing of information on the intersection of AI, nuclear security, and the nuclear supply chain will help to provide broader understanding, particularly via the sharing of case studies. Finally, these discussions need to be informed by and, as possible, integrated into broader international discussions on AI governance.

Capacity-building, awareness-raising, and training on the nexus between AI systems and models and nuclear security is needed, particularly with respect to the nuclear supply chain, among a range of stakeholders, including in national governments, regulators, international organisations, NGOs, and the nuclear industry. This capacity-building, awareness-raising, and training would highlight how AI systems and models can benefit nuclear security, the risks they could pose to nuclear security and the security of the nuclear supply chain, and the importance of data security.

AI systems can be used by nuclear and cyber security professionals to help secure the nuclear supply chain, including to detect deepfakes and other false credentials, map vulnerabilities in the supply chain, implement cyber security measures, and other tasks.

Continued research is needed on the relationship between nuclear security and AI systems and models to alert of emerging concerns and to help prepare for future disruptive developments related to AI and other advanced technologies.

AI-powered manufacturing plant Xunxi by the Chinese Alibaba Group. Credit: Xunxi via IAEA.

Artificial Intelligence, Nuclear Security, and the Nuclear Supply Chain

Artificial intelligence (AI), nuclear security,[1] and the nuclear supply chain will only increase in importance in the coming decades. AI science and applications are developing at a breakneck pace, while new geopolitical challenges and a need for increasing amounts of low-carbon energy will make nuclear energy – and the security of nuclear facilities – ever more relevant. At the same time, the supply chain for all manufactured goods is becoming more international and convoluted, providing new opportunities for infiltration by criminals and other malicious actors.

To ensure that society can benefit from the many positive aspects of AI and nuclear energy technologies while preventing malicious actors from exploiting new vulnerabilities, national governments and regulatory bodies, nuclear operators, and international organisations, among others, must consider the intersection of these three topics.

[1] In the context of this report, nuclear security refers to the security of nuclear materials and facilities, as per the definition in the IAEA "Objective and Essential Elements of a State's Nuclear Regime", where paragraph 1.1 states: "Nuclear security focuses on the prevention of, detection of, and response to, criminal or intentional unauthorized acts involving or directed at nuclear material, other radioactive material, associated facilities, or associated activities. Other acts determined by the State to have an adverse impact on nuclear security should be dealt with appropriately." IAEA Nuclear Security Series, No. 20, 2013, p. 1. Available at: https://www.iaea.org/publications/10353/objective-and-essential-elements-of-a-states-nuclear-security-regime.

While most nuclear security stakeholders are not experts in AI or the nuclear supply chain, they need to be cognisant of the rapidly shifting technological and geopolitical landscape of these topics, as they will affect the capabilities and methods of modern malicious actors. These potential capabilities and methods will, in turn, need to be considered in the development and implementation of policies, regulations, and international agreements that maintain nuclear security.

To provide insights and recommendations to help nuclear security stakeholders to better understand and manage rapidly shifting challenges in this area, the Vienna Center for Disarmament and Non-Proliferation (VCDNP) convened a workshop on 14 and 15 January 2025, entitled "Nuclear Security in a Changing World: Exploring Evolving Supply Chain Risks related to Artificial Intelligence".

This workshop, funded by Global Affairs Canada, convened 23 experts in AI, supply chain risks, cyber security, nuclear operations, and nuclear security. These various experts brought perspectives on these topics from around the world, including Africa, Europe, North America, South America, and the Middle East, as well as experience in the nuclear industry, nuclear suppliers, nuclear regulators, international organisations, national laboratories, academia, and national governments. The intense and fruitful discussions over the two days shaped the current report, which outlines the intersection of AI, nuclear security, and the nuclear supply chain, highlights key concerns, and provides conclusions and recommendations for governments, regulatory bodies, international organisations, and operators to consider.

Artificial Intelligence: History, Terminology, and Applications in the Nuclear Sector

The phrase "artificial intelligence" conjures up, for many, images of humanoid robots featured in science fiction novels and television series, such as Isaac Asimov's *I, Robot* or *Star Trek, the Next Generation*. Modern, real-world uses of AI, while less obviously humanlike, can also seem to have human or even superhuman skills, confounding our understanding of the world and leading humans to both overestimate and underestimate their capabilities. For this reason, as well as the relative novelty of AI for many in the nuclear field, before initiating the detailed discussion of the later sections of this report, we will briefly examine the current and near-term projected uses and capabilities of AI, with an eye to applications in the nuclear sector.

The term "artificial intelligence" was first coined by Professor John McCarthy to identify machines which could perform tasks that are characteristic of human intelligence.[2] While the exact definition of AI is currently controversial, this can serve as a useful working definition consistent with how the term is used in the technology sector. Google defines artificial intelligence as "a field of science concerned with building machines that can reason, learn, and act in such a way that would normally require human intelligence",[3] and IBM provides a similar definition, explaining that AI is "technology that enables computers and machines to simulate human learning, comprehension, problem solving, decision making, creativity and autonomy".[4]

Tasks characteristic of human intelligence include determining patterns from data, vision, speech, making decisions, and the creation of novel images and text, among others. Through research in the field of AI, many methods and techniques have been developed to allow computers to accomplish such tasks, such as understanding and producing speech (text or spoken), solving problems without being given explicit instruction, and creating novel images and videos.

[2] Calum McClelland, "The Difference between Artificial Intelligence, Machine Learning, and Deep Learning", medium.com, 4 December 2017. Available at: https://medium.com/iotforall/the-difference-between-artificial-intelligence-machine-learning-and-deep-learning-3aa67bff5991.

[3] Google Cloud, "What is Artificial Intelligence (AI)?" Available at: https://cloud.google.com/learn/what-is-artificial-intelligence.

[4] IBM, "What is artificial intelligence (AI)?" Available at: https://www.ibm.com/think/topics/artificial-intelligence.

AI as a field of research started in the 1950s, but until the incredible momentum of the last decade and a half, progress came in bursts of intense activity followed by fallow periods, frequently referred to as “AI winters”.[5] This sudden acceleration in AI development can be credited to three factors:

- **AI models or algorithms:** New models and algorithms have allowed for rapid progress.

- **High-quality data and data access:** Digitised data has become ubiquitous in the internet age, as vast databases of images, enabled rapid advances in image processing and recognition, and large amounts of digitised text have provided the needed input for these sophisticated machine learning algorithms.

- **Computational power:** Technological and materials science advances over the last few decades have enabled a massive increase in the computational power that can be devoted to AI systems, both their development and use.[6]

In particular, new algorithms and techniques related to **machine learning**, a field that seeks to mimic human learning, have been key to the recent explosion in powerful AI systems.[7] They have been able to exploit the availability of data and computational power cited above to make great leaps in the last decade. Such machine learning techniques, including neural networks and associated deep learning, have provided the basis for the development of ever more sophisticated **AI models**, software programmes that detect patterns from data sets. Machine learning requires a significant quantity of quality data on which it can be taught to make the needed connections for its intended application, known as **training data.**[8]

AI models can be broadly categorised as predictive or generative. **Predictive AI** models use statistical techniques and machine learning to analyse and solve computationally complex and data intensive problems. Traditionally, most AI applications in industry have focused on predictive AI systems.[9] **Generative AI** models identify patterns and relationships in the data they are trained on, such as images, computer code, and text, to “create” new or derived content, whether text, code, images, or videos, based on a user’s prompt. Generative AI systems require vast amounts of computer power and training data along with sophisticated machine learning and other AI methods and techniques.[10] These include the familiar subset of AI models referred to as **Large Language Models (LLMs)**, such as OpenAI’s ChatGPT, Anthropic’s Claude, and Meta’s Llama.

[5] More discussion of the history of AI science can be found in Artificial Intelligence, a Modern Approach, Fourth Edition by Stuart Russell and Peter Norvig, Hoboken: Pearson, 2021.

[6] For example, the development and use of GPUs (graphical processing units), which are better suited to parallel processing and floating point calculations than CPUs (central processing units), for machine learning applications.

[7] For more discussion of machine learning techniques as well as generative and predictive AI systems, see Donald Dudenhoeffer, “Past, Present, and Future Applications of AI in the Nuclear Sector”, 2025.

[8] Various types of training are used to train AI systems using machine learning, including supervised machine learning, unsupervised machine learning, and reinforcement learning. While it is beyond the scope of this paper to delve into detail on these processes, each of these types of learning has different potential strengths and weaknesses. More information on types of machine learning can be found here: https://www.ibm.com/think/topics/machine-learning-types.

[9] Donald Dudenhoeffer, “Past, Present, and Future Applications of AI in the Nuclear Sector”, 2025.

[10] More detailed information on generative AI and its implications can be found in N. Bajema, “Generative AI and WMD Nonproliferation: A Practical Primer for policy-makers and Diplomats”, CNS Occasional Paper, No. 63, December 2024. Available at: https://nonproliferation.org/wp-content/uploads/2024/12/generative_ai_and_wmd_nonproliferation_12042024.pdf.

These models are available as closed source, as open-source models, or as open-weight models. According to the Open Source Initiative (OSI), an **open-source AI model** permits one to "use the system for any purpose and without having to ask for permission, study how the system works and inspect its components, modify the system for any purpose, including to change its output [and] share the system for others to use with or without modifications, for any purpose."[11,12] In **open-weight AI models**, only the model weights, key parameters that define the internal functioning of an AI model, are provided, but this is not sufficient to meet the OSI definition, as open-weight models allow for fine-tuning of the AI model for specific tasks, but provide little insight into its inner workings.[13] Each of these options has different security implications, as they allow for different degrees of modification by the user (including malicious actors).

An **AI system**, as used throughout this report, refers to an application that uses AI models to perform a specified task that previous generations would have considered to be only possible for humans to accomplish.[14,15] Article 3 of the EU Artificial Intelligence Act specifically defines an AI system as "a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments."[16]

In Figure 1, an AI system is shown visually as a grey box, along with its interaction with the underlying model and the environment. The machine learning algorithm in the AI model is trained on vast amounts of training data, as shown in the box on the right. The AI system takes in information from the environment via **input data**, which it then processes according to the task it has been programmed to do, and returns either **output data** or an action.

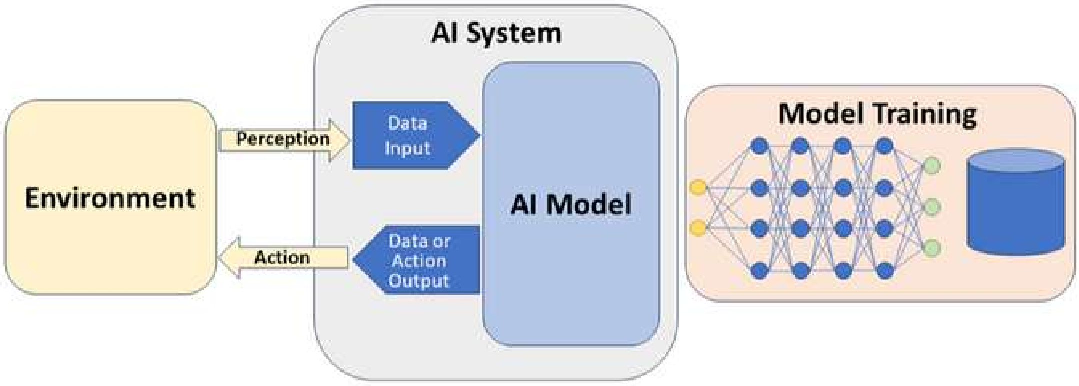


Fig. 1: AI system overview (adapted from OECD),[17] Dudenhoeffer (2025)[18]

[11] Open Source Initiative, "The Open Source AI Definition – 1.0", Available at: https://opensource.org/ai/open-source-ai-definition.

