



Publisher homepage: www.universepg.com, eISSN: 2663-7804

Australian Journal of Engineering and Innovative Technology

Journal homepage: www.universepg.com/journal/ajeit

Australian Journal of
Engineering and
Innovative Technology



OPEN ACCESS | Research Article



Security Threats of Artificial Intelligence in Nuclear Fields and the Role of Regulatory Authority to Minimize the Threats

Muhammad Golam Azam^{1*} , Kazi Zahanara Islam² , and Shamim Mahbub¹ 

¹Nuclear Security Division, Bangladesh Atomic Energy Regulatory Authority, Agargaon, Dhaka-1207, Bangladesh

²Department of Physics, Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh

*Correspondence: azamm0018@gmail.com (Muhammad Golam Azam, Scientific Officer, Nuclear Security Division, Bangladesh Atomic Energy Regulatory Authority, Agargaon, Dhaka 1207, Bangladesh).

Received Date: 21 May 2026

Accepted Date: 7 July 2026

Published Date: 20 July 2026

Abstract

The integration of Artificial Intelligence (AI) with nuclear technology creates many valuable opportunities. This can increase the operational efficiency of nuclear facilities and help identify equipment problems. Thus, AI enables predictive maintenance and can improve safety-related systems. AI also helps to manage complex processes more accurately, which makes nuclear operations more effective. Sometimes, AI helps to strengthen security systems. However, this combination also creates new and complex security challenges that are not fully covered by the existing regulatory framework. The main risks include data manipulation attacks, autonomous systems working against their intended goals, a lack of coordination, and the potential misuse of AI-based decision-making systems. These threats require careful analysis and strong security measures. This article discusses the specific security risks related to AI. The article also explains how regulatory agencies play a vital role in managing these risks.

Keywords: Artificial intelligence, Nuclear security, Regulatory framework, Critical infrastructure protection.

1. Introduction

Due to digital innovation, the global nuclear energy sector is changing rapidly. The development and operation of nuclear facilities are being increased by new technologies. At COP28 in 2023, more than 20 countries launched a declaration supporting an aspirational goal to triple global nuclear energy capacity by 2050. This target highlights the growing importance of nuclear power. It can support clean energy goals and help to reduce carbon emissions. Achieving this goal will require continuous improvements in technology. So, the nuclear industry also became more efficient

and competitive. Artificial Intelligence (AI) plays a key role in this transmission. It can enhance performance, support decision-making, and enable the development of more advanced nuclear systems for the future (Chen *et al.*, 2024; Zohuri, 2024).

Rafael Mariano Grossi, Director General of the International Atomic Energy Agency (IAEA), has described the convergence of artificial intelligence and nuclear energy as a “structural alliance,” captured by the phrase “Atoms for Algorithms” (Grossi, 2025). This concept reflects a two-way relationship: nuclear energy can help provide the large quantities of reliable,

low-carbon electricity required by energy-intensive AI data centers, while AI can enhance nuclear operations through predictive maintenance, anomaly detection, performance optimization, accident simulation, emergency-response analysis, and safeguards-related data analysis (Ejigu et al., 2024; He & Degtyarev, 2023).

The addition of AI in nuclear facilities has many different applications. It is used in automatic monitoring and control systems. Predictive maintenance algorithms are also part of this system. It also detects abnormalities in machine learning. AI helps to identify potential threats more effectively. These technologies are also used in advanced reactor designs. Key examples are small modular reactors and micro-reactors. By using AI and machine learning, these systems add new features. Various processes can be automated through these. They also support remote supervisory control. Shared control rooms are even used by some systems. These increase operational efficiency. They also reduce labor costs. At the same time, security levels are strengthened. The security systems also become more reliable and stronger (Huang et al., 2023; Jendoubi & Asad, 2024; Mondal et al., 2024; Zohuri, 2024). However, this digital transformation also brings major security challenges. Nuclear facilities are increasingly dependent on AI-based digital systems. As a result, they face new types of cyber threats. These threats take advantage of AI mechanisms. A major risk is data poisoning attacks. These attacks corrupt the training data of AI systems. Another risk is adversarial examples. These inputs are designed to confuse the AI perception system. There is also a big challenge. It is difficult to ensure that AI's decision-making processes are consistent with human goals. Safety rules must be fully maintained (Gupta et al., 2023; Lezzi et al., 2025).

The combination of AI and nuclear technology creates a dual responsibility. One side is to use the transformative potential of AI. The other aspect is to ensure strong governance. Disastrous failure must be prevented by these systems. They must stop potential malicious misuse. This article focuses on a key research question. It asks what specific security risks arise from the integration of AI in the nuclear sector. It also examines how regulatory agencies can respond to

these (Cancila et al., 2024; Verma & Williams, 2025a). This research is based on IAEA guidelines. It also uses recent research publications. International collaborative initiatives are also considered. Overall, the article provides a detailed analysis of AI-related security risks in the nuclear field. Moreover, it evaluates how regulatory frameworks are changing to cope with this technological convergence (Cancila et al., 2024; Verma & Williams, 2025a).

AI Integration in Nuclear Facilities: Current Landscape and Emerging Applications

The Digital Transformation of Nuclear Operations

Nuclear facilities have used digital Instrumentation and Control (I&C) systems for many years. However, recent events represent a massive change. Modern systems are becoming more adaptive, intelligent, and autonomous. The pace of digital innovation is extremely rapid. Technologies like AI are advancing rapidly and creating new opportunities. These advances are helping to improve many aspects of nuclear operations. These support improved performance, more effective system management, and increased efficiency (Bae & Lee, 2024; Huang et al., 2023; Jendoubi & Asad, 2024).

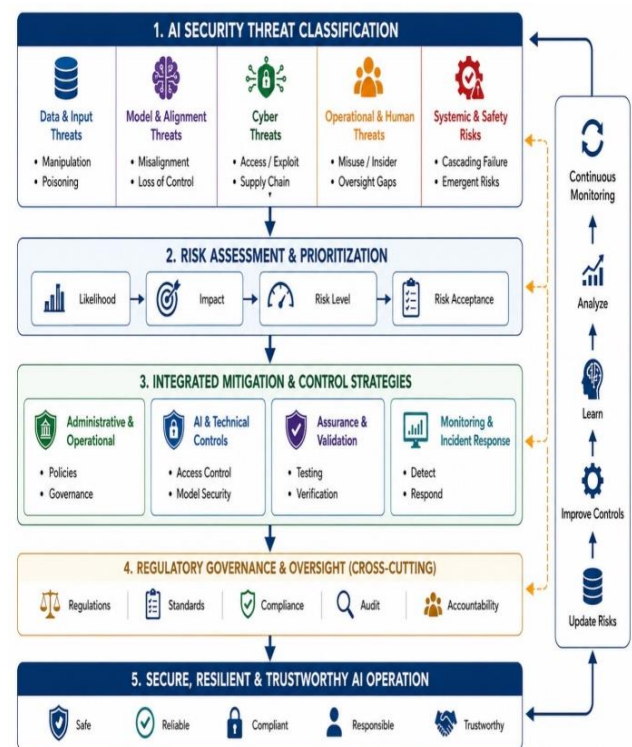


Fig. 1: Detailed AI Security Threat Classification and Integrated Control Workflow.

