VCDNP Vienna Center for Disarmament
and Non-Proliferation

April 2025

Counterfeiting, Artificial Intelligence, and Supply Chains in the Nuclear Sector

Prof. Christopher Hobbs
Zoha Naser

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

Authors

Dr. Christopher Hobbs is Professor in Science and International Security within the War Studies Department at King's College London. He also serves as Director of the King's Institute for Applied Security Studies (KIASS). He has collaborated extensively with practitioners around the world in support of his research and has published widely on security, non-proliferation, intelligence, and verification issues, including the Oxford University Press Handbook of Nuclear Security (2023).

Zoha Naser is a Research Assistant based in the Centre for Science and Security (CSSS) at King's College London and the King's Institute for Applied Security Studies (KIASS). Her research interests include nuclear security, non-proliferation, and space power. She She also works on broader issues related to chemical, biological, nuclear and radiological (CBRN) materials.

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

While AI can significantly reduce counterfeit risks, it also introduces new attack routes that could jeopardise the integrity of high-stakes infrastructures like nuclear facilities.

Executive Summary

This paper explores how artificial intelligence (AI) is transforming supply chain management and the risks and opportunities this presents, with a focus on the phenomenon of counterfeiting. This is a perennial problem in many sectors, with studies estimating counterfeiting costs businesses globally hundreds of billions of dollars each year. The nuclear sector is not immune to this threat, and there exist numerous examples of where counterfeit items have penetrated supply chains and been installed at nuclear facilities, with serious implications for safety and security.

Here, AI models offer potentially unique solutions to addressing this and other risks through identifying counterfeit items during the assembly stage, mapping and assessing the full extent of organisations' supply chain, and through predicting potential disruptions. However, the use of AI for these and other purposes, if inappropriately implemented, can also create vulnerabilities and new threat vectors that could be exploited by malicious actors. Risks may be especially magnified in critical industries, such as nuclear.

The key findings of this paper are summarised briefly below:

Key Findings

- **Counterfeits present a significant risk to the nuclear sector** and one which has been arguably underestimated to date, with scope for new initiatives in this area.

- **Organisations across a range of sectors are utilising AI tools for a variety of supply chain management purposes** such as forecasting, inventory management, supplier risk identification, transport optimisation, and warehouse operations. Although, the uptake of these by the global nuclear industry appears so far to be limited.

- **Firms have reported that the use of AI models for supply chain management has improved the efficiency of operations**, strengthened resilience to disruptions, and provided new unique insights into different risks and challenges.

- **The unchecked use of AI models in supply chain management and other areas, however, poses several new and diverse challenges.** For example, malicious actors could infiltrate the AI supply chain by introducing unlicensed software or by manipulating the data sets upon which the AI tools are trained to negatively impact the model outputs.

- **In addition to malicious attacks, unaware employees may share sensitive information with certain third-party AI applications,** which could inadvertently become publicly accessible.

- **The "black box" nature of AI technologies also presents an intrinsic challenge for understanding how decisions are made**, with this serving to potentially limit adoption in highly regulated industries such as the nuclear sector. Although, this risk could be mitigated through careful consideration of "human-in-the-loop" designs for AI integration alongside human input.

- **AI models can also be exploited by adversaries to support counterfeiting efforts in a variety of ways**, from making their own operations appear more legitimate to generating fake testing data and certification documents.

- **Conversely, however, AI technologies are also being used in a variety of sectors to identify counterfeits at the manufacturing stage** and to illuminate organisations' full supply chains, including relationships between different suppliers enabling more complete risk assessments of potential counterfeiters.

A 2006 incident in China highlights the fatal risks of counterfeit piping in nuclear power plants.

Counterfeits in the Nuclear Sector: Risks, Actors, and Routes

Counterfeits are materials, parts, equipment, and products that appear to be genuine but do not meet the rigorous standards that legitimate items do.[1] For example, steel piping that has been fraudulently certified and sold for use in high-temperature environments, often has not undergone the heat treatment required to improve its toughness, which ensures the steel is resistant to cracking. As such, counterfeits can be prone to malfunction or failure, compromising the integrity of the system or infrastructure they are embedded in or responsible for. This was seen at a nuclear facility in China in 2006, when counterfeit steel pipes unexpectedly failed and ruptured when used to circulate steam in the power plant leading to the deaths of two workers.[2]

[1] Please note that in guidance by the International Atomic Energy Agency (IAEA) and other industry documents, counterfeit items are often collectively referred to as counterfeit, fraudulent and suspect items (CFSIs); 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, p.9. Available at: https://www.kcl.ac.uk/csss/assets/securing-the-nuclear-supply-chain-a-handbook-of-case-studies-on-counterfeit-fraudulent-and-suspect-items.pdf.

[2] NNB Generation Company, "Managing Counterfeit, Fraudulent and Suspect Items Guide Book," 2022, p.14–15. Available at: https://www.hinkleysupplychain.co.uk/wp-content/uploads/2022/09/HPC_CFSI-Handbook-min.pdf.

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.[3] 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.[4] This incident served to highlight the potential for groups to weaponise the supply chain and create serious security risks.[5]

In the nuclear context, there are many potential items that can and historically have been counterfeited, including, but not limited to, fasteners, circuit breakers, fire protection equipment, and reactor coolant pumps.[6] It is difficult to estimate the extent of counterfeiting in the global nuclear sector, although the limited studies that exist indicate it may be widespread. A 1990 study by the US General Accounting Office reported that over 60 percent of operating nuclear power plants in the US had or were suspected to have counterfeit or non-conforming parts.[7] A survey conducted by the authors in 2024 revealed that over 40 percent of nuclear industry respondents had experienced cases of attempted counterfeit infiltration in their organisations.[8]

There are many routes by which counterfeits can penetrate industrial supply chains (as illustrated in Figure 1). This could occur at the manufacturing stages or during their distribution, where intermediaries may intentionally mislabel items for sale at a higher price point. Customers at the end of the supply chain can also facilitate this through knowingly procuring counterfeit items from an illegitimate firm.