[12] As noted by the the MIT Technology review in August 2024, there is debate whether all AI models claimed to be open-source truly are, according to this definition, or if they might be better described as partially open-source. For more details, consult Rhiannon Williams and James O'Donnell, "We finally have a definition for open-source AI", MIT Technology Review, 22 August 2024. Available at: https://www.technologyreview.com/2024/08/22/1097224/we-finally-have-a-definition-for-open-source-ai/.

[13] Model weights are learnable parameters in machine learning tools that reflect the learned connections between the training data, and are set based on the training data. For neural networks, they represent the strength and direction of connections between the individual neurons.

[14] The report will not regularly make the distinction between varying types of AI systems, including their underlying models, unless it is important to make that distinction in a particular context.

[15] Current AI systems, while they can be convincingly human-like in their interactions, lack what are called "general cognitive abilities". These kinds of broad abilities are what would characterise "artificial general intelligence", which remains in the future (as of April 2025) and which is not addressed in this report.

[16] EU Artificial Intelligence Act, Article 3. Available at: https://artificialintelligenceact.eu/article/3/.

[17] OECD, "Artificial Intelligence in Society", 2019, p. 23. Available at: https://www.oecd.org/en/publications/artificial-intelligence-in-society_eedfee77-en.html.

[18] Donald Dudenhoeffer, "Past, Present, and Future Applications of AI in the Nuclear Sector", 2025.

The idea of using AI systems to increase the effectiveness of operations in nuclear facilities is also not new, and in fact, some proposed applications date from the 1980s.[19] However, many envisioned applications of AI technology have not become practical until recently, with the acceleration in AI development described above. The present scope of proposed and actual uses of AI systems in the nuclear sector includes a range of applications, in reactor modelling, administration, safety, and security, including, among many others:

- Reactor core design
- Detection of counterfeit, fraudulent, and suspect items
- Predictive maintenance
- Process automation
- Report generation
- Knowledge management
- Remote surveillance
- Facial recognition for security purposes

The applications for AI systems listed above, as well as those highlighted in the fuller lists provided by Dudenhoeffer[20] and Huang et al.,[21] are likely to include the use of both generative and predictive underlying models, as well as other techniques to interact with the environment, such as natural language processing and computer vision.

Security of the Nuclear Supply Chain

According to Ganeshan and Harrison (2002), a supply chain is "a network of facilities and distribution options that performs the functions of procurement of materials, transformation of these materials into intermediate and finished products, and the distribution of these finished products to customers."[22]

Nuclear energy production specifically relies on the **nuclear supply chain** to provide products and services, including in design, construction, commissioning, operation, and decommissioning of nuclear facilities.[23] The nuclear supply chain provides materials and parts, such as concrete, pumps, electronics, wiring, computer systems (both hardware and software), and prosaic items, such as heating, ventilation, and air conditioning (HVAC) systems.

Supply chains for all complex manufactured items, from cars to televisions to nuclear power plants, increasingly feature intricate supply chains with myriad suppliers of individual items in various countries. Further, intermediate products are often prepared in separate countries from those who supply individual items, and final products are finished elsewhere. For example, a single car part constructed in North America for the US market might cross a national border multiple times.[24,25]

[19] Ibid.

[20] Ibid.

[21] Qingyu Huang et al., "A review of the application of artificial intelligence to nuclear reactors: Where we are and what's next", Heliyon, Volume 9, Issue 3, 2023. Available at: https://doi.org/10.1016/j.heliyon.2023.e13883.

[22] Ram Ganeshan and Terry P. Harrison, "An Introduction to Supply Chain Management", The University of Melbourne, Version 1,0, p. 1, 2002. Available at: https://static1.squarespace.com/static/5b9e942a8f5130f854dbef81/t/5be89d3b21c67c13123b21bd/1541971264501/an-introduction-to-supply-chain-management.pdf.

[23] The nuclear supply chain as discussed here is distinct from the nuclear fuel cycle, which addresses the preparation of nuclear fuel, its use in the reactor, and its storage and ultimate disposal after use.

[24] Vipal Monga and Santiago Perez, "Track One Car Part's Journey Through the U.S., Canada and Mexico – Before Tariffs", The Wall Street Journal, 1 March 2025. Available at: https://www.wsj.com/business/autos/track-one-car-parts-journey-through-the-u-s-canada-and-mexicobefore-tariffs-7c0d5dcb.

[25] Norman De Bono, "Trump tariffs: How one car piece crosses Canada, U.S., Mexico borders 7 times", The London Free Press, 4 March 2025. Available at: https://lfpress.com/news/local-news/trump-tariffs-car-part-crosses-canada-us-mexico-borders-7-times.

At the same time, the number of counterfeit items in circulation is increasing worldwide, and up to 2.5 percent of world trade is in counterfeited and pirated goods.[26] The nuclear industry is no exception.[27] According to the International Atomic Energy Agency (IAEA), "[e]xperience shows that [suspect, counterfeit, and fraudulent items] include a wide range of items, such as threaded fasteners, piping, [mechanical] components, and electrical [and electronic] components." Bulk materials and chemicals can also be of concern.[28] A 1990 study by the US General Accounting Office reported that over 60 percent of operating nuclear power plants in the United States had or were suspected to have counterfeit or non-conforming parts.[29] Further, a 2024 study by Hobbs et al. concluded that counterfeit, fraudulent, and suspect items pose a significant threat to both nuclear safety and security, and that counterfeits can infiltrate the nuclear supply chain in various ways.[30]

Nuclear safety and nuclear security, while related, are different and have different implications for the nuclear supply chain. According to the IAEA, **nuclear safety** has the fundamental objective of "protect[ing] people and the environment from harmful effects of ionizing radiation."[31] Also according to the IAEA, **nuclear security** has the objective of "protect[ing] persons, property, society, and the environment from harmful consequences of a nuclear security event",[32] in which a nuclear security event results from an intentional or criminal act by a person or group.

In this paper, the following definition of counterfeit, fraudulent, and suspect items (CFSI) based on guidance published by the Organisation for Economic Co-operation and Development (OECD)[33] and the IAEA,[34] as described by Hobbs et al. (2024):[35]

- **Counterfeit:** Items or goods that are intentionally altered, created, or restored to imitate original products, without legal authorisation.

- **Fraudulent:** Items or goods that are intentionally misrepresented to be something they are not. In industry, fraudulent products are often those with incorrect identification or falsified certification.

- **Suspect:** Items or goods that are suspected to be non-genuine or not to meet certain standards, specifications, or technical requirements. There is often indication of this via methods like visual inspection, testing, or other disclosed information. These items could be knowingly or unknowingly counterfeit or fraudulent.

[26] OECD and EU Intellectual Property Office, "Trends in Trade in Counterfeit and Pirated Goods", 18 March 2019. Available at: https://www.oecd-ilibrary.org/trade/trends-in-trade-in-counterfeit-and-pirated-goods_g2g9f533-en.

[27] Further detailed information on risks to the nuclear supply chain is provided in the paper "Counterfeiting, Artificial Intelligence, and Supply Chains in the Nuclear Sector", by Christopher Hobbs and Zoha Naser, 2025.

[28] International Atomic Energy Agency, "Managing suspect and counterfeit items in the nuclear industry", IAEA-TECDOC-1169, 2000, p.2. Available at: https://www-pub.iaea.org/MTCD/Publications/PDF/te_1169_prn.pdf.

[29] "Nuclear Safety and Health: Counterfeit and Substandard Products are a Governmentwide Concern", Report to the Chairman, Subcommittee on Oversight and Investigations, Committee on Energy and Commerce, House of Representatives, October 1990, p. 3. Available at: https://www.nirs.org/wp-content/uploads/reactorwatch/counterfeitparts/counterfeitpartsgao10161990.pdf.

[30] Christopher Hobbs, Zoha Naser, Daniel Salisbury and Sarah Tzinieris, "Securing the Nuclear Supply Chain: A Handbook of Case Studies on Counterfeit, Fraudulent and Suspect Items", King's College London Centre for Science and Security Studies, 2024. Available at: https://www.kcl.ac.uk/research/nuclear-security-implications-of-counterfeit-fraudulent-and-suspect-items.

[31] International Atomic Energy Agency, "Fundamental Safety Principles", Safety Fundamentals No. SF-1, 2006, p. 4. Available at: https://www-pub.iaea.org/MTCD/Publications/PDF/Pub1273_web.pdf.

[32] International Atomic Energy Agency, "Objective and Essential Elements of a State's Nuclear Security Regime", Nuclear Security Series No. 20, 2013. Available at: https://www-pub.iaea.org/MTCD/Publications/PDF/Pub1590_web.pdf.

[33] OECD, "Regulatory Oversight of Non-Conforming, Counterfeit, Fraudulent and Suspect Items (NCFSI)", Nuclear Energy Agency, Committee on Nuclear Regulatory Activities, NEA/CNRA/R (2012)7, 15 February 2013. Available at: https://www.oecd-nea.org/upload/docs/application/pdf/2020-01/cnra-r2012-7.pdf.

[34] International Atomic Energy Agency, "Managing Counterfeit and Fraudulent Items in the Nuclear Industry", IAEA Nuclear Energy Series, No. NP-T-3.26, 2019. Available at: https://www.iaea.org/publications/11182/managing-counterfeit-and-fraudulent-items-in-the-nuclear-industry.

[35] Hobbs et.al., "Securing the Nuclear Supply Chain: A Handbook of Case Studies on Counterfeit, Fraudulent and Suspect Items", King's College London Centre for Science and Security Studies, 2024, p. 15. Available at: https://www.kcl.ac.uk/research/nuclear-security-implications-of-counterfeit-fraudulent-and-suspect-items.

Often, discussions of CFSI, such as those above, focus on the safety implications, as counterfeits can be prone to malfunction or failure, compromising the integrity of the system or infrastructure they are embedded in or responsible for, and causing a safety incident. However, CFSI are also a security concern. Malicious actors might seek to insert CFSI into the nuclear supply chain, for example, with the goal of providing a backdoor for hackers to access the facility computer systems or to create an intentional failure of critical safety or other facility systems. This type of concern is the focus of the current report, when discussing security of the nuclear supply chain.

There are vivid examples of this phenomenon outside the nuclear industry. As noted in Hobbs and Naser (2025):[36]

In early 2024 the risks posed by counterfeit items and the importance of supply chain security was vividly illustrated by a series of detonations in Lebanon involving exploding pagers and walkie-talkies, which caused over 30 deaths and hundreds of injuries.[37] It was claimed that Mossad agents had embedded several pounds of explosives into the devices, which were then branded with a legitimate company logo and sold to Hezbollah.[38] This incident served to highlight the potential for groups to weaponize the supply chain and create serious security risks.[39]

In the remainder of this report, the emerging risks and opportunities for the security of the nuclear supply chain will be discussed in detail as ever more sophisticated AI systems are integrated into daily life as well as into industrial settings, including in nuclear facilities and government and regulatory operations.

[36] Christopher Hobbs and Zoha Nasser, "Counterfeiting, Artificial Intelligence, and Supply Chains in the Nuclear Sector", 2025.

[37] Kathleen Magramo, Antoinette Radford, Adriene Vogt, Elise Hammond, Aditi Sangal and Matt Meyer, "Lebanon rocked by deadly walkie-talkie and pager attacks", CNN, 20 September 2024. Available at: https://edition.cnn.com/world/live-news/lebanon-explosions-hezbollah-israel-09-19-24-intl-hnk/index.html.