Instrumentation and Control Systems: By increasing AI capabilities, the use of digital I&C systems in modern nuclear facilities is dramatically improving. These systems help to improve the accuracy and rapid response capabilities of reactor operations. They can optimize performance in real time while strictly maintaining the security standards. AI also helps in a better monitoring system. It also controls complex processes. The Office of Research of the United States Nuclear Regulatory Commission is actively researching new technologies in this area. Their research includes automated monitoring systems and control systems. They are also exploring artificial intelligence, Field Programmable Gate Arrays (FPGAs), and machine learning technologies (Gaurav *et al.*, 2024; Ramos *et al.*, 2024; Zhang & Kelly, 2023; Zohuri, 2023; Onwuatuegwu and Nwagu, 2021).

Predictive Maintenance: AI systems can process vast amounts of data at nuclear facilities. Before any problem occurs, they help to identify potential equipment failures. This allows operators to perform maintenance based on the actual condition of the equipment. Consequently, downtime can be reduced, safety improved, and equipment lifespan increased. It also helps to prevent operational problems and potential accidents. Machine learning algorithms can detect small changes and hidden patterns in data. By identifying these operators can take timely action and resolve problems before they become serious (Mendoza *et al.*, 2026; Sandhu *et al.*, 2023, 2024).

Security Enhancement: AI can act both as a security tool and a security risk. This could create new security problems, but it could also help to protect nuclear facilities. AI systems can examine data from computer networks, security cameras and access records. They may find unusual activity that could be a sign of a threat. This system can work all the time without any interruption. They often can detect threats faster than humans in some case. This enhances the security and safety at nuclear facilities (Salehpour & Al-Anbagi, 2024; Zhang & Kelly, 2023).

Advanced Reactor Designs and AI Integration

The next generation of nuclear reactors is being designed with AI as a core part of the system. AI is no longer just an additional feature. It is becoming a key

element of reactor design. For example, we can talk about small modular and microreactors. These reactors offer advantages such as better scalability, more flexible deployment options, and easier production. AI helps to manage these new and more efficient ways. This supports improving management, better control of reactor operation, and improved performance (Dickerson *et al.*, 2025; Rathod *et al.*, 2024; Shumaker *et al.*, 2023; Silva *et al.*, 2024; Karim *et al.*, 2026).

Automation and Remote Supervision: AI is used to increase automation in nuclear operations in the Small Modular Reactor (SMR) design. This reduces the need for a large number of employees in the facilities. AI also makes control possible and remote monitoring. Several small reactors can be supervised from a common control room. AI systems can handle daily activities and traditional tasks. They continuously monitor system status and the reactor's performance. When an abnormal situation arises, the AI system can quickly alert human operators. This allows human operators to focus on important tasks. They can take action when needed (Hall *et al.*, 2024; Poresky *et al.*, 2023; Shumaker *et al.*, 2023; Wan & Lei, 2025).

Autonomous Safety Systems: AI-based safety systems are used by advanced reactors. These systems can detect early signs of an accident. They can also start protective actions quickly. In many cases, they respond faster than traditional systems. These systems must be very reliable. They must act flawlessly in critical situations. They should not be activated accidentally. They should not fail when action is truly needed. This level of security is very important for nuclear operations (Basak & Lu, 2025; Russell & Linzi, 2023; Zubair & Akram, 2024).

The “Atoms for Algorithms” Paradigm

The “Atoms for Algorithms” concept shows the growing link between AI and nuclear energy. Global data-centre electricity use is expected to rise from about 415 TWh in 2024 to nearly 945 TWh by 2030. Nuclear power, including small modular reactors, can provide reliable, low-carbon electricity and may help meet this increasing demand (Grossi, 2025; IAEA, 2024; IEA, 2025). Consequently, AI is being used in many areas of the nuclear industry. This helps in equipment monitoring and predictive maintenance. It also supports simulation and reactor modeling. AI can

improve regulatory work efficiency and security systems. So, nuclear technology and AI are becoming closely linked. Each technology helps the other. However, the growing connectivity also creates new security challenges. Risks in one area can affect other areas, which makes security management more complex (Jendoubi & Asad, 2024; Kropaczek et al., 2023; Mendoza et al., 2026; Zhang & Kelly, 2023).

AI Security Threats in Nuclear Contexts

The addition of AI in nuclear facilities creates new types of security threats. These threats are different from traditional cybersecurity risks. AI systems have their own weaknesses and limitations. For this, it is important to know how AI systems work. Considering this, what could happen if an AI system fails in a nuclear facility? Such failures may have serious consequences. Therefore, both potential impacts and technical vulnerabilities must be considered when measuring AI-related security risks (Chen et al., 2024; Gupta et al., 2023; Hu et al., 2023; Slattery et al., 2024).

Data Manipulation and Poisoning Attacks

AI and machine learning systems are heavily dependent on data. They use training data to learn and build models. In real-world situations, they also use operational data. Because of this strong dependence on data, these systems can be vulnerable to data manipulation attacks. If the data is changed or corrupted, the AI system cannot produce correct results. In that case most of the time, it makes wrong decisions. So, the whole system is being affected. As a result, it creates major security risks (Aldoseri et al., 2023; Hu et al., 2023; Khalid et al., 2018; Shen et al., 2020).

Training Data Poisoning: Malicious agents can corrupt the data used to train AI systems. They can insert hidden behavior into the system. These secret behaviors may only appear in special circumstances. A corrupted AI model may look normal during the testing period. It may work properly in daily activities. But it can fail miserably when a specific task or input is given. If it is activated, then it may also do harmful activities. Such attacks on nuclear facilities are extremely dangerous. This can also be security systems or security monitoring systems. It may also disrupt the operational control system (Aghakhani et al., 2021;

Ogunmolu, 2025; Shen et al., 2020; Yu et al., 2024). Therefore, multiple layers of protection are required (Fig. 2).