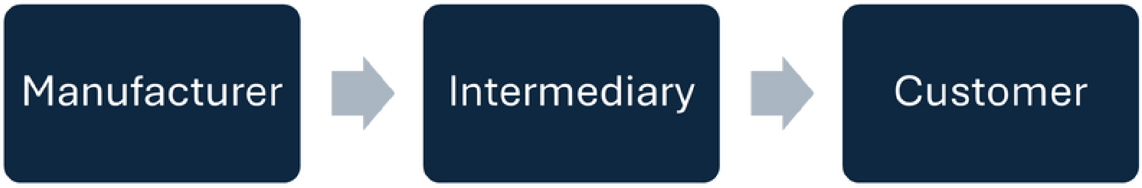

Fig. 1: Notional supply chain for tangible counterfeits (replicated from Hobbs et al. 2024)[9]

[3] 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.

[4] 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.

[5] 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.

[6] International Atomic Energy Agency (IAEA), “Managing Counterfeit and Fraudulent Items in the Nuclear Industry”, IAEA Nuclear Energy Series, No. NP-T-3.26, 2019, p.13-21. Available at: https://www.iaea.org/publications/11182/managing-counterfeit-and-fraudulent-items-in-the-nuclear-industry.

[7] US General Accounting Office, “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.

[8] Survey conducted by the authors, September 2024.

[9] 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, p.31. Available at: https://www.kcl.ac.uk/csss/assets/securing-the-nuclear-supply-chain-a-handbook-of-case-studies-on-counterfeit-fraudulent-and-suspect-items.pdf.

The routes and geographies that counterfeit items may travel are also complex, with illicit networks operating around the world and across multiple markets. Often, however, counterfeiters will target marketplaces or jurisdictions where a lax regulatory environment can allow them to operate with impunity. For example, transhipment hubs are utilised by actors to mask the origin of counterfeits and to establish distribution centres close to large international container ports.[10] Fraudulent items are also regularly moved through and manipulated within loosely regulated free trade zones (FTZs), where oversight can be lacking.[11]

Mitigating the risks of counterfeits requires proactive efforts by international organisations, national governments, and industry. Here, there are significant benefits to preventative approaches that identify and remove counterfeits from supply chains before they are installed at a facility. Chief amongst these are initiatives that include supply chain mapping and assessment, risk-informed supplier selection, and the development of programmes aimed at detecting the potential insertion of counterfeits, through audits and inspections.

This however can be a challenging task – for example, a study conducted by McKinsey in 2021 found that only two percent of companies surveyed were able to map their supply chain beyond the second tier.[12] It is here that AI can provide a useful solution: through supporting the mapping of supply chains; identifying how broader events might serve to alter procurement routes, potentially increasing the risk of counterfeits; and through the real-time monitoring of key elements of the supply chain. However, the use of AI in this area and broader supply chain management is not without risk. Concerns have been raised, as will be discussed later, over data protection, the reliability of AI models, their potential for manipulation, and how AI tools may be used by nefarious actors to further their counterfeiting goals.

[10] 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.

[11] Ibid.

[12] Knut Alicke, Ed Barriball, and Vera Trautwein, "How COVID-19 is reshaping supply chains", McKinsey & Company, 23 November 2021. Available at: https://www.mckinsey.com/capabilities/operations/our-insights/how-covid-19-is-reshaping-supply-chains.

Digital transformation through AI supports resilient and agile supply chain networks.

Artificial Intelligence in the Supply Chain

The use of AI to improve supply chain management and operations is not new, with research and practice in this area dating back decades.[13] However, increasing efforts have been made in recent years to integrate artificial intelligence and other digital technologies into the supply chain.[14] COVID-19 served as a wakeup call for many industries regarding the vulnerability of their supply chain to disturbances from unexpected events. This included the nuclear sector, where the pandemic caused supply chain disruptions, delaying the construction of nuclear power plants and complicating the distribution of radioisotopes.[15]

[14] Richard Wilding, “Supply Chain 4.0 – The Digital Era”, Cranfield School of Management, accessed 1 November 2024. Available at: https://blog.som.cranfield.ac.uk/execdev/supply-chain-4.0-the-digital-era.

[15] IAEA Director General, “The operation, safety and security of nuclear and radiation facilities and activities during the COVID-19 pandemic”, IAEA Board of Governors, GOV/INF/2020/8, 4 June 2020. Available at: https://www.iaea.org/sites/default/files/20/06/govinf2020-8.pdf.

There are many ways in which AI can be used to improve supply chain management and operations, including:

- **Forecasting:** This has been an active area for AI integration into the supply chain, with research demonstrating the value of machine learning and deep learning in monitoring real-time events, through analysing high-velocity changing data and adjusting forecasting immediately.[16] This has been utilised by firms in sectors, such as retail and electronics, both in terms of predicting potential disruptions to the supply chain and changes in consumer behaviour.[17]

- **Inventory management:** AI can be used to track inventory from manufacturers to warehouses and then to the point of sale. It can help firms better visualise how inventory is spread across their supply chain, tracking this in real time and identifying potential bottle necks and how they can be overcome, while also streamlining day-to-day operations and improving efficiency.[18]

- **Supplier risk identification:** As noted earlier, few companies have a detailed understanding of their extended supply chain. Here, AI tools can analyse large and disparate data sets to uncover information on suppliers that could present a potential risk. For example, if they are violating human rights or environmental standards.[19]

- **Transportation optimisation:** Issues at hubs and during border transportation can create unforeseen difficulties for firms and cause disruptions. At ports, for example, the process of getting ready for a shipment to dock can take a significant amount of time, creating add-on disruptions to shipping. Inefficient port layouts, limited handling capacity, and outdated infrastructure can create bottlenecks that lead to longer dwell times for shipping vessels.[20] With delays further magnified by disruptions due to events such as bad weather or unexpected maritime traffic. AI tools, combined with digital twins (a virtual replica of a real object or scenario), can be used to simulate these complex environments, helping port operators increase the efficiency of operations at these and other key points of the supply chain.[21] This, in turn, can minimise delays, reducing energy consumption and lowering carbon emissions.