[38] Matt Murphy and Joe Tidy, "What we know about the Hezbollah device explosions", BBC News, 20 September 2024. Available at: https://www.bbc.co.uk/news/articles/cew12r5qe1ro;
Craig R. Heeren, Charles E. Westerhaus and Justin O. Kay, "Exploding Pagers: Supply Chain Vulnerability and Strategies to Reduce Risk", Faegre Drinker Biddle & Reath LLP, 18 October 2024. Available at : https://www.faegredrinker.com/en/insights/publications/2024/10/exploding-pagers-supply-chain-vulnerability-and-strategies-to-reduce-risk.

[39] Ari Hawkins and Joseph Gideon, "Middle East pager attacks ignite fear of supply chain warfare", Politico, 19 September 2024. Available at: https://www.politico.com/news/2024/09/19/pager-attacks-supply-chain-warfare-00180136.

Next to potential supply chain dependencies, the integration of AI systems into nuclear facilities could increase risks to cyber and nuclear security.

Artificial Intelligence and the Nuclear Supply Chain: Applications and Risks

While AI systems may bring significant benefits to the nuclear sector, the current report focuses on potential security risks associated with AI systems and the supply chain for nuclear facilities, as well as opportunities for AI systems to help address them. This is not to minimise the potential benefits of these systems, but to serve as a signpost that security considerations also need to play a part when discussing and thinking about AI systems in the nuclear sector.

The risks associated with AI systems and models, nuclear security, and the supply chain can be separated into two broad categories:[40]

- **Nuclear security risks may be associated with the supply chain for AI systems integrated into nuclear facilities:** The integration of AI systems into nuclear facilities and activities, including in administration, may introduce supply chain dependencies that may lead to new vulnerabilities which will need to be addressed.

[40] Activities related to the overall nuclear security architecture, such as radiation portal monitors at national borders, nuclear forensics activites, and efforts to secure major public events against attacks are also subject to these supply chain risks, depending on the extent to which AI is integrated. While not the subject of this specific report, many of the points made in the following sections may also be applicable in this case.

- **Publicly available commercial AI systems and models may increase the capabilities of malicious actors:** In particular, generative AI models may enhance a malicious actor's ability to carry out various attacks, including on the nuclear supply chain.

In the following two sub-sections, each of these two categories is discussed in more detail, first, security of the supply chain directly associated with AI systems used in the nuclear sector, and second, securing the nuclear supply chain against the increased capabilities AI could provide to a malicious actor. At the end of each section, several key takeaways are highlighted for the awareness of nuclear security stakeholders, particularly policy-makers, regulators, and international organisations.

Securing the Supply Chain for AI Systems used in the Nuclear Sector

AI systems are actively under consideration by researchers as well as the nuclear industry seeking to improve safety, security, and operations of nuclear facilities. In a conversation moderated by experts at the VCDNP and the Stimson Center in February 2025, regulators from the Canadian Nuclear Safety Commission (CNSC), the United Kingdom Office of Nuclear Regulation (UKONR), and the US Nuclear Regulatory Commission (USNRC) confirmed that the nuclear industry has approached them with various potential applications of AI systems for integration into nuclear facilities.[41] However, these benefits are accompanied by new security risks that will need to be proactively managed. Both the benefits of AI integrated into the nuclear sector as well as some of the risks are outlined in the sub-sections to follow.

Benefits of AI Systems for the Nuclear Sector

Industry experts and researchers envision improvements in efficiency, cost, maintenance, operations, knowledge transfer, administration, safety, and security via the use of various types of AI systems. AI systems could be used, and in some cases are already in use, to streamline administrative work, assist with development of safety cases, provide real-time assessments of the status of the reactor and assist in preventive maintenance (in conjunction with a range of smart sensors), and more effectively identify intruders in vital areas at facility perimeters.

From the perspective of a researcher, Huang et al. (2023) list a large number of potential applications in nuclear reactors.[42] Scientists at the US Idaho National Laboratory are actively looking into a range of applications for AI systems (including those based on generative AI models) to improve safety and automate labour-intensive tasks, some of which are already in use by utilities.[43] Moreover, in 2024, a major nuclear vendor, Westinghouse, launched a proprietary nuclear-specific AI system using a generative AI model with the goal of facilitating access for customers to "more than 100 years of proprietary industry innovation and knowledge".[44] In addition, the World Institute for Nuclear Security (WINS) hosted a two-part workshop in 2024 focusing on the rapidly expanding role of AI in strengthening the security of nuclear facilities.[45]

[41] Stimson Center and VCDNP, "Atoms and Algorithms: A View from the Regulator", 3 February 2025. View the web report and recording here: https://www.stimson.org/event/atoms-and-algorithms-a-view-from-the-regulator/.

[42] Qingyu Huang et al., "A review of the application of artificial intelligence to nuclear reactors: Where we are and what's next", Heliyon, Volume 9, Issue 3, 2023. Available at: https://doi.org/10.1016/j.heliyon.2023.e13883.

[43] Addison Arave, "Artificial intelligence in nuclear: How computer and data scientists are enhancing the industry", Idaho National Laboratory, 14 August 2024. Available at: https://inl.gov/feature-story/artificial-intelligence-in-nuclear-how-computer-and-data-scientists-are-enhancing-the-industry/.

[44] Westinghouse Electric Company, "Westinghouse Unveils Pioneering Nuclear Genreative AI System", 4 September 2024. Available at: https://info.westinghousenuclear.com/news/westinghouse-unveils-pioneering-nuclear-generative-ai-system.

[45] World Institute for Nuclear Security, "WINS Virtual Workshop: Exploring the Role of Artificial Intelligence in Strengthening the Security of Nuclear Facilities", 10 – 11 December 2024. More details available at: https://www.wins.org/event/7901/wins-virtual-workshop%3A-exploring-the-role-of-artificial-intelligence-in-strengthening-the-security-of-nuclear-.

While there is ongoing discussion of the integration of AI into new nuclear builds, including in small modular reactor (SMR) designs, near-term use cases are also being developed for currently operating nuclear reactors around the world, as supported by a number of potential use cases collected by the IAEA International Network on Innovation to Support Operating Nuclear Power Plants.[46] While many outside the nuclear sector have the impression that current nuclear facilities are still running on analogue technology from the 1970s and 1980s, this is no longer the reality. The first fully digital instrumentation and control system (I&C) was integrated into Japan's Kashiwazaki-Kariwa-6 reactor nearly 30 years ago, in 1996. Since then, digital technology has become increasingly widespread in operating reactors, and approximately 40 percent of operating reactors currently incorporate digital I&C systems.[47] The integration of AI systems into facility operations is likely to follow in the coming decades.

AI systems are already being used in other parts of the energy sector, from predicting weather patterns for renewable energy farms to improving the efficiency of oil and gas extraction. The nuclear industry is known to be historically risk-averse, leading to concerns about the risk to competitiveness of nuclear from other energy sectors that may be quicker to adopt these technologies. At the same time, there are, of course, gains to be had from learning from the mistakes and solutions of the early adopters.

Key Takeaways

- Industry experts and researchers envision improvements in efficiency, cost, maintenance, operations, knowledge transfer, administration, safety, and security via the use of AI systems in the nuclear sector.
- The speed of integration of AI systems into the nuclear sector could affect competitiveness of nuclear versus other energy sectors, particularly if slow, but lessons can be drawn from other sectors that are earlier adopters.

Nuclear Security Challenges for AI Systems used in the Nuclear Sector

In response to the increasing interest in AI systems discussed above, some national regulatory bodies and international organisations have initiated discussions of how to regulate nuclear facilities integrating AI systems, including their safety. For example, a recent joint report issued by the CNSC, UKONR, and USNRC[48] considers the potential for developing AI systems for nuclear applications, and a 2023 international technical meeting at the IAEA focused on the safety implications of the use of AI systems in nuclear power plants.[49] However, while security concerns are briefly mentioned in the joint report, and there has been some consideration by WINS on security applications of AI, potential nuclear security challenges associated with AI systems have not yet received the same amount of attention as the benefits of using AI systems in nuclear facilities and associated safety concerns.[50]

[46] IAEA, International Network on Innovation to Support Operating Nuclear Power Plants (ISOP). Use cases are available via the members' area. More details can be found here: https://nucleus.iaea.org/sites/connect/ISOPpublic/SitePages/Home.aspx.

[47] Sonal Patel, "The Big Picture: Nuclear and I&C," Power Magazine, 1 February 2013. Available at: https://www.powermag.com/the-big-picture-nuclear-ic/.

[48] CNSC, UKONR, USNRC, "Considerations for Developing Artificial Intelligence Systems in Nuclear Applications," September 2024. Available at: https://www.onr.org.uk/media/03zl1osf/canukus_trilateral_ai_principles_paper_2024_08_28-final.pdf.

[49] IAEA, "Technical Meeting on the Safety Implications of Use of Artificial Intelligence in Nuclear Power Plants," 16 – 20 October 2023. More details available at: https://www.iaea.org/events/evt2103061.

[50] World Institute for Nuclear Security, "WINS Virtual Workshop: Exploring the Role of Artificial Intelligence in Strengthening the Security of Nuclear Facilities," 10 – 11 December 2024. More details available at: https://www.wins.org/event/7901/wins-virtual-workshop%3A-exploring-the-role-of-artificial-intelligence-in-strengthening-the-security-of-nuclear-.

AI systems considered for use in nuclear facilities can be sourced from a commercial supplier or developed in-house, likely based on an existing commercial AI model (see Fig. 1). The decision to choose a commercial supplier or develop an in-house solution will involve a range of considerations, including infrastructure requirements, in-house expertise, costs, and likewise security considerations. Particularly if an open-source AI model is used, the organisation may have greater control and security over the model and its training data. However, AI system development is non-trivial, and the development approach will need to be balanced against the need to not only develop but also maintain an in-house solution, which could require significant resources. The use of the application, the sensitivity of training and operational data, including national laws on data sovereignty, may likewise influence AI system development options.

The challenges associated with securing AI systems have several unique or “scaled up” aspects compared with other digital systems. These aspects need to be analysed for each AI system being considered, for example, ensuring that:

- The training data that was used for machine learning[51] of the AI model was not tampered with by a malicious actor (i.e., to enable data poisoning attacks).[52]
- The output of the AI system is reliable in a broad variety of situations that are likely to be encountered, or are unlikely to be encountered but could have significant consequences if an inaccurate answer is provided.[53]
- The input data for the AI system has not been tampered with, for example, the facility data being provided to an AI system is not manipulated such that the model cannot interpret the data correctly.[54]
- Critical decision-making is done by humans, not directly by the AI system.
- Human-AI system interactions are taken into account, including ways in which the human could be influenced by the AI system.
- Cyber security of computing resources and data storage used for AI systems is robust, particularly if networked or cloud resources are needed.[55]

While these challenges are not unique to nuclear facilities, overcoming them reliably is essential for applying them in nuclear facilities or other critical infrastructure where an error has the potential for catastrophic consequences. The cost, efficiency, and other benefits of a given AI system need to be balanced with the security risks associated with the adoption of that system.

Several of the challenges listed above are relevant to procurement processes and security of the supply chain. First, depending on how the AI system was procured, it needs to be established that the training data for the AI model was not tampered with, and that the output of the system is reliable. Both aspects will be discussed in more detail in the next section.

[51] See page 5 and 6 for more information on machine learning and training data.

[52] A data poisoning attack is a type of cyber attack in which the training data used for machine learning algorithms in an AI system is intentionally corrupted or otherwise manipulated. This can be done in multiple ways. For example, via the injection of fabricated data points into the data set (data injection) and attacks that introduce subtle manipulations that cause anomalous behaviour if a trigger is encountered (backdoor attacks).