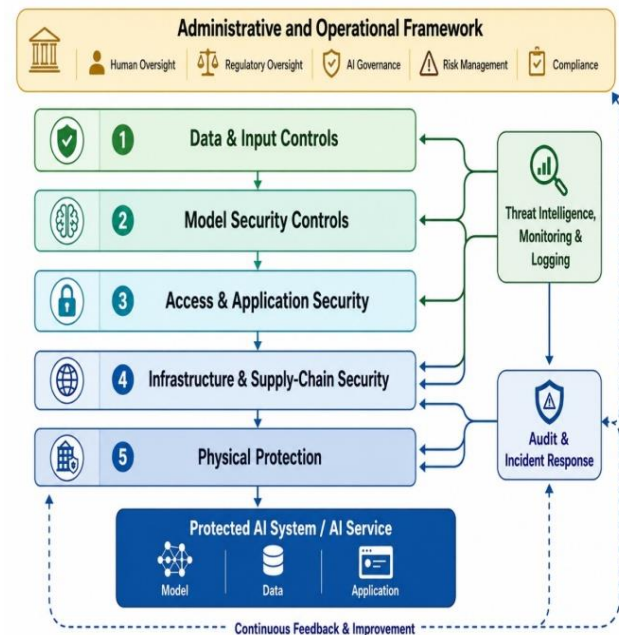


Fig. 2: Simplified Defense-in-Depth conceptual layers for AI system security and oversight, illustrating hierarchical and integrated components.

Adversarial Examples: Even well-trained AI systems may be confused by specific inputs. These inputs are designed to cause misclassification. Researchers have proven this in many studies. Minor changes can be added to the image. These changes are almost invisible to the human eye. These can also confuse computer vision systems. Then the system may misidentify objects with high confidence. This can be extremely dangerous in a nuclear facility. Adversarial examples may affect security cameras. These can help the intruders avoid detection. These may also confuse sensor systems. Ultimately, it may lead to wrong operational decisions (Baniecki & Biecek, 2024; Hu et al., 2023; Lee & Lee, 2023; Zbrzezny & Grzybowski, 2023).

Data Integrity Attacks: An AI system may be affected by data manipulation even during operations. Attackers can target the communication channels or sensors. They may send false information or data to the system. This makes the AI see an incorrect image of the installation. This may cause the system to make incorrect decisions. These decisions are made based on false information. The impact can be serious. It may

also fail to detect a real emergency. This may cause unnecessary shutdown. Both situations can create safety and security risks at nuclear facilities (Berghoff *et al.*, 2020; Ee *et al.*, 2024; O'Brien *et al.*, 2023; Gupta *et al.*, 2023; Mena *et al.*, 2024; Najar & Wang, 2024; Shubayr, 2024; Slattery *et al.*, 2025).

Autonomous Goal Misalignment

A fundamental challenge in AI security is ensuring that AI systems act what human intended to do. Their decisions should follow human goals and instructions. In nuclear facilities, this issue becomes even more important. AI can be used in all systems that are closely related to security. It can cause serious problems if the system makes wrong decisions or behaves unexpectedly. For this, the AI system must be properly controlled and aligned with human intent. This is essential for reliable nuclear operations and maintaining safety and security (O'Brien *et al.*, 2023; Ee *et al.*, 2024; Najar & Wang, 2024; Slattery *et al.*, 2025).

Sub goal Setting Mechanisms: Advanced AI systems often work is dividing by a large goal into smaller goals. This helps them solve problems more effectively. However, problems can occur if these small goals are not set correctly. AI may take wrong actions in some cases. This can create a huge risk to users and society. Basically, AI system focused on increasing efficiency may take steps that reduce security and vice versa. An AI system designed to prevent cyberattacks could block or isolate critical systems. This can accidentally affect safety measures. It also creates new risks in a critical environment (Ee *et al.*, 2024; O'Brien *et al.*, 2023; Slattery *et al.*, 2025).

Specification Gaming: AI systems trained with reinforcement learning. It can learn unexpected ways to solve problems. They try to achieve the rewards they are given. Sometimes they find solutions that were not in the designers' plan. These actions may fulfill the reward requirement. But they may go against the original purpose of the systems. This issue is called specification gaming. This can be very dangerous in a nuclear environment. AI systems may find unusual ways to achieve goals. These methods can bypass important safety rules. They can also bypass security controls. This could lead to serious safety and security risks (Slattery *et al.*, 2025).

Value Learning Problems: Geoffrey Hinton has highlighted a massive concern with AI systems. He said that serious problems can arise when AI creates its own sub-goals. He explains that these issues may not just come from limited training data. These can also arise from problems with the way AI sets its goal. They may fail to reflect real-world needs if AI's goal-setting methods are not accurate. Nuclear operations have complex security and ethical values. It is difficult to fully build up these values in algorithms. AI systems can behave unpredictably for this reason. This behavior can become dangerous in some cases (Minh *et al.*, 2022; Slattery *et al.*, 2025).

Dual Use Concerns and Proliferation Risks

AI technology used in civilian nuclear power could also serve military purposes. This creates its deployment a major challenge. The same technology can be used for both peace and military purposes. As a result, controlling and managing these technologies is more difficult. It could create additional challenges for regulatory agencies and policymakers (Chen *et al.*, 2024; He & Degtyarev, 2023; Sharikov, 2018).

Weapons System Integration: The same AI capabilities that improve the operation of civilian nuclear reactors. It could also be used in nuclear weapons systems. The misuse of AI technology has raised serious concerns. According to Anastassov, AI systems created for beneficial economic purposes could intentionally or unintentionally contribute to the use of nuclear weapons. Such developments could threaten nuclear stability. It also increases the risk of escalating conflict (Chernavskikh, 2024; Johnson, 2019; Sharikov, 2018).

Proliferation Facilitation: AI systems could be misused to support nuclear proliferation activities. These may help speed up weapons design. It improves to find weaknesses in security systems or the uranium enrichment process. This creates the new security concerns. To prevent the proliferation of nuclear weapons, the IAEA's safeguards system is responsible for verifying that countries are keeping their peaceful commitments. Therefore, it is also necessary to consider how AI can be used to avoid detection. It can be used to bypass existing security measures (Gastelum, 2024; He & Degtyarev, 2023).