- **Warehouse operations:** AI-driven robotics and automation technologies can be used to navigate, identify, select, and pack items for shipping to consumers, which can increase operational efficiencies and improve worker safety.

[16] For an example see Vitakesh Mani, Catarina Delgado, Benjamin T. Hazen, and Purvishkumar Patel, "Mitigating Supply Chain Risk via Sustainability Using Big Data Analytics: Evidence from the Manufacturing Supply Chain", Sustainability 9, no.4, 2017. Available at: https://doi.org/10.3390/su9040608.

[17] Rudrendu Kumar Paul and Bidyut Sarkar, "Transforming Supply Chain Management: The Impact of AI-Powered Demand Forecasting", Manufacturing & Logistics IT, 10 August 2023. Available at: https://www.logisticsit.com/articles/2023/08/04/transforming-supply-chain-management-the-impact-of-ai-powered-demand-forecasting; Adulyasak, Benomar, Chaouachi, Cohen, and Khern-am-nuai, "Using AI to detect panic buying and improve products distribution amid pandemic", AI and Society, Vol. 39, No. 1, pp. 2099-2128, 2024. Available at: https://doi.org/10.1007/s00146-023-01654-9.

[18] Adulyasak, Benomar, Chaouachi, Cohen, and Khern-am-nuai, "Using AI to detect panic buying and improve products distribution amid pandemic", AI and Society, Vol. 39, No. 1, pp. 2099-2128, 2024. Available at: https://doi.org/10.1007/s00146-023-01654-9.

[19] Oliver Telling, "Multinationals turn to generative AI to manage supply chains", Financial Times, 12 August 2023. Available at: https://www.ft.com/content/b7fafed2-9d00-49b0-a281-c1002b139865.

[20] Saranya Senguttuvan, "Why is my spot freight delayed? 6 reasons for ocean shipping delays (+ Maersk Spot FAQS)", Maersk, 9 October 2024. Available at: https://www.maersk.com/insights/digitalisation/2024/10/09/spot-freight-delays.

[21] Helene Hoffman, "Eye on the Future AI in Supply Chains and Logistics", Maersk, 2 July 2024. Available at: https://www.maersk.com/insights/digitalisation/2024/07/02/ai-in-logistics-and-supply-chains.

It is difficult to estimate the true extent to which companies are already utilising AI for the aforementioned purposes. Although, it is clear that many large global companies have sought to embrace this technology, either partnering with third-party providers and/or developing their own in-house tools. For example, companies such as Walmart, Tyson Foods, Koch Industries, Maersk, Siemens, and Unilever, are already using AI to plan and adapt for supply chain disruptions.[22]

Meanwhile large e-commerce companies such as Amazon and Alibaba are using AI to optimise warehouse operations.[23] Companies have also released information on the benefits of AI in this area, with Amazon, which has arguably the world's most complex supply chain and has invested heavily in AI, reporting that time taken to identify and store inventory was reduced by 75 percent thanks to the use of AI for warehouse management.[24]

However, despite the promise of AI and the benefits that are already being reaped, there are implementation challenges and potential risks that need to be considered and overcome for its wider adoption to be realised. According to a 2024 study by Cannas et al. organisations looking to use AI for supply chain management and operations, must overcome interrelated financial, organisational, strategic, and technological barriers.[25]

For instance, the use of AI can be hampered by a lack employee readiness and resistance to change, as well as the difficulty of upskilling staff on the use of AI technologies and the absence of sufficient qualified staff available in the marketplace. The use of AI also introduces a range of safety, security, and operational risks, that will need to be managed as part of its successful adoption and utilisation in critical sectors such as nuclear.

[22] Remko Van Hoek and Mary Lacity, "How Global Companies Use AI to Prevent Supply Chain Disruptions", Harvard Business Review, 21 November 2023. Available at: https://hbr.org/2023/11/how-global-companies-use-ai-to-prevent-supply-chain-disruptions.

[23] Amazon, "Amazon unveils the next generation of fulfilment centres powered by AI and 10 times more robotics", 9 October 2024. Available at: https://www.aboutamazon.com/news/operations/amazon-fulfillment-center-robotics-ai.

[24] Kris Van Cleave and Analisa Novak, "Amazon is using AI to deliver packages faster than ever this holiday season", CBS News, 27 November 2023. Available at: https://www.cbsnews.com/news/amazon-faster-deliveries-ai-holiday-season-cyber-monday-deals/.

[25] Violetta Giada Cannas, Maria Pia Ciano, Mattia Saltalamacchia, and Raffaele Secchi, "Artificial intelligence in supply chain and operations management: a multiple case study research", International Journal of Production Research, Vol. 62, No.9, 2024. Available at: https://doi.org/10.1080/00207543.2023.2232050.

AI systems introduce new attack vectors, including data sharing risks and vulnerabilities from nefarious actors.

Potential Vulnerabilities and New Attack Routes

While AI technology can and is being used to improve supply chain management and operations, its adoption in this and other areas poses several risks.

Data Qualification, Reliability, and Sharing

AI models are built on large data sets and how these are used to train an AI model has a substantial impact on the way it acts and interprets information. Should data be unreliable or even intentionally misleading, there could be significant consequences to the operation of the model. One way in which this could manifest is through the advent of "data poisoning" attacks.[26] These occur when a nefarious actor influences an AI model's training data by injecting false information during the training process, significantly altering the actions and results of the model. This could happen via a cyber hack, with nefarious actors exploiting vulnerabilities during the training process to degrade its performance.