[53] This is similar for other types of software tools, as will be discussed in the next section; however, the explainability challenges related to AI systems (also known as “black box” behaviour) are frequently more complex than for other software tools.

[54] See page 6 and Fig. 1 for more information on input data.

[55] The high computing resource needs and data storage associated with many AI systems can necessitate the use off-site or cloud computing or storage.

Second, as noted previously, the supply chain for any computing and data storage resources planned for use with the AI system need to be considered and secured to the extent feasible, including for off-site or cloud resources. Off-site and cloud resources, while commonly used to enable AI systems with high computing and data needs, by their very nature, introduce further challenges, discussed later in this report.

Key Takeaways

- Novel aspects of the supply chain for AI systems integrated into nuclear facilities need to be analysed and associated nuclear security challenges addressed.
- The benefits of AI systems need to be balanced with the potential nuclear security risks associated with their adoption, including when deciding between in-house and commercial solutions.

Establishing Trustworthiness of AI Systems in Nuclear Applications

For AI systems to fulfil their promise to improve safety, security, and operations in nuclear facilities, operators and regulators need to have confidence that they will reliably function as expected in all foreseeable situations. Beyond safety and reliability, confidence is needed that the system has not been tampered with by a malicious actor.

Existing verification and validation processes to ensure that other types of digital systems or assets in critical applications function as expected may provide a valuable starting point for building such confidence in AI systems.[56] For all digital assets used in critical applications, it is necessary for the vendor and end-user to establish sufficient assurance that the asset functions as intended. While it is theoretically possible to track and explain each step of data manipulation in digital assets that do not rely on AI, in reality, this is also a complicated process, and the end-user often does not have access to all information needed to undertake it. While a wide range of validation and verification is done for such assets, it is generally not exhaustive. New methods and approaches may be necessary to establish sufficient "trust" in systems that integrate AI capabilities.

A typical approach is to establish the **trustworthiness** of the digital asset indirectly, through thorough testing and qualification.[57] More information on digital supply chain security in the nuclear sector, including cyber security approaches in procurement, can be found in the IAEA publication "Computer Security Approaches to Reduce Cyber Risks in the Nuclear Supply Chain".[58]

[56] Other types of digital assets currently in use in nuclear facilities include digital instrumentation and control systems and computer systems used in administration, among others.

[57] Trustworthiness is a property of a system that provides evidence that it is dependable when used and end-users have awareness of its capabilities during use. For more information see John D. Lee and Katrina A, See paper on "Trust in Automation:Designing for Appropriate Reliance", Human Factors and Ergonomics Society, Sage Journals, Volume 46, Issue 1, Spring 2004, p. 50 -80. Available at: https://journals.sagepub.com/doi/abs/10.1518/hfes.46.1.50_30392?journalCode=hfsa.

[58] IAEA, "Computer Security Approaches to Reduce Cyber Risks in the Nuclear Supply Chain", 2022. Available at: https://www-pub.iaea.org/MTCD/Publications/PDF/TDL-011web.pdf.

When discussing how to establish the trustworthiness of digital systems, three concepts are often used:

- **Explainability**, or the ability to determine the cause of system behaviour.
- **Interpretability**, or the ability for humans to understand the basis for the output.
- **Transparency**, or the ability to understand what has been learned by the tool and why it has been learned.

While these concepts also apply to establishing trustworthiness of AI systems, the use of AI can complicate establishing explainability, interpretability, and transparency versus other digital systems.

As discussed earlier in this report, in AI models (on which AI systems are built), machine learning algorithms are used to determine connections between data and create outputs without explicitly being given rules for making these connections by a human.[59] Even after the AI model has been trained and is ready for use, it may not be clear which connections have been made or what rules the AI model is applying to make them. This phenomenon is often referred to as the "black box problem".[60] Further, the training data used to train the AI model in the tool may also be unknown, or at best, only partially known.

Thus, in addition to the security considerations associated with all digital systems, for AI systems, it is also necessary to:

- Establish the reliability of the training data and ensure that the training data was not compromised.[61,62]
- Establish the reliability of the model and ensure that it was not compromised.
- Establish the trustworthiness of the AI system's output in a broad variety of situations that are likely to be encountered, or are unlikely to be encountered but could have significant consequences if an inaccurate answer is provided. This recognises the impossibility of envisioning and testing a complete set of possible scenarios.
- AI experts as well as cyber security experts need to be involved in the procurement and qualification process for any AI systems to be integrated into nuclear facilities.

The security of the supply chain of the AI system as well as the hardware the tool is operating on is also important to establish, as many layers deep into the supply chain as feasible given the sensitivity of the particular application in question.[63] Any cloud computing or storage resources involved in the use of the AI system need to be considered.

[59] See pages 5 and 6 and Figure 1 for more details on AI models.

[60] In this analogy, the AI system is an opaque black box into which inputs are fed and outputs are obtained, but whose inner workings are inscrutable.

[61] While not a supply chain concern, the test data may also be compromised by a malicious actor.

[62] As discussed in the previous sections, this may not be able to be directly established for commercial models or for applications built on commercial models.

[63] It is practical to employ a graded approach to cyber security in the supply chain, as discussed in the IAEA publication entitled "Computer Security Approaches to Reduce Cyber Risks in the Nuclear Supply Chain", cited previously.

Beyond testing and qualification, reliable AI control in the form of norms and technologies will help to increase confidence in AI systems in critical applications, such as nuclear facilities and activities. Technologies like blockchain[64] could increase the explainability and interpretability of AI decisions, leading to greater trustworthiness.[65] Further, an AI system could be used to detect anomalies in other AI systems, although the trustworthiness of this tool would need to be established as well. Validation of the trustworthiness of AI systems in practical applications is an ongoing area of research, which will likely provide lessons learned and advancements in the coming years that the nuclear sector can leverage.[66,67]

Even with the most thorough testing and qualifcation, assuring the perfect trustworthiness of an AI application is not possible, given the nature of machine learning. For high-consequence applications like nuclear facilities, AI systems are likely to have limited autonomy. In these settings, AI systems can assist or augment human decision-making, but humans must retain decision-making authority, accountability, and responsibility for actions taken by AI systems.

Key Takeaways

- Operators and regulators need to have confidence that AI systems will reliably function as expected in all foreseeable situations. Verification and validation processes used to ensure that other types of digital systems function as expected can provide a starting point for building this confidence.
- For AI systems, it is necessary to:
    - Establish that the training data was reliable and not compromised.
    - Establish the reliability of the model to ensure that it was not compromised.
    - Establish the trustworthiness of the AI system's output in a broad variety of situations that are likely to be encountered.
    - Involve AI and cyber security experts in the procurement and qualification process to address AI system-specific challenges.
- In high consequence applications, like nuclear facilities, humans must retain decision-making authority, accountability, and responsibility for actions taken by AI systems

[64] Blockchain technologies provide a ledger of records across a digital system, providing more transparency for a system's interactions.

[65] Scott Zoldi and Jordan T. Levine, "Using Blockchain to Build Customer Trust in AI", Harvard Business Review, 20 January 2025. Available at: https://hbr.org/2025/01/using-blockchain-to-build-customer-trust-in-ai.

[66] Lalli Myllyaho, Mikko Raatikainen, Tomi Männistö, and Jukka K. Nurminen, "Systematic literature review of validation methods for AI systems", Journal of Systems and Software, Vol. 191, November 2021. Available at: https://www.sciencedirect.com/science/article/pii/S0164121221001473.

[67] Georg Stettinger, Patrick Weissensteiner and Siddartha Khastgir, "Trustworthiness Assurance Assessment for High-Risk AI-Based Systems", IEEE Access, 8 February 2024. Available at: https://wrap.warwick.ac.uk/id/eprint/183638/1/Trustworthiness_Assurance_Assessment_for_High-Risk_AI-Based_Systems.pdf.

Limitations to Managing Risks Associated with the Supply Chain for AI Systems

In integrating an AI system into nuclear facility operations it is essential to establish 1) that the software in the AI systems, including the AI models they are built on, has not been tampered with by a malicious actor before arriving in the facility, and 2) the security of the supply chain for the computing and data storage resources planned for use with the AI system, including for off-site or cloud resources, has been considered. However, as with other types of security risk, the security risk associated with the supply chain for AI systems integrated into nuclear facilities can only be managed, not eliminated. Two key factors influence the management of this risk.

First, there is a limit to how precisely the full supply chain can be mapped, including for the types of computing hardware and AI systems that may be used in nuclear facilities. Current globalised supply chains involve many steps in which items are procured, manufactured, and combined, and can be highly complex. While there are commercial services emerging that specialise in using AI systems to undertake deep supply chain mapping, including for security purposes, these services are just beginning to appear and have not yet, to our knowledge, been applied extensively in the nuclear sector. With that in mind, they may become an increasingly valuable tool for supply chain security, both for the supply chain for AI systems integrated into the nuclear sector and the nuclear supply chain more broadly, in the coming years.

Second, while globalised supply chains often involve many companies and countries, there are a number of bottlenecks where only one supplier provides necessary parts for a broad variety of industries, for example, for the computer chips, GPUs, needed to run AI systems.[69] In these cases, there may be no alternative suppliers even if some security risk is identified. Other constructed components that support different sectors in managing industrial processes are also converging. Thus, limited options may be available to change suppliers for increased supply chain security, or even if suppliers of a component appear to be shifted, they may be relying on the same lower-level suppliers themselves for parts of that component. Even AI systems developed in-house (as discussed on page 13) are likely to rely on the same limited number of suppliers, both for the underlying AI models and for needed hardware.

Ultimately, security concerns will need to be balanced against the cost, efficiency, and other benefits that these tools bring. Some amount of risk in the supply chain for AI systems integrated into nuclear facilities will always need to be mitigated, and, where it cannot be mitigated, accepted. How much risk to accept is a decision that will depend heavily on the criticality of the use case for each particular AI system, as well as the extent of the security risk the tools are assessed to pose.

Key Takeaways

- There are limits to how thoroughly supply chain security risk can realistically be analysed, although this may change in the future. Supply chain risk for AI systems integrated into nuclear facilities can be managed but not eliminated.

- Only a limited number of commercial suppliers of some components may be available, particularly for AI systems and associated digital and hardware assets, and in-house solutions may also rely on these same suppliers.

[69] By the end of 2024, US company Nvidia held between 70 and 90 percent of the market share for AI chips. More details can be seen in Kitty Wheeler, "How Nvidia's AI Made It the World's Most Valuable Firm", Technology Magazine, 8 November 2024. Available at: https://technologymagazine.com/articles/how-nvidias-ai-made-it-the-worlds-most-valuable-firm

Risks Associated with Commercial AI Models used by Malicious Actors

In the previous section, the focus was on security of the supply chain for AI systems integrated into nuclear facilities, with a focus on the benefits that AI systems, when properly secured, could bring to these facilities. However, the increasingly broad availability of powerful and general-purpose AI models to the general public also has the potential to increase the capabilities of malicious actors who would seek to find and exploit vulnerabilities in the broader nuclear supply chain for their own purposes.

Both criminal and terrorist organisations are likely to use such tools to upgrade their attack methods. Criminal organisations are well-known to engage in cyber organised crime, including to exploit new online criminal markets.[70] Further, as noted in a 2021 joint report by the United Nations Office of Counter-Terrorism (UNCCT) and the United Nations Interregional Crime and Justice Research Institute (UNICRI), terrorists also might increasingly seek to use AI to further their aims: "as soon as AI becomes more widespread, the barriers to entry will be lowered by reducing the skills and technical expertise needed to employ it."[71]

Within the context of the nuclear supply chain, as described earlier in this paper, a major concern is the insertion of CFSI by a malicious actor. While most counterfeit items are primarily a danger due to their low quality, intentionally and maliciously inserted counterfeit items could also be used to, for example trigger a high-consequence safety event at a nuclear facility or enable an attack on a transport of nuclear material.