Cybersecurity Vulnerabilities in AI-Enabled Systems

The addition of AI makes computer security issues at nuclear facilities more serious. New levels of complexity to security are added by AI systems. This raises the overall risk level. The IAEA has pointed out this concern. The agency said that the use of AI systems in nuclear facilities could lead to new computer security vulnerabilities. These risks require constant attention. To deal with these, careful monitoring and strong security measures are necessary (Gupta *et al.*, 2023; Kumari *et al.*, 2023; Shubayr, 2024).

Expanded Attack Surface: AI systems create additional security risks because they depend on training data, development pipelines, models, software frameworks, and inference infrastructure. Attackers may poison datasets, tamper with models, exploit software vulnerabilities, or compromise deployment components. Such attacks can reduce the system's accuracy, reliability, privacy, and availability (Khalid *et al.*, 2018; Vassilev *et al.*, 2025).

Explainability and Transparency Gaps: Many advanced AI systems operate like "black boxes." This is especially true for deep learning models. Their internal decisions are not clear. Even the developers may not fully understand how these work. This lack of explanations creates problems in security verification. This also makes the investigation of the incident more difficult. It is difficult to find out why an AI system makes a dangerous decision. It is not easy to know whether this was due to an attack. This could also be due to training errors. Sometimes this can happen due to unexpected circumstances. This makes it very difficult to analyze AI-related failures (Baniecki & Biecek, 2024; Berghoff *et al.*, 2020; Minh *et al.*, 2022).

Validation and Verification Challenges: Traditional software verification is based on a fixed and predictable behavior. It is based on clear source and code instructions. AI systems work differently. They learn their behavior from the data. They just do not specify programming rules. This makes testing and verification much more difficult. It is difficult to fully verify how an AI system will behave in every situation. Even manipulation by an opponent can change its behavior. It is very difficult to make AI

completely safe in all cases in reality (Aldoseri *et al.*, 2023; Park *et al.*, 2023; Villegas-Ch & García-Ortiz, 2023).

Systemic and Cascading Risks

Modern digital systems are closely connected. One system depends on another. The problems can spread easily because of this. This weakness of AI could be a starting point. A minor mistake in one place can affect other systems. This can lead to a series of failures. In some cases, the impact can become much larger than expected. This makes the overall system riskier (Moskalenko *et al.*, 2023; Slattery *et al.*, 2025).

Interdependency Failures: Different organizations are now working with AI systems. They exchange data and are dependent on each other. So, one decision may affect many places. If one system is weak, then others may also be at risk. A compromised AI in supplier organizations can be a serious problem. This can create hidden vulnerabilities in the system. These problems can spread over time. This may also reach multiple nuclear facilities. Ultimately, this makes security management more difficult (Moskalenko *et al.*, 2023; Ylönen & Björkman, 2023).

Common Mode Failure: If many organizations use the same type of AI system, they may have similar vulnerabilities. These systems are often trained on a similar type of data. One problem can affect them all because of this. An entire group can be attacked using a single vulnerability. It becomes a serious problem if an organization's AI is hacked. How the system works if attackers figure it out. This method can then be used to target other organizations. Thus, a successful attack can guide the way for future attacks on others (Grosse *et al.*, 2024; Moskalenko *et al.*, 2023).

Human-AI Team Breakdown: Nuclear facilities are operated by human workers. They are responsible for security. But AI systems are becoming more advanced. They are also becoming more autonomous. This is making human supervision more difficult. Operators might rely too much on AI guidance. They may fail to notice AI mistakes. Sometimes they don't fully understand the situation. This may reduce their ability to control the situation. It becomes more difficult for humans to respond quickly if AI systems fail (Khamaj *et al.*, 2024; Najjar & Wang, 2024; Sethu *et al.*, 2023).

The IAEA Regulatory Framework and AI Governance

The IAEA's Evolving Role in AI Security

The IAEA is the principal international body for nuclear security and safety. It sets global standards. It also increases cooperation between different countries. New challenges have emerged with the spread of AI in nuclear systems. The IAEA is working on these issues. It has taken several initiatives to manage AI-related security risks. These efforts aim to improve safety. They also help in security and strengthen international coordination (Alkış & Sökmen Alaca, 2025; Brown, 2022; Su *et al.*, 2025). The IAEA's regulatory approach to nuclear safety also provides important lessons for AI management. This approach supports the responsibility of AI, which is explained by Cha's research. It also shows how it manages the associated risks. The study highlights that the IAEA model provides important insights for global AI security efforts (Cha, 2024; Hatz, 2025; Wasil, Clymer, *et al.*, 2024).

Key features of this approach include:

- International standards development that establishes baseline requirements while accommodating national differences (Cha, 2024; Klebanov & Lizikova, 2025).
- Technical cooperation programs that help member states develop skills and knowledge (Alkış & Sökmen Alaca, 2025; Su *et al.*, 2025).
- Peer review mechanisms that enable continuous improvement and mutual learning to each other (Cha, 2024).
- Safeguards and verification that ensure countries follow their international commitments (Hatz, 2025; Klebanov & Lizikova, 2025).

Coordinated Research Projects on AI Security

To address emerging AI security challenges, the IAEA has launched a new Collaborative Research Project (CRP) called "Enhancing Computer Security of Artificial Intelligence Applications for Nuclear Technologies" (Spirito *et al.*, 2023; Su *et al.*, 2025). The main goal of this project is to improve the computer security of AI-based technologies used in nuclear facilities. It also supports the safe use of AI in small modular reactors and other nuclear applications

(Arhouni *et al.*, 2025; Spirito *et al.*, 2023; Yockey *et al.*, 2023).

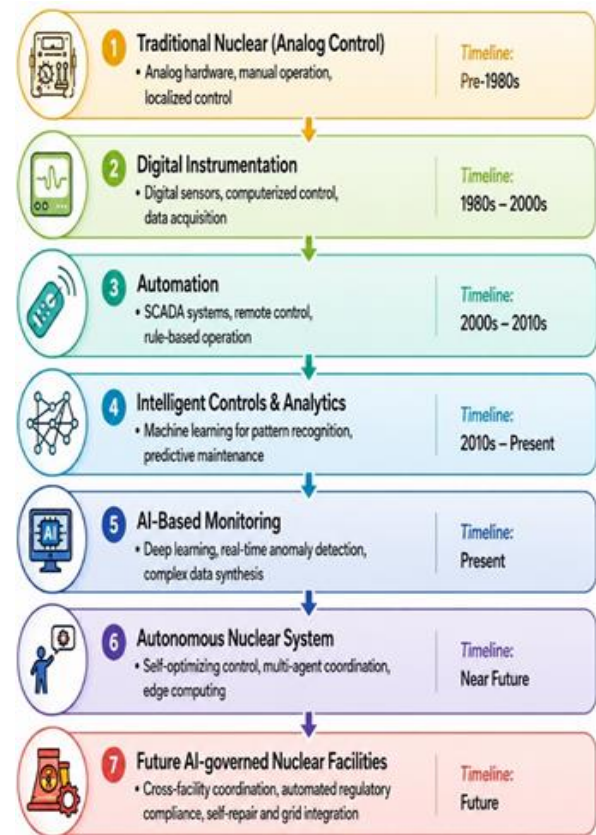


Fig. 3: Evolution of AI in Nuclear Technology.