[26] Bart Lenaerts-Bergmans, "Data Poisoning: The Exploitation of Generative AI", Crowdstrike, 20 March 2024. Available at: https://www.crowdstrike.com/en-us/cyber security-101/cyber attacks/data-poisoning/?srsltid=AfmBOop0Gtl4evDF53iLIHE27S2mOWUe0hNlRAtmcR0IdPBdG9Q3K9iP.

Due to the sheer amount of data being used in training a large model, the point at which malicious attackers infiltrated the data set can also be difficult to pinpoint, meaning data poisoning attacks could easily go undetected. Attacks could be carried out by internal actors ("insiders") who have intimate knowledge of the model, in what is known as a "white box attack". These are especially dangerous and more likely to succeed due to the insider's understanding of the AI and its vulnerabilities.[27]

A 2019 study by New York University found that AI models can be very susceptible to fraudulent or tampered training data, with potentially devastating consequences.[28] In this study, the research team were able to create a fake set of training data that encouraged autonomous vehicle software to register a stop sign as a speed limit sign, which if implemented could mean that vehicles fail to stop, resulting in an accident. This experiment demonstrated how a malicious actor could infiltrate a safety-critical AI system and compromise its integrity. In another project, a group of researchers were able to successfully train a model to hide the fact that it had been "poisoned" by unreliable data, leading it to act in a "deceptive" manner, avoiding detection during model safety training.[29]

These risks have been recognised in proposed legislation in this area, with Article 10 of the European Union's AI Act recognising the need for safeguards on data used to train AI in "high-risk systems".[30] In the context of supply chain management, "data poisoning" in AI models could disrupt ordering systems and distribution centres, wasting time and money.[31] For an AI model focused on evaluating supplier risks, this type of attack could lead to erroneous assessments, potentially resulting in orders being concluded with illicit suppliers that demonstrate patterns of behaviour suggestive of counterfeiting activity.

While some larger firms, such as Amazon or Microsoft, have their own in-house AI systems, the majority rely on third-party AI companies to provide these services and tools. "Third-party AI" typically refers to an AI system where at least one stage of the lifecycle – be it development, design, or deployment – occurs partially or wholly outside of the end user's control.[32] Companies will often enter into contracts with AI firms that specialise in a specific type of tool set, such as supply chain forecasting or regulation monitoring, and integrate them in their local systems by paying for access, acquisition, or licensing.[33] Outside of specific agreements, many firms also make use of widely available AI tools, such as ChatGPT and Midjourney, for more general supply chain management demands.[34] However, the use of some of these programmes has raised serious questions about data security and how sensitive data that leaves company servers may impact the firm.

[27] Eirini Anthi, Lowri Williams, Matilda Rhode, Pete Burnap, and Adam Wedgbury, "Adversarial attacks on machine learning cyber security defences in Industrial Control Systems", Journal of Information Security and Applications, Vol. 58, May 2021. Available at: https://doi.org/10.1016/j.jisa.2020.102717.

[28] Tianyu Gu, Kang Liu, Brendan Dolan-Gavitt, and Siddharth Garg, "BadNets: Evaluating Backdooring Attacks on Deep Neural Networks", IEEE Access, 7, 2019. Available at: https://ieeexplore.ieee.org/ielaam/6287639/8600701/8685687-aam.pdf.

[29] Evan Hubinger, Carson Denison, Jesse Mu, Mike Lambert, Nicholas Schiefer, and Ethan Perez, "Sleeper Agents: Training Deceptive LLMs That Persist Through Safety Training", arXiv preprint, Cornell University, January 2024. Available at: https://doi.org/10.48550/arXiv.2401.05566.

[30] European Commission, "Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts", 21 April 2021. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206.

[31] Usman Javed Butt, Osama Hussein, Krison Hasanaj, Khaled Shaalan, Bilal Hassan, and Haider al-Khateeb, "Predicting the Impact of Data Poisoning Attacks in Blockchain-Enabled Supply Chain Networks", Algorithms, Vol. 16, No.12, 2023. Available at: https://doi.org/10.3390/a16120549.

[32] This definition is lifted from: Rosamund Powell and Marion Oswald, "Assurance of Third-Party AI Systems for UK National Security", Centre for Emerging Technology and Security Research Report, The Alan Turing Institute, January 2024, p.8. Available at: https://cetas.turing.ac.uk/sites/default/files/2024-01/01.17.2024_assurance_report.pdf.

[33] Maria Diaz, "Third-party AI tools are responsible for 55% of AI failures in business", ZD Net, 26 September 2023. Available at: https://www.zdnet.com/article/third-party-ai-tools-are-responsible-for-55-of-ai-failures-in-business/.

[34] Steve Banker, "The Potential (And Peril) of ChatGPT in Supply Chain Applications", Forbes, 1 June 2023. Available at: https://www.forbes.com/sites/stevebanker/2023/06/01/the-power-and-peril-of-chatgpt-for-supply-chain-management/.

Use of third-party AI poses a potential risk in that companies may have little oversight or management of data security of the model and who in the third-party AI firm has access to it. This poses a greater risk when the data being inputted is sensitive or comes from high-security industries, like the nuclear sector. A prominent example of this was seen in April 2023, when internal meeting documents and source code from Samsung was leaked after employees ran the data through ChatGPT.[35] Data used by chatbots like ChatGPT is retained and stored on large servers that become publicly accessible as training data for the model to learn from. As large language models (LLMs) generate responses from learned data, confidential material stored on the server could be inadvertently exposed in a question-answer prompt or be uncovered by an individual who gains access to the training data.[36]

This new avenue for sensitive data leakage was flagged as a significant risk that needed to be carefully managed in a joint 2024 report by the nuclear regulatory authorities of Canada, the United States, and the United Kingdom.[37] For example, if sensitive information regarding the routes, times, or type of material contained in nuclear shipments was shared with an AI model and then accessed by a nefarious actor, this could significantly heighten the security threat.