A range of commercially available AI models and tools available now and in the near future have the potential to significantly assist malicious actors in reaching such goals, if these risks are not mitigated. For example, commercial LLMs, which have been trained on vast amounts of publicly available information, may reach security-sensitive conclusions that could assist malicious actors in identifying vulnerabilities, including in the nuclear supply chain. With respect to cyber security and the nuclear supply chain, this increases an existing risk posed by search engines and information freely available on the internet, which a 2015 Chatham House report noted already "can readily identify critical infrastructure components that are connected to the internet" in nuclear facilities.[72]

With the current generation of LLMs, this risk is balanced by their significant tendency to reach false conclusions and assert them with certainty (commonly referred to as "hallucinations"), which could mislead a malicious actor choosing to use them.[73] However, trends have been towards decreasing hallucinations over the past few years, and methods are being actively sought by the AI community to address them.[74] As these rates decrease even further, these tools could become more effective not only for the general public but for those who would seek to misuse them.

[70] UNODC, "Criminal groups engaging in cyber organized crime", UNODC Teaching Module Series: Cybercrime, Module 13: Cyber Organized Crime. Available at: https://sherloc.unodc.org/cld/en/education/tertiary/cybercrime/module-13/key-issues/criminal-groups-engaging-in-cyber-organized-crime.html.

[71] UNCCT and UNICRI, "Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes", UN Office of Counter-Terrorism, 2021. Available at: https://unicri.it/News/Algorithms-Terrorism-Malicious-Use-Artificial-Intelligence-Terrorist-Purposes.

[72] Caroline Bayon, Roger Brunt and David Livingstone, "Cyber Security at Civil Nuclear Facilities: Understanding the Risks", Chatham House Report, September 2015, p. viii. Available at: https://www.chathamhouse.org/archive/cyber-security-civil-nuclear-facilities-understanding-risks.

[73] Hugging Face, a platform where the machine learning community collaborates on models, data sets, and applications, provides a list of hallucination rates for a broad range of LLMs. As of 02 April 2025, these rates ranged from around 1 percent up to around 30 percent. The latest rates are available at:
https://huggingface.co/spaces/vectara/Hallucination-evaluation-leaderboard.

[74] Billy Perrigo, "Scientists Develop New Algorithm to Spot AI 'Hallucinations'", TIME, 19 June 2024. Available at: https://time.com/6989928/ai-artificial-intelligence-hallucinations-prevent/.

Many of the large commercial players in this sphere are focused on the development of model guardrails to prevent LLMs from providing useful and detailed information on the development of chemical, biological, radiological, and nuclear weapons.[75] To some extent, this could be useful in securing information relevant to the overall nuclear supply chain, particularly related to dual-use items. A further step could be to consider guardrails more specific to critical national infrastructure, such as nuclear facilities.

However, the impact of such guardrails may be limited in the coming years, as other commercial players have adopted an open-source or open-weight approach to their LLMs (see page 6).[76] While open-source and open-weight approaches are intended to improve accessibility and speed up the development of new applications, the availability of the model weights also permits the removal of guardrails through a process called abliteration, allowing a malicious actor to use an uncensored version of the AI model.[77]

While the power of LLMs to process data has the potential for misuse as described above, these models may provide even more useful opportunities for malicious actors, particularly with respect to nuclear supply chain attacks. For example, the generation of highly convincing fake video and audio using AI, known colloquially as "deepfakes",[78] has regularly made the news recently.[79] Deepfakes are frequently discussed in the context of the spread of misinformation and disinformation, particularly on social media. However, the potential for deepfakes to enable criminal activity is a significant concern in many industries, particularly in the financial sector, as described in a 2024 article by Deloitte.[80] Deepfakes could also be valuable for counterfeiters and other malicious actors seeking to be perceived as a known trusted supplier. For example, as noted in Hobbs and Naser (2025), "[a] 2022 investigation by Insider identified a number of companies that used AI-generated images of fake employees to make their companies appear legitimate, including a cyber company contracted by the City of Austin police department."[81]

In seeking to be seen as a trusted supplier of components, malicious actors could also use AI systems to generate virtual proofs of trust, such as fake testing data, quality assurance, or certification documents. They could also use machine learning tools to generate more convincing phishing or other emails, gather information on targets of such emails, and tailor attacks. They could even use deepfakes to infiltrate companies or pose as trusted suppliers.[82] As generative AI models continue to advance, the opportunities provided to malicious actors are likely to become more sophisticated as well.

[75] For examples, see OpenAI Safety Update, 21 May 2024, available at: https://openai.com/index/openai-safety-update/ and Anthropic "Announcing our updated Responsible Scaling Policy," 15 October 2024, available at: https://www.anthropic.com/news/announcing-our-updated-responsible-scaling-policy.

[76] Aruna Kolluru, "Exploring the World of Open Source and Open Weights AI", Medium, 29 March 2024. Available at: https://medium.com/@aruna.kolluru/exploring-the-world-of-open-source-and-open-weights-ai-aa09707b69fc.

[77] Mate Valko, "AI Safety is Dead: People Stripping Guardrails from Every Open-Source Model (And Why Governments Can't Stop Them)", 27 January 2025. Available at: https://blog.namilink.com/open-weight-models-losing-their-guardrails-5d8e3652bb86.

[78] Deepfakes are typically produced using advanced machine learning algorithms such as Generative Adversarial Networks (GAN)(More on GAN can be found at Amazon Web Services, "What is a GAN?" available at: https://aws.amazon.com/what-is/gan/) or Diffusion Models (More on diffusion models can be found at IBM, "What are diffusion models", available at: https://www.ibm.com/think/topics/diffusion-models).

[79] For examples, see CNN Business's explanation, compilation of examples, quizzes and methods for detecting "When seeing is no longer believing", available at: https://edition.cnn.com/interactive/2019/01/business/pentagons-race-against-deepfakes/.

[80] Satish Lalchand, Val Srinivas, Brendan Maggiore, and Joshua Henderson, "Generative AI is expected to magnify the risk of deepfakes and other fraud in banking", Deloitte Center for Financial Services, 29 May 2024. Available at: https://www2.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-predictions/2024/deepfake-banking-fraud-risk-on-the-rise.html.

[81] Ryan Hogg and Evan Ratliff, "That company's 'About Us' page may be full of fake pictures of 'people' who don't actually exist", Insider, 16 October 2022. Available at: https://www.businessinsider.com/ai-generated-images-fake-staff-appearing-on-companies-websites-2022-10.

[82] An example of how damaging a deepfake can be is demonstrated by the hiring of a North Korean hacker by the security awareness training firm KnowBe4 in 2024. This hiring was facilitated by the use of deepfake technologies to create a convincing false identity, despite video interviews and background checks. For more details on this case see Avantika, "KonwBe4 Uncovers Fake Employee: How a North Korean Hacker Was Hired into the Team", The Cyber Express, 24 July 2024. Available at: https://thecyberexpress.com/knowbe4-fake-employee-north-korean-hacker/.

Malicious actors may also use LLMs to generate computer code with the goal of facilitating increasingly sophisticated cyberattacks on critical infrastructure like nuclear facilities. As noted by the same Chatham House report cited above, "[t]here is a pervading myth that nuclear facilities are 'air gapped' - or completely isolated from the public internet - and that this protects them from cyber attack."[83] However, cyberattacks are levelled at critical infrastructure on a regular basis, and at least four significant cyberattacks have occurred since 2003 at nuclear facilities.[84] Further, cyberattacks on nuclear facilities could have significant consequences: the IAEA notes that "[c]yber-attacks at nuclear facilities can contribute to causing physical damage to the facility and/or disabling its security or safety systems (i.e. sabotage), to obtaining unauthorized access to sensitive nuclear information, or to achieving unauthorized removal of nuclear material."[85]

However, only the minority of cyberattacks aimed at the nuclear sector, as in other sectors, involve actors who surpass a threshold of sophistication marking them as dangerous or requiring significant time and effort from cyber security staff to resolve. A major risk of AI models in the hands of malicious actors is the advantage such tools can give in planning and executing cyberattacks, including by helping identify vulnerabilities and the use of automated or partially autonomous "bots". This can increase significantly the number of attacks that pass the above threshold of sophistication. Even if it remains that few attacks are extremely dangerous, the workload of cyber security staff could be increased to the point at which it is more difficult to mount a defence, increasing the possibility of success of any given attack.

Whether through analysing freely available information, enabling deepfakes, or scaling up cyberattacks, the misuse of commercial AI models as described here primarily provides methods to "upgrade" existing modes of attack on nuclear facilities. At present, it is unlikely that the current generation of AI models will enable substantively new and innovative forms of attacks, but that the main risk to the nuclear sector, including the nuclear supply chain, will come from their ability to increase the capabilities of ordinary malicious actors. For the nuclear supply chain in particular, the risk that deepfakes could pose needs to be recognised and mitigated to the extent possible.

While at present the main risk is an increase in the severity of existing risks, as AI technology advances, the possibility of new and innovative attacks is likely to increase. Thus, it is essential for the nuclear security community to remain alert and continue to re-evaluate the potential threat that a malicious actor with access to commercial AI systems could pose, including to the nuclear supply chain.

[83] Caroline Bayon, Roger Brunt and David Livingstone, "Cyber Security at Civil Nuclear Facilities: Understanding the Risks", Chatham House Report, September 2015, p. viii. Available at: https://www.chathamhouse.org/archive/cyber-security-civil-nuclear-facilities-understanding-risks.

[84] Han et. al. (2022) cite four major cyberattacks that have occurred at nuclear facilities over the past two and a half decades, including a 2003 attack at Davis-Besse NPP in Ohio, United States, and a 2014 attack on the Republic of Korea's Korea Hydro and Nuclear Power plant operator. See more details in the article by Sang Min Han, Chanyoung Lee, and Poong Hyun Seong, "Estimating the frequency of cyber threats to nuclear power plants based on operating experience analysis", International Journal of Critical Infrastructure Protection, Volume 37, July 2022. Available at: https://www.sciencedirect.com/science/article/pii/S1874548222000142.

[85] IAEA, "Computer Security Techniques for Nuclear Facilities", IAEA Nuclear Security Series, No. 17-T (Rev. 1), 2021, p. 1. Available at: https://www-pub.iaea.org/MTCD/Publications/PDF/PUB1921_web.pdf.

Key Takeaways

- Commercial and open-source AI systems available today, particularly those built on generative AI models, could be used by malicious actors to enhance existing types of attacks on the nuclear supply chain.
- The use of commercial and open-source AI systems may make nuclear supply chain attacks by less sophisticated malicious actors more dangerous by increasing their capabilities.
- New and innovative modes of nuclear supply chain attacks may develop alongside AI technology advancements, and vigilance will be needed in the coming years.

Data Management and Data Security

The capacity of AI systems of all types to provide humans – including malicious actors – with new and powerful methods to process and draw conclusions from vast quantities of data means that data security is more essential to nuclear security than ever before. In recognition of this new reality, many States are increasingly concerned about where and how sensitive data, such as that from the nuclear sector, is stored. Thus, discussing nuclear security in the modern age, particularly with respect to AI systems, requires an understanding of the potential uses, misuses, and security of data as well as export control requirements relevant to data.