The primary goals of these projects are as follows:

- Developing techniques and frameworks to find security vulnerabilities in AI systems used in nuclear technology (Blakely *et al.*, 2025; Sample & Eggers, 2020; Spirito *et al.*, 2023; Yockey *et al.*, 2023).
- Developing security assessment tools based on AI that can strengthen defenses against AI-related attacks (Arhouni *et al.*, 2025; Blakely *et al.*, 2025).
- Creating safeguards and security controls for AI-enabled nuclear applications (Sample & Eggers, 2020; Spirito *et al.*, 2023; Yockey *et al.*, 2023).
- Develop a framework for training to increase knowledge and skills on AI security (Arhouni *et al.*, 2025; Su *et al.*, 2025).

The project recognizes the growing application of AI and machine learning in the nuclear industry. These technologies have the potential to increase operational efficiency and improve security systems. They can also

help with identification. However, they also introduce new threats to computer security. These challenges require new and effective solutions (Arhouni *et al.*, 2025; Blakely *et al.*, 2025; Jendoubi & Asad, 2024). The CRP is open to research organizations from all IAEA member countries. The IAEA encourages the participation of women and young researchers. This helps to build a skilled and diverse workforce in the expanding field of AI safety for nuclear technology (Arhouni *et al.*, 2025; Hatz, 2025).

IAEA Safety Standards and AI Integration

A general foundation for nuclear safety is provided by the IAEA safety standards. Nuclear facilities may operate safely under these regulations. As the use of AI-powered systems increases, these standards might need to be clarified and expanded (Cancila *et al.*, 2024; Cha, 2024; Jendoubi & Asad, 2024; Su *et al.*, 2025). Several crucial factors must be considered, such as:

Defense in Depth: Defense in depth is a key principle of nuclear security. Multiple independent layers of protection are relied on it. As AI is added to digital systems, this principle must be maintained. However, AI systems can use similar development methods and training data. This can cause general weakness. As a result, this reduces their independence. Multiple layers of protection may be affected by the same vulnerability (Blakely *et al.*, 2025; O'Brien *et al.*, 2023; Shah *et al.*, 2025; Spirito *et al.*, 2023).

Safety Security Interface: Both safety and security must be balanced by nuclear facilities. The main goal of safety is to prevent accidents. To provide protection from malicious attacks are the main goal of security. The use of AI systems is increasing in both areas. This can sometimes lead to conflict. A safety measure may affect security functions or may weaken security. For example, we can say that an AI safety system could isolate a compromised network. Important security monitoring systems could be accidentally disabled. Similarly, an AI security system can override safety controls and may prioritize uninterrupted operations. These situations can create additional challenges and risks (Arhouni *et al.*, 2025; Blakely *et al.*, 2025; Sample & Eggers, 2020; Spirito *et al.*, 2023).

Graded Approach: The IAEA adopts a graded approach. This means that more stringent conditions apply

to systems with greater security concerns. The same concept must be followed by AI systems. First of all, we need to assess how important each AI function is for security. After that, we can determine the level of risk. The security conditions can be adjusted based on this. AI systems that perform highly critical security functions should be subject to stronger protection measures and more rigorous controls, whereas less critical systems may require fewer security measures. This approach helps ensure that security measures are proportionate to the level of risk and aligned with actual security needs (Cha, 2024; O'Brien *et al.*, 2023; Jendoubi & Asad, 2024; Wasil, Clymer, *et al.*, 2024).

Nuclear Security Guidance and AI

The IAEA Nuclear Security Series provides guidance on protecting nuclear facilities from malicious activity. This includes important documents related to computer security at nuclear facilities. (Alkış & Sökmen Alaca, 2025; Law & Ho, 2023; Spirito *et al.*, 2023).

An important structure is provided by these documents. These include:

- Identifying critical digital assets that need protection (Blakely *et al.*, 2025; Spirito *et al.*, 2023; Yockey *et al.*, 2023).
- Creating protective measures with clear security boundaries (Blakely *et al.*, 2025; Law & Ho, 2023; Spirito *et al.*, 2023).
- Implementing technical security controls to secure systems (Blakely *et al.*, 2025; Law & Ho, 2023; Spirito *et al.*, 2023; Yockey *et al.*, 2023).
- Managing supply chain security for digital systems (Sample & Eggers, 2020; Spirito *et al.*, 2023).
- Conducting security assessments and ongoing monitoring (Arhouni *et al.*, 2025; Blakely *et al.*, 2025; Spirito *et al.*, 2023; Yockey *et al.*, 2023).

This guidance can also be applied to AI systems. But specific AI-related topics need to be expanded. This includes keeping data secure and safe. This also includes validating and verifying AI models. Another concern is protection against machine learning attacks. These additions are important for the safe use of AI in nuclear environments (Cancila *et al.*, 2024; Cha, 2024; Jendoubi & Asad, 2024; Spirito *et al.*, 2023).

Regulatory Challenges and Governance Gaps

The Adaptation Challenge: From Nuclear Safety to AI Governance

The IAEA's nuclear security framework provides the useful lessons for managing AI (Cha, 2024). However, to directly apply this framework, we run into important limitations (Cha, 2024). Therefore, AI and nuclear technology are fundamentally different. Hence, the same method cannot always be used for both fields (Cha, 2024; Judge *et al.*, 2025).

Maturity and Stability: Nuclear technology is built on well-established scientific foundations. It also benefits from decades of management experience (Judge *et al.*, 2025; Verma & Williams, 2025b). AI technology is very different. It is developing at a very fast pace. Its capabilities are often faster than we realize. It is also growing rapidly. For this, regulatory frameworks designed for stable technologies often struggle to keep up with these changes (Cha, 2024; Judge *et al.*, 2025; Li, 2026; Verma & Williams, 2025b).