In the context of counterfeiting, revealing certain inventory-related information, for example, relating to obsolete and hard-to-procure items, could provide illicit suppliers looking to penetrate a nuclear facility with useful knowledge with which to focus their counterfeiting efforts. As such, careful consideration will need to be given as to what type of data is shared with third-party AI models, and how third parties manage and secure that data. In addition, development of clear policies and training for staff is vital to help ensure that sensitive information is not accidently released.

Limited Visibility into AI Operations

An intrinsic challenge around the use of AI technologies, such as deep learning models, is their “black box” nature, where it is very difficult and sometimes near impossible to understand how decisions are made.[38] Scientists are still unclear as to how such systems process the data that is inputted and come to the conclusions that they do. Researchers are making some attempts to overcome the AI “black box”, and techniques, such as “red-teaming”, where a group simulates an attack on an AI model to test its strength and identify its weaknesses, are providing insight into how and why specific technologies behave in a certain way.[39] Red-teaming also allows researchers to identify how and why an AI may be making radical decisions or if it is being unduly influenced by dangerous or inappropriate content, helping to create better barriers to prevent this.[40] However, this research is still in its early stages and the varied and rapidly developing nature of AI technologies means that cracking the decision-making rationale of most AI models will likely take significant time and effort.

[35] Diaz, “Third-party AI tools are responsible for 55% of AI failures in business”.

[36] Siladitya Ray, “Samsung Bans ChatGPT Among Employees After Sensitive Code Leak”, Forbes, 2 May 2023. Available at : https://www.forbes.com/sites/siladityaray/2023/05/02/samsung-bans-chatgpt-and-other-chatbots-for-employees-after-sensitive-code-leak/.

[37] Canadian Nuclear Safety Commission, UK Office for Nuclear Regulation, and US Nuclear Regulatory Commission, “Considerations for Developing Artificial Intelligence Systems in Nuclear Applications”, September 2024, p.14. Available at: https://www.nrc.gov/docs/ML2424/ML24241A252.pdf.

[38] Lou Blouin, “AI’s mysterious ‘black box’ problem, explained”, University of Michigan-Dearborn News, 6 March 2023. Available at: https://umdearborn.edu/news/ais-mysterious-black-box-problem-explained.

[39] Billy Perrigo, “No One Truly Knows How AI Systems Work. A New Discovery Could Change That”, TIME Magazine, 21 May 2024. Available at: https://time.com/6980210/anthropic-interpretability-ai-safety-research/.

[40] Billy Perrigo, “The Scientists Breaking AI to Make It Safer”, TIME Magazine, 26 October 2023. Available at: https://time.com/6328851/scientists-training-ai-safety/.

This has clear implications for how certain AI models are used in highly regulated industries, such as the nuclear sector, where clear claims, arguments, and evidence need to be provided when making safety and security-related judgements. However, this does not necessarily preclude their use, rather careful thought will need to be given to the "human-in-the-loop" – the need for "human interaction, intervention, and judgment to control or change the outcome of an AI model", with the degree and type of intervention necessary assessed for different applications.[41] In the context of the supply chain, one would envision that assessments around the use of certain suppliers will always require human review.

Overreliance on AI Tools

Building on the previous sections, it is clear that companies must be careful in placing too much reliance on the outputs of AI models and establish appropriate checks and balances. As discussed above, their data sets can be "poisoned" or manipulated, making the AI model less effective. The importance of this point cannot be overstated, as studies have shown that if as little as 0.1 percent of training data is poisoned, an AI model's associated decision can be significantly altered.[42] Given that many AI supply chain solutions are designed for use by regular employees as opposed to AI experts, it is crucial that sufficient awareness-raising and training is given as to how their outputs could potentially be misleading, and that the model is actively monitored for signs that it has been compromised.

Another potential impact of overreliance in this area is how the use of AI may serve to limit opportunity for the development of human expertise.[43] For example, the inspection of items via visual means is one of the oldest and most common techniques for counterfeit detection.[44] Should this move more and more into the domain of AI models, then the ability of humans to inspect items and identify potential issues independent of technological aid could diminish.

To help overcome these issues, a gradual introduction of AI technologies while keeping humans involved in the loop (human-machine collaboration), as opposed to on the loop (human serves in a supervisory role only), is likely an important step towards integrating AI in sensitive industries, such as the nuclear sector. AI, by its very nature, is constantly learning and growing, meaning that there will always be the potential for it to make mistakes or require adjustment and fine tuning.

While AI continues to train and develop, human intervention will still be required to check decision-making and intervene in high-stakes operations to ensure that outputs are acceptable and accurate.[45] "Human-machine teaming", utilised in the defence sector, could be a viable strategy for the implementation of AI in security-critical industries like nuclear.[46] This concept recognises that humans and AI both have value to contribute to the decision-making process, and finding the right balance of human-machine integration can aid in effectively introducing advanced technologies to security-critical operations and environments.

[41] Xiao-Li Meng, "Data science and engineering with human in the loop, behind the loop and above the loop", HDSR, Issue 5.2, 27 April 2023.

[42] Ericka Chickowski, "AI and the software supply chain: AppSec just got way more complicated" ReversingLabs, 24 July 2023. Available at: https://www.reversinglabs.com/blog/ai-further-complicates-software-supply-chain-security.

[43] Robert Glenn Richey Jr., Soumyadeb Chowdhury, Beth Davis-Sramek, Mihalis Giannakis, and Yogesh K. Dwivedi, "Artificial intelligence in logistics and supply chain management: A primer and roadmap for research", Journal of Business Logistics, Vol. 44, No.4, 2023, pp.532-549.

[44] Gianmarco Baldini, Igor Nai Fovino, Riccardo Satta, Aris Tsois, and Enrico Checchi, "Survey of techniques for the fight against counterfeit goods and Intellectual Property Rights (IPR) infringement", JRC Technical Reports, European Commission, 2015. Available at: https://publications.jrc.ec.europa.eu/repository/bitstream/JRC98181/lbna27688enn.pdf.