Potentially sensitive data, including routine procurement data and other administrative data that could provide insights into the nuclear supply chain, must be secured against unauthorised disclosure, and staff made aware of the sensitivity of this data. Further, the types of beneficial AI applications for the nuclear sector discussed earlier in this paper will use significant amounts of sensitive data, potentially supplied by smart sensors and other networked devices, the security of which will need to be carefully considered and managed.[86] Further, when storing and communicating this data, data sovereignty concerns as well as export controls will need to be considered.[87]

Caution should be extended to the use of commercial AI systems that may be already in use, even informally, by staff, and which may acquire sensitive data in the course of use. Notably, the use of commercial AI systems by employees may even be larger than managers are aware of, as is the case outside of the nuclear sector. While many managers may believe that their employees are not using AI systems, a recent report by McKinsey stated: "business leaders underestimate how extensively their employees are using gen[erative] AI [tools]. C-suite leaders estimate that only 4 percent of employees use gen[erative] AI for at least 30 percent of their daily work, when in fact that percentage is three times greater, as self-reported by employees."[88]

[86] In cases where internal AI systems are used with proprietary data, one option for increased cyber security is to limit access to those AI systems and data to only those who need to use the tools and are properly trained to do so.

[87] As defined by IBM, data sovereignty is the concept that data is subject to the laws of the country or region where it was generated. See more at IBM, "What is data sovereignty?" https://www.ibm.com/think/topics/data-sovereignty#:~:text=Data%20sovereignty%3A%20Data%20that%20is,and%20requirements%20surrounding%20data%20residency.

[88] Hannah Mayer, Lareina Yee, Michael Chui, and Roger Roberts, "Superagency in the Workplace", McKinsey & Company, McKinsey Digital, 28 January 2025, Chapter 2. Available at: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work.

In general, but particularly when cloud storage is under consideration, it is essential for operators to understand the obligations on their data – such as data protection requirements, export controls, and third-party data use – the operationalisation of their data, as well as all national regulations applicable to the storage and use of this data. States and operators will need to carefully consider their options for data storage and management, in line with an increasing number of national laws and regulations regarding data storage and management. They will need to decide whether to use a commercial cloud service or on-site resources and assess the security of any sensitive nuclear facility data provided to a cloud service, including the legality of providing such data to the service, in particular with respect to data sovereignty.

The supply chain for digital assets used for cloud storage may also involve nuclear security risks, and the decision to use cloud storage also needs to consider the security of the supply chain of the company supplying these services, to the extent reasonable. This is of particular importance when AI systems are used in nuclear facilities as external computing resources and cloud storage are key elements of many AI models.

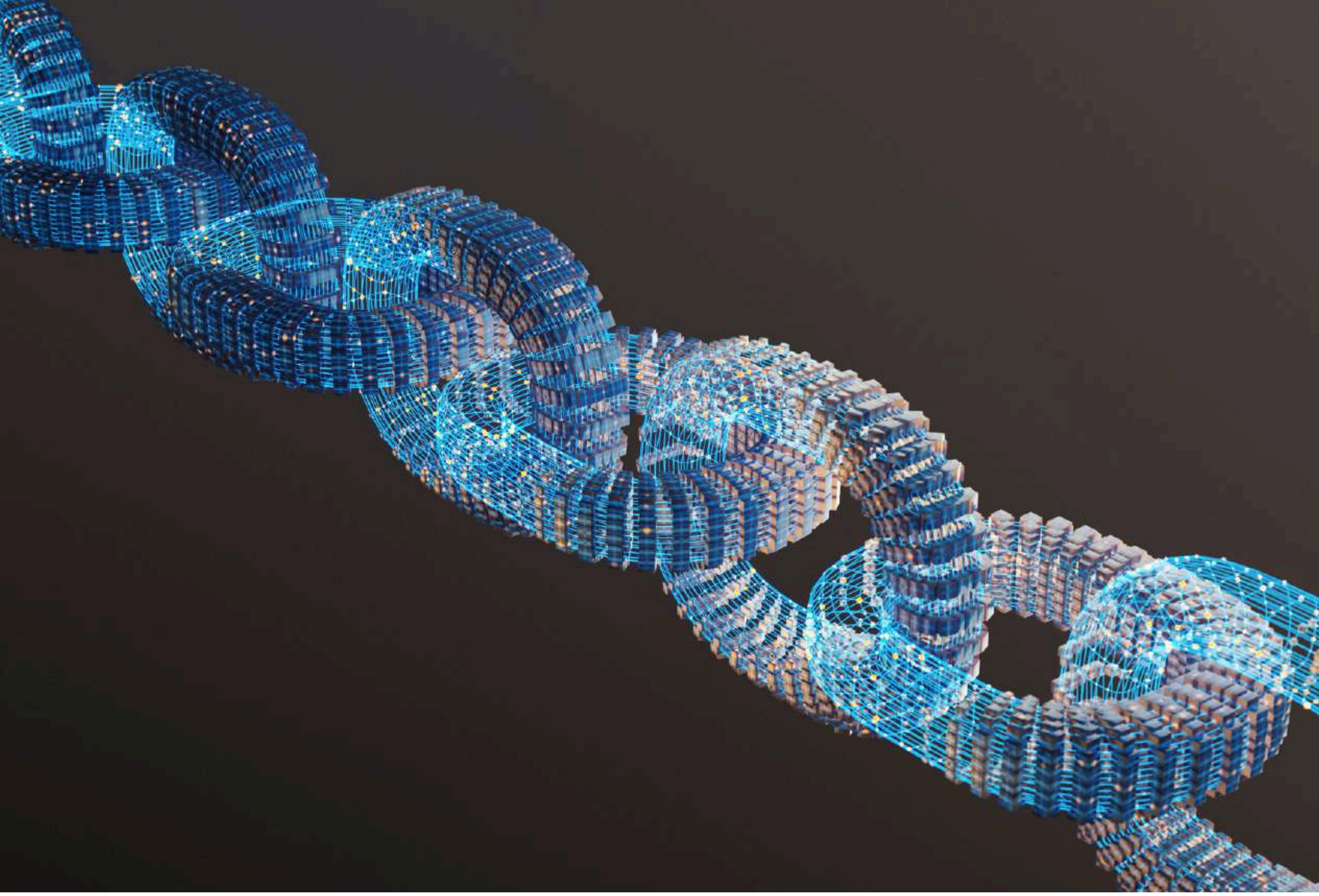
This situation could be managed via the establishment of focal points for data management and security in the national regulator and operators, in coordination with export control regulators.

Key Takeaways

- The integration of AI systems into nuclear facilities, including on the business side, may lead to the use and generation of significant amounts of potentially sensitive and export controlled data that needs to be managed.

- Operators should understand the obligations on and operationalisation of their data as well as all national regulations applicable to the storage and use of this data, particularly when cloud computing and storage resources are under consideration.

- The establishment of “data management and security focal points” in the national regulator as well as in operators in close coordination with export control personnel could be beneficial.



The integration of AI systems in the nuclear sector is closely tied to nuclear security considerations along the nuclear supply chain.

Strengthening Security of the Nuclear Supply Chain in the Age of Artificial Intelligence

The previous sections have described a number of benefits as well as nuclear security challenges that are already emerging as AI models become more advanced and widespread. They have also highlighted areas where the nuclear security community, including policy-makers, needs to be aware of how these potential future technological developments may pose challenges for nuclear security.

Over the last 50 years, the international community has established a strong and flexible international legal framework to ensure the security of nuclear materials, facilities, and activities against malicious non-State actors, consisting of conventions, bilateral and trilateral agreements, and non-binding international guidance on implementing nuclear security. Implementing those international rules and norms, States have established rules and regulations for nuclear security at the domestic level.

More recently, cyber security has become recognised as an essential part of the protection of nuclear materials, facilities, and activities, and has been integrated into this framework. While the nuclear security challenges posed by AI systems and models are deeply entwined with those posed by cyber security and often similar, they have unique aspects that need to be considered.

In the following sections, the description of the risk landscape at the intersection of AI, nuclear security, and the nuclear supply chain is used as a basis to suggest next steps in this area for national policymakers, diplomats, international organisations, and other nuclear security stakeholders. This includes the implementation of nuclear security obligations; protecting effectively against AI-enhanced threats; opportunities to strengthen security of the supply chain using AI; building capacity, training, and rasing awareness; and recognising and accounting for the technological pace of change.

Implementing International Nuclear Security Obligations

There are two sets of considerations related to international law at the nexus of AI and nuclear security. These are the need to ensure that:

- Considerations linked to the use of AI systems and models are adequately addressed in the international legal infrastructure for nuclear security.

- Nuclear security concerns are accounted for in ongoing discussions on the potential development of an international legal infrastructure for AI safety and governance.

Each of these will be discussed in turn, but the primary focus will be on the legal infrastructure for nuclear security, given that an infrastructure for AI safety and governance does not yet exist.

In the near term, the bulk of international discussion focused specifically on AI and nuclear security will involve how States implement their obligations under international conventions related to nuclear security, including related international guidance for its implementation. The nexus of AI and the legal framework for nuclear security is discussed in more detail in Nilsson (2025).[89]

The centrepiece of the international legal framework focused on the protection of nuclear materials and facilities is the Convention on the Physical Protection of Nuclear Material (CPPNM)[90] and its 2005 Amendment.[91] Although neither AI nor cyber security are mentioned directly in the CPPNM as amended, the nexus of AI and nuclear security is covered because the relevant provisions are sufficiently broad and flexible, including supply chain concerns, such as those discussed in this report.

Fundamental Principle G of the CPPNM as amended addresses threat as: “The State’s physical protection should be based on the State’s current evaluation of the threat.” This flexible formulation encompasses all technological tools and varying threats that are deemed relevant by a State, thus implicitly including AI-enhanced capabilities if they are evaluated to be part of a threat. Further, Fundamental Principle H of the CPPNM as amended addresses the need to take a graded approach to protecting vulnerabilities in the facility, accounting for this evaluation of the threat. For example, if the sabotage of an AI system integrated into a facility poses a risk of unacceptable radiological consequences, then the security of this system will need to be addressed. In the same broad sense, it could be argued that, security risks associated with the nuclear supply chain, including those posed by or exacerbated by AI models used by malicious actors, need to be addressed if they are evaluated to be included in the threat to a facility (i.e., as a tool used by malicious actors).

Finally, Fundamental Principle L of the CPPNM as amended addresses information confidentiality and State responsibility, specifically: “The State should establish requirements for protecting the confidentiality of information, the unauthorized disclosure of which could compromise the physical protection of nuclear material and nuclear facilities.”

[89] Anita Nilsson, “Artificial Intelligence, Nuclear Security and the International Legal Framework”, 2025.

[90] The full text of the CPPNM (INFCIRC/274/Rev.1) is available on the IAEA website at: https://www.iaea.org/sites/default/files/infcirc274r1.pdf.

[91] The full text of the 2005 Amendment to the CPPNM (INFCIRC/274/Rev.1/Mod. 1 (Corrected)) is available on the IAEA website at: https://www.iaea.org/sites/default/files/publications/documents/infcircs/1979/infcirc274r1m1c.pdf.

Another relevant legal instrument on nuclear security is the International Convention for the Suppression of Acts of Nuclear Terrorism (ICSANT),[92] which is focused on the criminalisation of acts and threats of nuclear terrorism. ICSANT also does not explicitly mention cyber security or AI, but the broad definition of offences in Article 2 would include the use of AI in aiding a person to commit such offences.[93]

As discussed above, data is essential to machine learning and in the development and use of an AI system. Data is required for training, fine-tuning, as an input and as an output. To make the AI system of practical use in a nuclear facility, this data will likely include operational data of which some could be considered sensitive. Data security in AI systems is one of the greatest concerns that potential users have expressed, according to a recent survey by Dudenhoeffer.[94] Guardrails need to be established and assurance activities conducted to protect against unintentional sensitive data exposure and, likewise, malicious data extraction or compromise.