Predictability and Verification: Though nuclear systems are complex, their behavior is generally predictable. These can be verified, modeled, and tested against established specifications (Judge *et al.*, 2025; Verma & Williams, 2025a). AI behaviors cannot always be predicted from their design. Even after passing all tests, an AI system may fail unexpectedly during real-world operations (Cha, 2024; Judge *et al.*, 2025).

International Consensus: The nuclear sector benefits from decades of international cooperation. Different countries have developed common policies and practices over the years (Cha, 2024; Li, 2026). AI governance is different. It still lacks a coordinated approach and is fragmented. Different policies are followed by different countries. These differences are varying levels of risk-taking, economic interests, and are influenced by cultural values (Cha, 2024; Li, 2026; Wasil, Barnett, *et al.*, 2024).

Liability and Legal Responsibility

The addition of AI in nuclear facilities raises complex questions about liability. These questions become important when AI systems contribute to nuclear damages (Klebanov & Lizikova, 2025; Trout, 2025).

Attribution Challenges: It becomes very difficult to find the cause if there are any damages associated with an AI system. It is hard to know what exactly went wrong. The AI may have made a wrong decision. This may be due to faulty training data. This could also be due to limitations of the algorithm. It could also be due to malicious manipulations. Sometimes humans can also be responsible. They may have failed to control the AI in time. The developer may be at fault in other cases. System integrators also may be responsible. The operator of the organization may perform the main responsibilities. This makes accountability complex and unclear (Judge *et al.*, 2025; Klebanov & Lizikova, 2025; Trout, 2025).

Existing Liability Regimes: International nuclear liability agreements generally place liability on operators of nuclear facilities. This makes it faster and easier to pay compensation to victims. But it also could reduce the pressure on AI developers to fully ensure security. Klebanov and Lizikova clearly highlight this point. They point out that the use of AI in nuclear power plants challenges the current liability system. It becomes a serious issue when AI technology is part of the cause of an accident (Klebanov & Lizikova, 2025; Trout, 2025).

Insurance Implications: The behavior of AI failures is often unpredictable. This makes it very difficult for insurance companies to assess risk. Insurers may not fully understand when or how failure may occur. They may not even be aware of the potential for these failures. They may hesitate to provide insurance for AI-related risks for this reason. This uncertainty can create obstacles. This could slow down the use of AI in nuclear systems (Klebanov & Lizikova, 2025; Trout, 2025).

Non-Proliferation and Safeguards Implications

AI integration affects the implementation of IAEA safeguards. It also creates new challenges for the non-proliferation regime. As AI adoption grows, both the risks and opportunities increase. These changes may have an impact on the nuclear sector and the international non-proliferation regime (Allison & Herzog, 2025; Cha, 2024; Klebanov & Lizikova, 2025; Nelson, 2021; Verma & Williams, 2025a).

Safeguards Enhancement: AI is revealing new ways to strengthen security systems. This can improve the open-source information, analysis of inspection data and satellite imagery. Undeclared nuclear activity might be indicated by machine learning algorithms, which can find patterns. Some of these patterns may be difficult for people to identify. This helps to enhance surveillance and supports non-proliferation efforts (Allison & Herzog, 2025; Klebanov & Lizikova, 2025; Nelson, 2021; Verma & Williams, 2025a).

Safeguards Vulnerabilities: On the other hand, reliance on AI for safeguards can create new risks. Inspectors may become dependent on AI tools to detect unusual activity. They may hide activity in a way that evades the AI's notice if adversaries understand how these tools work. This may weaken the effectiveness of the security system. Therefore, an AI-enriched security system must be continuously improved and updated. This is essential to remain effective against evolving avoidance tactics (Allison & Herzog, 2025; Klebanov & Lizikova, 2025; Nelson, 2021; Verma & Williams, 2025a).

Verification in the AI Era: An even greater challenge is maintaining trust in nuclear verification. Advanced AI technology is used by both the inspected state and the inspector. This makes the verification process more complicated. Transparency and trust-building measures need to be developed. They need to work on AI-driven stealth methods. They must protect sensitive secrets while maintaining rightful public trust at the same time (Allison & Herzog, 2025; Klebanov & Lizikova, 2025; Nelson, 2021; Su *et al.*, 2025; Verma & Williams, 2025a; Wasil, Barnett, *et al.*, 2024).

Export Control and Technology Governance

AI technologies used in nuclear applications are subject to difficult export control systems. To prevent nuclear proliferation, these rules are designed. They also support peaceful international cooperation at the same time (Allison & Herzog, 2025; Klebanov & Lizikova, 2025; Nelson, 2021; Wasil, Barnett, *et al.*, 2024).

Dual Use Classification: Determining which AI capabilities require export controls is not easy. This requires a careful assessment of current capabilities. It is also necessary to think about future possibilities for

this. These days, an AI algorithm may seem harmless. For example, we can say what improves reactor performance may not have much to do with weapon use. But this situation may change over time. The same algorithms may become more sensitive as AI improves. Its relevance may increase for weapons-related uses. This makes decision-making controls more complex (Allison & Herzog, 2025; Klebanov & Lizikova, 2025; Nelson, 2021; Wasil, Barnett, *et al.*, 2024).

Regulatory Arbitrage: Different countries have different rules to control the export of AI. A risk called regulatory arbitrage, which is created by this. Developers can transfer sensitive technology through countries with weak regulations. This could result in gaps in the control system. Such gaps make control less effective. International coordination is very important. The IAEA and other international forums assist in resolving the issue. They work to close these gaps and improve international cooperation (Allison & Herzog, 2025; Klebanov & Lizikova, 2025; Li, 2026; Nelson, 2021; Wasil, Barnett, *et al.*, 2024).

Open-Source Challenges: Much AI research, including software, technical methods, and model information, is openly published. Although this openness supports innovation, it also makes sensitive knowledge easier to access and harder to control. Some publicly available information could potentially contribute to nuclear proliferation. Traditional export-control systems were mainly designed to regulate physical goods and restricted technical information, making them less effective when digital knowledge can be shared instantly across borders. As a result, controlling the spread of sensitive AI-related knowledge remains a significant challenge (Allison & Herzog, 2025; Klebanov & Lizikova, 2025; Nelson, 2021; Wasil, Barnett, *et al.*, 2024).