[45] Andrew McAfee and Erik Brynjolfsson, "Big Data: The Management Revolution", Harvard Business Review, 1 October 2012. Available at: https://hbr.org/2012/10/big-data-the-management-revolutionhttps://hbr.org/2012/10/big-data-the-management-revolution.

[46] United Kingdom Ministry of Defence, "Joint Concept Note 1/18: Human-Machine Teaming", Development, Concepts and Doctrine Centre, May 2018. Available at : https://assets.publishing.service.gov.uk/media/5b02f398e5274a0d7fa9a7c0/20180517-concepts_uk_human_machine_teaming_jcn_1_18.pdf.

For the nuclear supply chain, "human-machine teaming" could involve active human oversight on AI-driven decisions for the procurement of critical parts, decisions made on whether to partner with certain suppliers, and the conduct of additional checks on associated testing and certification data. Human intervention in these tasks would ensure that the AI system is constantly being monitored for defects or mistakes, as well as checked for potential breaches. Here, a careful balancing act will be necessary to avoid some of the potential pitfalls in human-machine teaming. On the one hand, too much human intervention could lead to a system where the AI model cannot make a decision without human input, limiting the benefits of its use. On the other hand, too much AI autonomy and trust in the model could create a human overseer who becomes complacent and less able to identify potential issues.[47]

Counterfeits in the Artificial Intelligence Supply Chain

Another area of potential concern is the AI supply chain itself, and the potential for this to be undermined by the insertion of counterfeit technology and software. An AI model's "lifecycle" is extensive, involving actors both internal and external to the organisation where it is deployed. This includes the development stage, where training and data collection help build the foundations for the system, the deployment phase where the model is adapted and trialled, and the adaptation phase where the deployed model is routinely monitored for development purposes and augmented with local data (summarised in Figure 1).[48] What makes AI unique from other digital technologies is the rapid and ongoing pace of its development, with it constantly learning and adapting as it takes on an increasing amount of data in real time.[49] Researchers have noted that a single model can be made up of hundreds of data libraries and individual modules that are constantly absorbing new information, making it difficult to be fully aware of all aspects of a model.[50]

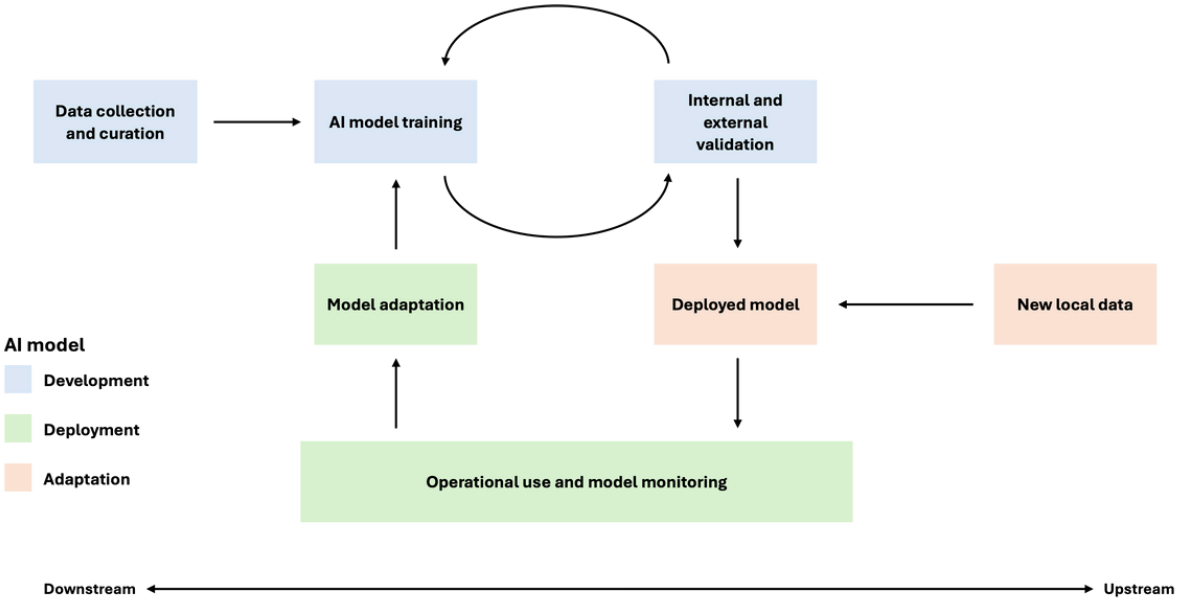

Fig. 2: Example of an AI system's lifecycle, adapted from Brown (2020)[51]

[47] Noah Greene, "Controlling the danger: managing the risks of AI-enabled nuclear systems", Outlook, NATO Defence College, No.7, November 2024. Available at: https://www.ndc.nato.int/download/downloads.php?icode=831.

[48] Ian Brown, "Allocating accountability in AI supply chains: a UK-centred regulatory perspective", Ada Lovelace Institute, June 2023. Available at: https://www.adalovelaceinstitute.org/wp-content/uploads/2023/06/Allocating-accountability-in-AI-supply-chains-June-2023.pdf.

[49] Ibid.

[50] Qixue Xiao, Kang Li, Deyue Zhang, and Weilin Xu, "Security Risks in Deep Learning Implementations", IEEE Security and Privacy Workshop, May 2018. Available at: https://ieeexplore.ieee.org/document/8424643.

[51] Ian Brown "Allocating accountability in AI supply chains"; Yipeng Hu et al., "The challenges of deploying artificial intelligence models in a rapidly evolving pandemic", Nature Machine Intelligence, Figure 1, 2020. Available at: https://www.nature.com/articles/s42256-020-0185-2/figures/1.