Given the increasingly rapid pace of change of AI systems and models, it is not realistic to seek to include mention of all related technological advancements directly into international legal instruments. International legal instruments, particularly those that are legally binding on all parties that adhere to them, require years and sometimes decades of intense negotiations to arrive at final agreement. General text, such as that cited above from the CPPNM as amended, does not need to be changed every six months to a year as technology progresses, and also does not risk being outdated in five years. Further, a lack of direct mention in legal instruments, such as the CPPNM as amended, does not negate the applicability of States Parties' obligations to protect against a malicious actor using AI models to aid an attack or exploit vulnerabilities in AI systems or models.

While existing legal instruments for nuclear security may be sufficient in the face of technological change, the way in which they are implemented will unquestioningly need to reflect the reality of technological shifts. Depending on the type of facility and the national regulatory structure, a new threat assessment may need to be undertaken. This new threat assessment would account for the effects of AI systems on nuclear security, including ensuring security measures put in place by the operator meet either performance-based or prescriptive regulations developed by the State. The national regulator will also need to develop a plan to evaluate the application of the AI system and enforce relevant regulations.

Detailed international guidance on the robust implementation of national nuclear security obligations under legal instruments like the CPPNM as amended is provided in the IAEA Nuclear Security Series (NSS). This guidance, built by consensus from IAEA Member States, provides four levels of detail, from the high-level Fundamentals of Nuclear Security, through Recommendations and Implementing Guides, to Technical Guidance on specific detailed technical issues of nuclear security. Each of these levels is of value to national regulators and policy-makers, and many provide needed information for operators.

While the IAEA is actively working to meet States' needs for guidance on AI and nuclear security via international Technical Meetings and the drafting of informational documents, at present, there is little to no mention of AI systems and models in the NSS. Providing guidance in the NSS on the nuclear security aspects of AI systems and models, at an appropriate level of detail, could assist the international nuclear security community in the longer term in addressing their challenges and taking advantage of their benefits. However, NSS publications are developed by consensus, and given the amount of time it takes to reach consensus among many States, they often need to remain at a high level.

[92] The full text of ICSANT (2005), is available on the UNODC website at: https://www.unodc.org/uploads/icsant/documents/ICSANT_Text/English.pdf.

[93] An independently reasoning AI that would undertake an attack on a nuclear facility on its own initiative at this point remains in the future, and is not addressed here. However, this is a case in which AI advancements may need to be addressed directly in legal instruments, and discussion on related topics is underway in the context of lethal autonomous weapon systems (LAWS) and AI governance more broadly.

[94] Donald Dudenhoeffer, "Past, Present, and Future Applications of AI in the Nuclear Sector", 2025.

Given the length of this process compared to the pace of change of AI, in the short term, it is sensible for relevant organisations to develop informational publications on the nexus of AI and nuclear security, aimed at policy-makers, regulators, and operators. These types of efforts are in many cases already underway. To address the pace at which AI is developing, it might also be warranted to establish a standing working group to provide continuous guidance. This working group could be organised by an international organisation or another body.

Given limited budget and human resources, few States have the capacity to adequately research and provide guidance on the impact of rapidly changing emerging threats on nuclear security. International, freely available information and guidance for policy-makers, regulators, and operators like that provided by the IAEA and others is invaluable in building an international, cohesive response to emerging threats, such as those at the nexus of AI and nuclear security, ensuring fewer and more difficult targets for malicious actors. Further, such information could address the many potential beneficial uses of AI systems for nuclear security, such as facial recognition and perimeter monitoring, among others.

This section has focused on the legal infrastructure for nuclear security, given that an infrastructure for AI safety and governance does not yet exist. However, it is worth noting the ongoing discussions in this area centred at the United Nations in Geneva and in New York.[95] To this point, chemical, biological, radiological, and nuclear terrorism (CBRN) more broadly has been a relatively minor aspect of these discussions. For example, while the UN High-Level Advisory Body on Artificial Intelligence's September 2023 Governing AI for Humanity: Interim Report briefly mentioned CBRN risks specifically,[96] its September 2024 Final Report mentions only "malicious use by non-State actors".[97]

Even where CBRN concerns are discussed, the risks associated with AI and civilian nuclear facilities are often deprioritised compared to risks associated with AI and non-State actors seeking to develop chemical and biological weapons. While chemical and biological weapons risks are essential to address, the risks associated with AI systems and models and nuclear facilities should not be neglected. To this end, policy-makers knowledgeable in nuclear security are well placed to contribute, where possible, to national and regional discussions on AI safety and governance.

Key Takeaways

- The legal framework for nuclear security is broad and flexible enough to address threats related to AI systems, as it obligates States to protect against relevant threats and secure confidential information.

- States and operators could benefit from reliable guidance and information from international organisations and others on mitigating the impact of AI systems and models on nuclear security, and using AI systems to aid nuclear security.

- Discussions on nuclear security and AI need to be informed by and integrated into broader international discussions on AI governance.

[95] Regional, national, and sub-national discussions about AI governance safety are also ongoing, for example in the EU and in the US state of California.

[96] United Nations, "Interim Report: Governing AI for Humanity," December 2023. Available at: https://www.un.org/sites/un2.un.org/files/un_ai_advisory_body_governing_ai_for_humanity_interim_report.pdf.

[97] United Nations, "Governing AI for Humanity," September 2024, p.29. Available at: https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf.

Protecting Against Threats Involving AI Systems and Models

The use of commercial AI models (such as LLMs) by the general public is rapidly growing, and it is no longer possible to prevent malicious actors from acquiring and using them. Moreover, many parts of the AI community have emphasised the importance of sharing information and models to accelerate progress, resulting in a proliferation of open-source and open-weight models, as mentioned on page 6 of this report, increasing the risk of misuse of these tools. The nuclear security community has no choice but to adapt to this situation and its consequences. Further, the cost, efficiency, safety, and security benefits of integrating AI systems into nuclear facilities are significant enough that their use will likely increase in the future. The security of these AI systems will also need to be considered and ensured to the extent possible, including against a range of potentially AI-enabled cyberattacks.

It is not possible to protect against all threats, particularly threats that could change rapidly in response to developments in AI technology. Effectively protecting against such threats, whether in the nuclear supply chain or more broadly, will rely on maintaining awareness of the threat and emphasising the importance of resiliency in nuclear security systems.

Maintaining Awareness of the Threat

At the State level, the nuclear security implications of AI systems and models as discussed throughout this report should be integrated into the national threat assessment, to the extent possible. However, given that the revision cycle of the national threat assessment is unlikely to provide an opportunity to sufficiently reflect rapid changes in threat characteristics, further information will likely be needed for regulators and operators on this topic.

Different States will also have different useful perspectives on nuclear security and AI, given different experiences, use cases, and incidents. The international sharing of information on the intersection of AI, cyber security, and supply chain can be sensitive and challenging, but it is necessary to engage, given the potential value of such exchanges. Even if direct information on potential threats or past threats cannot be shared, providing information on procedures that work to combat or mitigate AI-related threats in the supply chain are invaluable. The organisation of relevant forums and information-sharing activities by international organisations, such as the IAEA, given its central role in nuclear security, could help to fill this gap.

Resiliency

Even if the State and the operator are sufficiently aware of the threat, nuclear security systems will need to be resilient against the rapidly evolving capabilities that AI systems and models could provide to a malicious actor. In the cyber security domain, it is recommended practice to assume that the adversary will be able to access sensitive systems and to proactively ensure that this access will not trigger unacceptable consequences. This can protect against the emergence of new capabilities that make a previously inaccessible system suddenly accessible. This case should be anticipated and system resilience developed to mitigate its impact.

A similar approach can be taken to the security of the nuclear supply chain, whether digital, physical, or related to AI systems, by assuming malicious items or CFSI will find their way into the supply chain and planning mitigation measures around this assumption. As a proactive measure, while cyber security and AI experts may not need to be directly involved in procurement, they should advise and provide expertise to procurement processes for AI systems. Cyber security considerations, including impact analysis, need to be integrated into all digital technology procurement.

This integration requires an understanding by procurement personnel of cyber security considerations and often an evaluation by cyber security experts to assess technical risks and potential impact of malicious exploitation of the system. The cyber security team needs to also maintain situational awareness related to the organisation's digital landscape and the risks emanating from AI-enhanced attacks. It is essential to raise awareness in States, national regulators, and operators on the risk that AI systems and models could provide opportunities to malicious actors seeking to infiltrate the nuclear supply chain.[98]

Key Takeaways

- Nuclear security implications of AI systems and models need to be considered as part of the national threat assessments, to the extent possible.
- The international sharing of information on the intersection of AI, cyber security, and the nuclear supply chain is needed and valuable, despite the potential sensitivity of such discussions.
- Nuclear security systems, and particularly cyber security systems, need to be resilient against AI-enhanced threats.
- To protect against AI-enhanced nuclear supply chain attacks and infiltration of the supply chain, cyber security and AI experts are well placed to provide expertise to procurement and qualification processes.

Opportunities for AI to Improve Security of the Nuclear Supply Chain

AI systems used by nuclear security staff at facilities can actively improve defence against the types of risks outlined in the previous sections of this report. In some cases, these tools are already being used and in other cases, they still need to be developed, but they will be essential in managing these risks.

First, there are opportunities for AI systems to improve operators' and suppliers' understanding of the full supply chain for AI systems integrated into nuclear facilities as well as for the broader nuclear supply chain. This type of technology, which uses an AI model to track myriad and branching supply chain links, already exists commercially and is likely to develop further in the coming years.

It is already being used in supply chain applications, for example, to identify the origin of cotton fibres in clothing manufacturing with the goal of tracking suppliers engaging in forced labour practices.[99] In the coming years, the ability to determine the source country and suppliers for key parts will be highly desirable, both to be able to predict supply chain difficulties and work around them, and to be able to identify potential vulnerabilities to a supply chain attack.

[98] Cyber security is a concern for a broad range of frequently-used industrial components and devices that are also used in nuclear facilities. Attacks on these systems could have significant consequences, for example, when a malware attack (referred to as "FrostyGoop") on Lviv-based energy facility Lviveploenergo interacted directly with the industrial control systems, leaving 600 households without heat in winter. More on this use case can be found in the article by Daryna Antoniuk, "FrostyGoop malware left 600 Ukrainian households without heat this winter", The Record, 23 July 2024, at: https://therecord.media/frostygoop-malware-ukraine-heat.

[99] Altana, "Illuminating the Xinjiang Forced Labor Ecosystem," 15 July 2022. Available at: https://altana.ai/resources/illuminating-the-xinjiang-forced-labor-ecosystem.

Further, some AI attack vectors, such as deepfakes and the falsified certificate examples discussed earlier in this report, will be most effectively identified as fakes in the coming years via specialised AI systems, as the markers signalling them as AI-generated become increasingly subtle. Cyber security professionals will increasingly need to be well acquainted with AI systems, which also may help to manage an increased volume of sophisticated, AI-enhanced attacks by malicious actors, by identifying and dealing quickly with falsified emails and attempts to enter secure parts of the facility network. Digital signatures, such as blockchain, can be used to track details of the supply chain of components, when used in conjunction with networked devices (Internet of Things or IoT) and other AI systems, as described in a recent article from the Technische Universität Wien.[100]

In addition, AI systems are likely to be useful for nuclear security and cyber security more broadly as they help to inform the design of defence-in-depth systems, security by design, and lead to advances in cyber-informed engineering.

Key Takeaway

- AI systems in the hands of security professionals will be increasingly useful in securing the nuclear supply chain, including by detecting deepfakes and other false credentials, mapping vulnerabilities in the supply chain, assisting cyber security professionals, and other tasks.