The Role of National Regulatory Authority

The national regulatory authority has critical role to ensure the nuclear security with the update of technology. AI integrated systems of nuclear infrastructure can be a breaching point for cyber-attacks which might have adverse consequences on safety and security of the nuclear installations. Therefore, regulatory body needs to establish strict regulations and guidance to ensure AI integrated systems are

protected against hackers. Again, the loss of human control in case of sensitive decisions during the operation of NPP using AI can be one of the greatest threats. To mitigate this problem, regulatory authority shall demand through regulations that well-trained personnel must verify and authorize in case of each critical command. The US Nuclear Regulatory Commission (NRC) is a prime example of active national engagement in addressing AI security challenges in the nuclear sector. Its research department conducts technical research to support regulatory activities related to cybersecurity (Nuclear Energy Agency, 2019). These efforts include the assessment of autonomous systems and characterization of cybersecurity conditions using AI or ML (Nuclear Energy Agency, 2019). They also focus on assessing AI performance gaps for security-sensitive systems and risk analysis of wireless technologies (Nuclear Energy Agency, 2019). The NRC collaborates closely with the Department of Energy and educational institutions such as Purdue University (Zhang & Kelly, 2023). It also partners with industrial organizations like the Electric Power Research Institute to coordinate cybersecurity research (Zhang & Kelly, 2023). This collective effort recognizes that effective AI governance requires multi-disciplinary expertise (Zhang & Kelly, 2023). This collaborative framework directly assists in the development of regulatory guidelines (Klevtsov *et al.*, 2016). An example is Regulatory Guide 5.71 regarding cyber-security programs (Klevtsov *et al.*, 2016). The NRC is actively looking forward to future updates to this guidance to address emerging AI applications (Klevtsov *et al.*, 2016). International cooperation is essential for a strong AI security regime (Cha, 2024). This global challenge is clearly beyond the capabilities of any single country (Cha, 2024). The IAEA serves as the primary multilateral forum for nuclear regulations (Cha, 2024). It successfully facilitates bilateral exchanges, quality coordination, and shared learning (Cha, 2024). Specialized technical working groups under this type of structure enable cross-border expert consensus (Rakitin & Chebyshov, 2022). These groups establish specific AI security requirements, assessment methods, and best practices (Rakitin & Chebyshov, 2022). Furthermore, advanced developed regulatory jurisdictions actively contribute to global resilience (Paperiello, 2011). They do this through capacity

building initiatives such as the IAEA's Technical Cooperation Programme (Paperiello, 2011). These programs provide training and important knowledge transfer to help developing countries establish a credible AI oversight System (Paperiello, 2011). Emerging regulatory philosophies reflect diverse national contexts but converge strongly on three core principles (Borysiewicz *et al.*, 2015). Firstly, risk-based control regulation balances the intensity of direct supervision with the potential consequence of AI failure (Borysiewicz *et al.*, 2015). This approach requires structured procedures to assess both feasibility and operational impact (Borysiewicz *et al.*, 2015). Secondly, the technology-neutral framework emphasizes performance-based and risk - informed criteria (Nuclear Energy Agency, 2019). As explored by the NRC, these criteria remain applicable across evolving AI architectures without continuous regulatory revision (Nuclear Energy Agency, 2019). Thirdly, adaptive regulation supports flexible, interactive models capable of responding to the rapid evolution of AI (Cha, 2024). This requires new rules to be made compatible with the existing legal framework (Cha, 2024). At the same time, it will enable timely, context - sensitive intervention (Cha, 2024). Together, these approaches aim to balance technological innovation and the imperative of maintaining nuclear safety and security. Therefore, a regulatory authority can ensure the nuclear security with the evolution of new technology through establishing proper legal infrastructure by proper collaboration with national and international organizations.

2. Conclusion

The addition of AI to the nuclear industry has afforded it novel opportunities to enhance safety and operational efficiency. However, it raises digital vulnerability, regulatory complexity, and autonomous misalignments. When new threats arrive, policies must be changed. In addition, regulations must be flexible and always updatable as AI technology is developing very rapidly. The regulatory authority can play critical role to ensure the nuclear security with the update of technology by updating national legislation. The implementing guide and technical guides need to be updated properly to address the upcoming threat on nuclear security in case of using AI in the nuclear

industries. Though the use of AI facilitates the safety, it can be threat to the nuclear security especially to the cyber security and thus safety and security interface need to be balanced before implementing AI in the nuclear industries. This safety-security interface balancing should be address accordingly in the national legislation. Global nuclear legislation must be rearranged from a rigid framework to a highly adaptive and risk-informed pattern to ensure safe utilization of AI in “Atoms for Algorithms”. Establishments like the IAEA and national regulators have a crucial role in enforcing global security standards. Finally, the secure use of AI in critical nuclear infrastructure requires compatibility between technological innovation and rigorous safety and security standards.

3. Author Contributions

M.G.A.: Study conception and design, Data collection, Analysis, and Manuscript preparation; K.Z.I.: Study conception and design, Data collection, Analysis, and Manuscript preparation; S.M.: Study conception and design, Data collection, Analysis, and Manuscript preparation.

4. Acknowledgment

We would like to acknowledge our colleagues in our field who consistently encourage us to advance and develop our scientific research.

5. Conflicts of Interest

The authors have no conflicts to disclose.