In this context, one of the biggest threats to the AI supply chain could come from its software components. A 2023 report on Cloud Native security conducted by cyber security firm Venafi found that 75 percent of security professionals cited their software supply chain as the biggest security blind spot at their organisation.[52] Cases of counterfeit and fraudulent software are extremely common, with the Business Software Alliance's 2018 report finding that 37 percent of all software used globally that year was unlicensed.[53] Counterfeit software can open up a variety of risks, and there are a number of cases involving malicious actors selling fake software to steal user data, including a high-profile 2023 case involving the Lazarus Group, a North Korean hacking ring.[54] In this example, the Lazarus Group shared malicious and fraudulent software packages on coding repositories that when downloaded by users, would install malware on their devices, and steal cryptocurrency.

Python, an integral software language for AI development, is often the subject of counterfeiting, including in a case in March 2024, where malicious actors "poisoned" a fake python programme to steal user data and hack programmes on their devices.[55] Though an attack of this nature does not seem to have yet targeted an AI system, many models are adapted from existing software applications, and hence, are potentially subject to the same vulnerabilities.[56]

Use of AI by Nefarious Actors

The use of generative AI has exploded in recent years, and while LLMs present a wide range of benefits in many areas, counterfeiters may also exploit such technology to enhance their own malicious activities. A common example of this is the use of AI to create false company profiles for non-existent employees. 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.[57] In another example, fake LinkedIn profiles with AI-generated images and profiles were used on the platform to promote sales.[58] While these examples may seem mundane and harmless, their impact should not be understated.

Faking profiles and qualifications could enable illicit networks to appear more legitimate, helping them sell counterfeit items. An example of such a case was seen with the fraudulent airline manufacturing company AOG Technics, which purported to be a legitimate manufacturer of parts for commercial aircraft, but instead sold fraudulent goods.[59] The company was ultimately publicly exposed in 2023, and an investigation by Bloomberg noticed that LinkedIn accounts for several executive officials at the company, including the chief quality assurance officer, were fake.

[52] Venafi, "The Impact of Machine Identities on the State of Cloud Native Security in 2023", 2023. Available at: https://venafi.com/lp/cloud-native-security-report-2023/#read.

[53] The Business Software Alliance, "Software Management: Security Imperative, Business Opportunity", BSA Global Software Survey, June 2018. Available at: https://gss.bsa.org/wp-content/uploads/2018/05/2018_BSA_GSS_Report_en.pdf.

[54] Yehuda Gelb, "Lazarus Group Launches First Open-Source Supply Chain Attacks Targeting Crypto Sector", Checkmarx Zero, 2 August 2023. Available at: https://zero.checkmarx.com/lazarus-group-launches-first-open-source-supply-chain-attacks-targeting-crypto-sector-cabc626e404e.

[55] Checkmarx Security Research Team, "Over 170k Users Affected by Attack Using Fake Python Infrastructure", 25 March 2024. Available at: https://checkmarx.com/blog/over-170k-users-affected-by-attack-using-fake-python-infrastructure/.

[56] Qixue Xiao et al., "Security Risks in Deep Learning Implementations", IEEE Security and Privacy Workshop, May 2018.

[57] 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.

[58] Shannon Bond, "That smiling LinkedIn profile face might be a computer-generated fake", NPR, 27 March 2022. Available at: https://www.npr.org/2022/03/27/1088140809/fake-linkedin-profiles.

[59] Siddharth Vikram Philip, Sabah Meddings, and Supriya Singh, "Bogus Supplier of Jet-Engine Parts May Have Faked Employees Too", Bloomberg, 8 September 2023. Available at : https://www.bloomberg.com/news/articles/2023-09-08/linkedin-profiles-expose-bogus-claims-at-fake-parts-supplier-to-jet-engines.

The implications of this case were significant, as AOG Technics sold parts to large, well-established airlines, such as United Airlines, American Airlines, and Virgin Australia Airlines.[60] Though this example does not appear to have involved the use of AI, it demonstrates that the risk is not unfounded, even when it comes to selling counterfeit items to critical industries.

Another potential use of AI for malicious purposes could be to generate fake testing data or quality assurance and certification documents. The nuclear industry has seen a number of these cases over the years, including a quality assurance certificate falsification scandal in South Korea between 2012 and 2014, a quality assurance data falsification at the Creusot Forge in France in 2014, and at the Sellafield MOX Plant in the UK in 2000.[61] While none of these cases involved AI, they demonstrated that fraudulent testing data, quality assurance, and certification have gone undetected at nuclear facilities before.

A counterfeiter could therefore feasibly use a tool like ChatGPT to generate a believable document aimed at presenting a counterfeit item as legitimate.[62] This concept is rooted in existing AI programmes that generate fake documents for cyber security purposes. These models, known as FORGE systems, create fake copies of important documents and populate them on an organisation's server, so that, should a cyber attack on the firm occur, would-be attackers would have to sort through a large number of fake documents to find the real version. This solution was specifically built for major corporations and government organisations who may be targeted due to their sensitive or security-critical information.[63] While the intended use of these tools is not malicious, it demonstrates that AI technology exists that can produce hyper-realistic technical documents.[64]

[60] Ian Molyneaux, "WestJet and Delta are latest airlines drawn into AOG Technics parts scandal", Aerotime, 4 October 2023. Available at: https://www.aerotime.aero/articles/westjet-delta-aog-technics.

[61] Benjamin Mallet, Chine Labbe, and Bate Felix, "French court probes forged documents case at Areva nuclear foundry", Reuters, 8 December 2016. Available at: https://www.reuters.com/article/world/french-court-probes-forged-documents-case-at-areva-nuclear-foundry-idUSL5N1E34XV/; BBC News, "Sellafield nuclear records faked", 17 February 2000. Available at: http://news.bbc.co.uk/1/hi/uk/646230.stm.

[62] Sarah Tait, "CDD and Generative AI – Risks of KYC Fraud", KPMG, July 2024. Available at: https://kpmg.com/uk/en/home/insights/2024/07/cdd-and-generative-ai-risks-of-kyc-fraud.html.