Capacity-Building, Training, and Awareness-Raising

Given the rapidly rising interest in AI systems in the nuclear sector along with the growing risk of AI-enhanced threats, more capacity will be needed to address the nexus of nuclear security and AI systems and models, both in operating organisations and at national regulators. As observed by Dudenhoffer (2025), "This is like the case of cyber security in the nuclear industry, where one of the largest challenges for organisations (both licensees and competent authorities) has been building and sustaining the human capital, i.e. individuals with the knowledge, understanding, and the experience to build and sustain effective governance."[101]

The need to provide tailored training for individuals throughout the nuclear sector, at operators, regulators, and other competent authorities, is likely to grow as AI becomes even more widespread. As Dudenhoffer (2025) further notes, "[a]mong existing staff, AI has the potential to dramatically alter current job functions. Continuous training and upskilling of the workforce at all levels may be necessary to prepare and enable the workforce."[102]

Cyber security professionals employed in the nuclear sector already have the responsibility for protecting against cyberattacks, including evolving, AI-enhanced cyber threats. Further, they have acquired in many cases the additional responsibility of ensuring security of new technology integration, such as AI systems in the nuclear sector. However, cyber security experts are not typically themselves experts in AI, which is a field that has exploded in relevance only in the last few years.

[100] TU Wien, "Revolutionizing Global Supply Chains: How AI, Blockchain, and IoT Enhance Efficiency and Resilience", 13 September 2024. Available at: https://www.tuwien.at/en/ace/news/news/global-supply-chains-ai-blockchain-iot-efficiency-resilience.

[101] Donald Dudenhoeffer, "Past, Present, and Future Applications of AI in the Nuclear Sector", 2025.

[102] Ibid.

Continuing training for these professionals, as well as the integration of AI-focused professionals into the nuclear security field, will be increasingly necessary. However, the reality that AI experts are currently in demand in many sectors is likely to complicate the acquisition of new talent to meet this need, and, as possible, capacity to manage the new challenges posed by AI may need to be built in the near term by training staff already in place.[103]

Relevant training specific to nuclear security and AI systems and models can be provided by organisations such as the IAEA, where available, and may also be delivered on shorter timeframes by organisations such as WINS,[104] who have already delivered two workshops focused on applications of AI to strengthen nuclear security in 2024.[105] Training on AI systems and models for cyber security professionals in the nuclear sector is of particular importance for assuring the security of the nuclear supply chain, and specific training at this nexus by such organisations would be useful. In addition, AI systems may provide new opportunities for delivering trainings, streamlining the needed up-skilling of this broad range of professionals.

In addition to building workforce capacity, awareness-raising on the potential uses of AI sytems and models in relation to the security of the nuclear supply chain is essential. For example, without awareness of how AI could be used to mimic a trusted supplier and provide seemingly valid data and certifications to support the sale of what is ultimately a counterfeit part, it is far more difficult to intercept this sort of attack. Without awareness of the risk of data poisoning and the need to build confidence in the AI outputs, a new and beneficial AI system for the facility could become a significant security risk.

This awareness needs to be built among a range of nuclear security stakeholders, notably within the nuclear industry, including operators and nuclear suppliers, within national regulators and other government competent authorities, and with key decision-makers. While this might appear to be too niche of a concern for the broader AI governance community and the commercial AI industry, awareness-raising work still remains to be done on the importance of the intersection of AI and nuclear security as a whole. Such awareness can be built through events focusing on the nexus of AI and nuclear supply chain security aimed at policy-makers and organisational leaders as well as via championing of this topic by knowledgeable individuals within organisations.

[103] This problem is not limited to the nuclear sector: according to a recent Forbes article, "more than half of the business leaders surveyed (55%) expressed worry about their ability to find sufficient talent to fill positions within the next year. Employers have aggressively sought out technical AI talent, resulting in a 323% increase in hiring over the past eight years." Ref: Jack Kelley, "AI-Skilled Workers are the New, Hot, In-Demand Professionals", Forbes, 1 August 2024. Available at: https://www.forbes.com/sites/jackkelly/2024/08/01/ai-skilled-workers-are-the-new-hot-in-demand-professionals/.

[104] WINS has several resources and training opportunities available. More information at: https://www.wins.org/.

[105] WINS, "WINS Virtual Workshop: Exploring the Role of Artificial Intelligence in Strengthening the Security of Nuclear Facilities", 10 – 11 December 2024. More details available at: https://www.wins.org/event/7901/wins-virtual-workshop%3A-exploring-the-role-of-artificial-intelligence-in-strengthening-the-security-of-nuclear-
WINS, "Introduction to the Role of Artificial Intelligence in Strengthening the Security of Nuclear Facilities", 6 to 8 February 2024. More details available at: https://www.wins.org/event/7877/introduction-to-the-role-of-artificial-intelligence-in-strengthening-the-security-of-nuclear-facilities.

Key Takeaways

- Capacity-building and training on nuclear security challenges and opportunities related to AI technology is needed in regulators, governments, and in the nuclear industry, particularly on the intersection of AI and security of the nuclear supply chain.
- Tailored training for certain stakeholders, such as cyber security professionals, could be of particular value.
- Awareness-raising on the potential uses of AI systems and models in relation to the security of the nuclear supply chain is essential among all stakeholders, and could be built through events aimed at policymakers and organisational leaders, and through championing of this topic by knowledgeable individuals within organisations.

Recognising the Pace of Change

Aside from ensuring that guidance and capacity are available to address the risks posed by AI systems for the nuclear sector, and particularly for the security of the nuclear supply chain, it is necessary to recognise the pace of change of AI technologies. The workshop that underlies this report was held in January 2025; this report was written in the first quarter of 2025. However, the change in this area between 2020 and 2025 was enormous, and the change expected between 2025 and 2030 could be even greater.

At the end of 2024 and beginning of 2025 alone, there have been significant shifts in the AI landscape. Smaller models such as DeepSeek are emerging, competing with the LLMs developed by US companies, such as OpenAI, Microsoft, Google, and Anthropic. Open-source and open-weight models, such as DeepSeek and Meta's Llama, are growing in popularity, despite the guardrail-related drawbacks discussed at the start of this report. AI agents, built for unpredictability and able to operate autonomously, may be the next leap forward. The intersection of robotics, networked devices, and AI are on the cusp of revolutionising the industrial sector. The digital transformation of industry, including the integration of AI will continue at a blistering pace. What we know today could change tomorrow.

Further, nuclear energy and thus nuclear facilities will be with us into the foreseeable future. There continues to be impetus not only to update and keep ageing facilities in service (and bring some back from decommissioning),[106] but to massively increase their contribution to the worldwide energy mix, to fight climate change.[107] This is exemplified by the COP28 pledge to triple nuclear capacity by 2050.[108]

Nuclear security must adapt to these changes and cannot risk being static against such a rapidly changing backdrop. The benefits of these new technologies will lead to pressure to use them, and at the same time, threat actors will gain new capabilities. Governments, regulators, and operators will need to keep pace.

[106] C. Mandler, "Three Mile Island nuclear plan will reopen to power Microsoft data centers," National Public Radio, 20 September 2024. Available at: https://www.npr.org/2024/09/20/nx-s1-5120581/three-mile-island-nuclear-power-plant-microsoft-ai#:~:text=The%20agreement%20will%20span%2020,of%20Constellation's%20former%20parent%20company.

[107] Pippa Stevens and Spencer Kimball, "Amazon, Google and Meta support tripiling nuclear power by 2050", CNBC, 12 March 2025. Available at: https://www.cnbc.com/2025/03/12/amazon-google-and-meta-support-tripling-nuclear-power-by-2050.html.

[108] US Department of Energy, "At COP28, Countries Launch Declaration to Triple Nuclear Energy Capacity by 2050, Recognizing the Key Role of Nuclear Energy in Reaching Net Zero", Energy.gov, 1 December 2023. Available at: https://www.energy.gov/articles/cop28-countries-launch-declaration-triple-nuclear-energy-capacity-2050-recognizing-key.

The nuclear industry is traditionally conservative with respect to the adoption of new technologies, and for good reason, given the potentially catastrophic consequences of a failure in a nuclear facility. However, given the impetus towards the beneficial use of AI in a range of sectors, it is unlikely that its beneficial use can or should be avoided in the nuclear sector. The nuclear industry does not, however, need to be a first adopter, and lessons can be drawn from other high-consequence sectors that are quicker to integrate AI systems, such as other parts of the energy sector or the air transportation sector.

Any guidance and information written on AI today, including this report, could be out of date within years, or even months, given the pace of change. Thus, policies and regulations for nuclear security and AI, including in the supply chain, need to be flexible and future-proof, and regulatory staff must be assigned to the task of staying current with the rapid shifts of technologies, both to understand the risks from malicious actors using AI and risks and mitigation strategies for AI integrated into facilities. Further, lessons learned from the past should not be forgotten and the implications of aging AI systems in facilities should be considered now rather than later.

Drafters of international guidance and national regulations will need to ensure that their guidance and regulations are flexible, agile, and forward-looking, but at the same time provide enough detailed guidance to ensure security. It will be a difficult balance to strike, and further research may be needed to determine the best way to go about it.

For that reason, it is essential for States and international bodies to continue to support proactive research on the nexus of AI and nuclear security, and in particular, AI and the security of the supply chain. The international community must stay ahead of the curve by funding forward-looking research to alert them of emerging concerns before they happen, to avoid falling into the trap of simply reacting to known threats that may be rapidly outdated.

This includes the development of guidance, information, and trainings by international organisations such as the IAEA, as well as quick-turn training by organisations such as WINS, which can help to keep staff at regulators and operators around the world informed in this shifting area.

Finally, a word with respect to the security of the nuclear supply chain and nuclear security as a whole: today's AI systems are just the beginning. This report has elaborated at length the risks and benefits associated with these AI systems, however, it is difficult to imagine realistically what may be coming next and what the "next big thing" might be. The nuclear security community must remain prepared and vigilant against new risks, but also recognise that the same next big thing might bring great benefits.

Key Takeaways

- Security of the nuclear supply chain in the age of AI needs to be aware, informed, and proactive.
- Guidance, policies, and regulations at the intersection of nuclear security and AI systems and models largely still need to be developed. They will need to be flexible and forward-looking, given the pace of change.
- Research is needed on the relationship between nuclear security and AI systems and models to alert of emerging concerns and to help prepare for the next big thing.

A Note of Thanks

The VCDNP wishes to thank all the individuals who generously gave their time to be interviewed for this research project. Your insights, experiences, and reflections helped shape the research direction, the workshop programme, and this publication.

The VCDNP is also deeply grateful to the participants of the workshop on "Nuclear Security in a Changing World: Exploring Evolving Supply Chain Risks related to Artificial Intelligence", held on 14 and 15 January 2025. Your active engagement, thoughtful contributions, and diverse approaches enriched this research endeavour and aided the production of this publication.

A special thanks go to the authors of the three papers commissioned over the course of this project, whose research provided a strong basis for the workshop discussions and this report. Your work has been central to furthering our understanding of the nexus between nuclear security, the supply chain, and artificial intelligence.

All this work would not have been possible without the support of Global Affairs Canada. The VCDNP appreciates your commitment to this project.

VCDNP
Vienna Center for Disarmament and Non-Proliferation
Austrian Federal Ministry for European and International Affairs

VCDNP

Vienna Center for Disarmament
and Non-Proliferation

The VCDNP is an international non-governmental organisation that conducts research, facilitates dialogue, and builds capacity on nuclear non-proliferation and disarmament.

🌐 vcdnp.org

✉ info@vcdnp.org

𝕏 @VCDNP