[63] Tanmoy Chakraborty, Sushil Jajodia, Jonathan Katz, Antonio Picariello, Giancarlo Sperli, and V.S. Subrahmanian, "FORGE: A Fake Online Repository Generation Engine for Cyber Deception", IEEE Transactions on Dependable and Secure Computing, Vol.18, No. 2, 2021. Available at: https://doi.org/10.1109/TDSC.2019.2898661.

[64] Chakraborty, Jajodia, Katz, Picariello, Sperli and Subrahmanian, "FORGE: A Fake Online Repository Generation Engine for Cyber Deception".

AI models play a key role in detecting and eliminating counterfeit components in critical supply chains.

Harnessing Artificial Intelligence in the Fight Againsts Counterfeiting

The previous sections have explored some of the risks and challenges of integrating AI models into supply chain management, and how this could enable the insertion of counterfeits into nuclear facilities if implemented and used inappropriately. However, there is a lot that AI can and already is bringing to the detection and removal of counterfeit items from supply chains.

As discussed earlier, AI models are being developed for applications that include supply chain forecasting and supplier risk identification. As noted in the aforementioned study by McKinsey, many companies have lacked the ability to map and assess their extended supply chain, given the considerable effort that this would take if done manually. In an effort to close this gap, a number of firms are developing AI tools to support supply chain visualisation and network analysis. For example, Altana, a US-based AI startup, uses public and private data to generate dynamic maps of an organisation's supply chain tiers.[65] This is done by drawing on customs declarations, shipping information, order descriptions, and other data points. The tool is used by government agencies, shipping companies, and other organisations to both show dependencies in the supply chain that might be vulnerable to disruptions from, for example, natural disasters or geopolitical events, and potentially high-risk sub-suppliers.[66]

[65] Altana, https://altana.ai/, website accessed on 16 November 2024.

[66] Eric Revell, "Altana using AI to map and analyze global supply chain with launch of next-gen Atlas," Fox Business, 8October 2023. Available at: https://www.foxbusiness.com/technology/altana-using-ai-map-analyze-global-supply-chain-with-launch-next-gen-atlas.

In another example, Exiger, another US-based AI company, uses machine learning models to offer a comprehensive assessment of supply chain risks, including cyber; financial; environmental, social and governance; reputational, criminal and regulatory; operational; foreign ownership control or influence; and product.[67] Exiger works with a wide range of firms in sectors, such as financial, healthcare, defence, energy, and the US nuclear sector. These and other companies are supporting the identification of counterfeits by enabling companies to understand where the components and materials that go into a product truly originate from and the risks associated with certain suppliers and geographies.

AI models have also been deployed within the supply chain to help visually identify defective parts and counterfeits during the manufacturing process. This includes programmes like Google Cloud Visual Inspection AI, which automatically detects items on the production line that are defective or do not meet certain standards and regulations.[68] Other examples include the Israeli startup Cybord, whose AI model uses visual and spectroscopic data on electronic components gleamed during assembly, assessing these against large data sets of known and counterfeit items.[69] This AI model authenticates manufacturer markings, while also identifying potential low-quality or aged components through examining defects caused by corrosion, cracks, and other means.[70]

Beyond the supply chain, firms like Clearway and Hikvision provide cameras or programmes that use AI to detect unauthorised worker behaviour that violate safety or environmental protocols. For example, their tools can identify those who are not wearing the correct personal protective equipment (PPE) needed for their worksite.[71] This could be expanded to detect other forms of unauthorised behaviour that might be suggestive of insertion of counterfeits or actions likely to result in defect products.

[67] Exiger, website accessed on 16 November 2024, https://www.exiger.com/.

[68] Mandeep Wariach and Thomas Reinbacher, "Visual Inspection AI: a purpose-built solution for faster, more accurate quality control", Google Cloud Blog, 22 June 2021. Available at: https://cloud.google.com/blog/products/ai-machine-learning/improve-manufacturing-quality-control-with-visual-inspection-ai.

[69] E. Afreen Banu, R. Priyanka, P. Thiruramanathan, T. Senthilnathan, Vithya V T, and K. Vinoth, "Robust AI-Enabled Electronic Components Authentication and Anti-Counterfeiting", 2024 Ninth International Conference on Science Technology Engineering and Mathematics.

[70] Cybord, https://cybord.ai/, website accessed on 16 November 2024.

[71] Clearway, "PPE Monitoring & Compliance", July 2023, https://www.clearway.co.uk/wp-content/uploads/2023/07/Clearway-PB-PPE-Monitoring-Compliance_July23.pdf; Hikvision, "Smart Detection Ahead of the Risks", accessed 08 November 2024, https://www.hikvision.com/uk/core-technologies/ai-analytics/ppe-detection/; Joseph Tsidulko, "Benefits of an AI Supply Chain", Oracle UK, 11 January 2024, https://www.oracle.com/uk/scm/ai-supply-chain/.

The successful integration of AI in supply chain management hinges on balancing innovation with rigorous controls to ensure safety and security.

Conclusions

This paper has sought to explore both the benefits and risks of integrating AI into supply chain management. On the one hand, AI has already been successfully adopted in industries and firms to help streamline supply chain operations and procurement activities. AI tools can help companies avoid interruptions and wider market disruptions as well as better visualise their supply chains and have more control over what items and products are procured. AI also offers significant potential benefits in the detection of counterfeits, through identifying specific fraudulent items and highlighting high-risk suppliers.

Conversely, the use of AI, if not carefully considered, may create several new and novel vulnerabilities, which could have significant implications for safety and security, especially in critical sectors, such as the nuclear industry. These include challenges around data protection and the black box nature of AI decision-making. The use of AI by nefarious actors may also complicate the detection of counterfeits, through the creation of seemingly trustworthy suppliers and genuine items. Consequently, it is essential for governments, regulators, and companies to carefully consider the implication of AI use when it comes to supply chain management and other areas of operation.

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