6. References

- Aghakhani, H., Meng, D., Wang, Y.-X., Kruegel, C., & Vigna, G. (2021). *Bullseye Polytope: A Scalable Clean-Label Poisoning Attack with Improved Transferability*. <http://arxiv.org/abs/2005.00191>
- Aldoseri, A., Al-Khalifa, K. N., & Hamouda, A. M. (2023). Re-Thinking Data Strategy and Integration for Artificial Intelligence: Concepts, Opportunities, and Challenges. *Applied Sciences*, 13(12), 7082. <https://doi.org/10.3390/app13127082>
- Alkış, M. A., & Sökmen Alaca, A. İ. (2025). IAEA and Nuclear Security: An Evolutionary Perspective. *Journal of Nuclear Sciences*. <https://doi.org/10.59474/nuclear.2023.63>
- Allison, D. M., & Herzog, S. (2025). Artificial Intelligence and Nuclear Weapons Proliferation: The Technological Arms Race for (In)visibility. *Risk Analysis*, 45(11), 3839–3859. <https://doi.org/10.1111/risa.70105>
- Arhouni, F. E., Abdo, M. A. S., Ouakkas, S., Bouhssa, M. L., & Boukhair, A. (2025). Artificial intelligence-driven advances in nuclear technology: Exploring innovations, applications, and future prospects. *Annals of Nuclear Energy*, 213, 111151. <https://doi.org/10.1016/j.anucene.2024.111151>
- Bae, J., & Lee, S. J. (2024). Current Progress in the Application of Artificial Intelligence for Nuclear Power Plant Operation. *Korean Journal of Chemical Engineering*, 41(10), 2851–2870. <https://doi.org/10.1007/s11814-024-00246-7>
- Baniecki, H., & Biecek, P. (2024). Adversarial attacks and defenses in explainable artificial intelligence: A survey. *Information Fusion*, 107, 102303. <https://doi.org/10.1016/j.inffus.2024.102303>
- Basak, S., & Lu, L. (2025). Safety assessment of passive safety systems in nuclear reactors using artificial neural networks. *Reliability Engineering & System Safety*, 264, 111355. <https://doi.org/10.1016/j.res.2025.111355>
- Berghoff, C., Neu, M., & von Twickel, A. (2020). Vulnerabilities of Connectionist AI Applications: Evaluation and Defense. *Frontiers in Big Data*, 3. <https://doi.org/10.3389/fdata.2020.00023>
- Blakely, B., Li, Y., Dave, A., Kultgen, D., & Vilim, R. (2025). *AI-Driven Cybersecurity Testbed for Nuclear Infrastructure: Comprehensive Evaluation Using METL Operational Data*.
- Borysiewicz, M., Kowal, K., & Potemski, S. (2015). An application of the value tree analysis methodology within the integrated risk informed decision making for the nuclear facilities. *Reliability Engineering & System Safety*, 139, 113–119. <https://doi.org/10.1016/j.res.2015.02.013>
- Brown, D. (2022). *Artificial Intelligence for Accelerating Nuclear Applications, Science, and Technology*. <https://doi.org/10.2172/1913826>
- Cancila, D., Daniel, G., Sirven, J.-B., Chihani, Z., Chersi, F., & Vinciguerra, R. (2024). Research Directions on AI and Nuclear. *EPJ Web of Conferences*, 302, 17005.

- <https://doi.org/10.1051/epjconf/202430217005>
- Cha, S. (2024). Towards an international regulatory framework for AI safety: lessons from the IAEA's nuclear safety regulations. *Humanities and Social Sciences Communications*, 11(1), 506. <https://doi.org/10.1057/s41599-024-03017-1>
- Chen, Y., Zhang, R., Lyu, J., & Hou, Y. (2024). AI and Nuclear: A perfect intersection of danger and potential? *Energy Economics*, 133, 107506. <https://doi.org/10.1016/j.eneco.2024.107506>
- Chernavskikh, V. (2024). *Nuclear Weapons and Artificial Intelligence: Technological Promises and Practical Realities*. <https://doi.org/10.55163/VBQX6088>
- Dickerson, K., Watkins, H., Sparks, D., Hughes Green, N., & Morrow, S. (2025). What Can the History of Function Allocation Tell Us About the Role of Automation in New Nuclear Power Plants? *Energies*, 19(1), 220. <https://doi.org/10.3390/en19010220>
- Ee, S. K. E., O'Brien, J., Williams, Z., El-Dakhkhni, A., Aird, M., & Lintz, A. (2024). *Adapting cybersecurity frameworks to manage frontier AI risks: A defense-in-depth approach* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2408.07933>
- Ejigu, D. A., Tuo, Y., & Liu, X. (2024). Application of artificial intelligence technologies and big data computing for nuclear power plants control: a review. *Frontiers in Nuclear Engineering*, 3. <https://doi.org/10.3389/fnuen.2024.1355630>
- Gastelum, Z. N. (2024). Quantifying Safeguards Innovations: Impacts of Innovations on International Atomic Energy Agency Safeguards Inspection Effort. *Journal for Peace and Nuclear Disarmament*, 7(1), 236–263. <https://doi.org/10.1080/25751654.2024.2364962>
- Gaurav, K., Singh, B. K., & Kumar, V. (2024). Intelligent fault monitoring and reliability analysis in safety-critical systems of nuclear power plants using SIAO-CNN-ORNN. *Multimedia Tools and Applications*, 83(22), 61287–61311. <https://doi.org/10.1007/s11042-023-17707-6>
- Grosse, K., Bieringer, L., Besold, T. R., & Alahi, A. (2024). *Towards more Practical Threat Models in Artificial Intelligence Security*. <http://arxiv.org/abs/2311.09994>
- Grossi, R. M. (2025, December 3). *The atom and the algorithm: Nuclear energy and AI are converging to shape the future* [Speech]. International Atomic Energy Agency (Accessed on 02 June, 2026).
- Gupta, K. D., Talukder, S., & Alam, S. B. (2023). *Adversarial AI threats on digitization of Nuclear Power Plants*. <https://doi.org/10.36227/techrxiv.23853510>
- Hall, A., Murray, P., Boring, R. L., & Agarwal, V. (2024). Human-Centered and Explainable Artificial Intelligence in Nuclear Operations. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 68(1), 1563–1568. <https://doi.org/10.1177/10711813241276463>
- Hatz, S. (2025). *The Nuclear Analogy in AI Governance Research*.
- He, J., & Degtyarev, N. (2023). AI and atoms: How artificial intelligence is revolutionizing nuclear material production. *Bulletin of the Atomic Scientists*, 79(5), 316–328. <https://doi.org/10.1080/00963402.2023.2245251>
- Hu, Y., Kuang, W., Qin, Z., Li, K., Zhang, J., Gao, Y., Li, W., & Li, K. (2023). Artificial Intelligence Security: Threats and Countermeasures. *ACM Computing Surveys*, 55(1), 1–36. <https://doi.org/10.1145/3487890>
- Huang, Q., Peng, S., Deng, J., Zeng, H., Zhang, Z., Liu, Y., & Yuan, P. (2023). A review of the application of artificial intelligence to nuclear reactors: Where we are and what's next. *Heliyon*, 9(3), e13883. <https://doi.org/10.1016/j.heliyon.2023.e13883>
- International Atomic Energy Agency. (2024). Small modular reactors: Advances in SMR developments 2024. IAEA. <https://doi.org/10.61092/iaea.3o4h-svum>
- International Energy Agency. (2025). Energy and AI. IEA.
- Jendoubi, C., & Asad, A. (2024). A Survey of Artificial Intelligence Applications in Nuclear Power Plants. *IoT*, 5(4), 666–691. <https://doi.org/10.3390/iot5040030>
- Johnson, J. (2019). The AI-cyber nexus: implications for military escalation, deterrence and strategic stability. *Journal of Cyber Policy*, 4(3), 442–460. <https://doi.org/10.1080/23738871.2019.1701693>

- Judge, B., Nitzberg, M., & Russell, S. (2025). When code isn't law: rethinking regulation for artificial intelligence. *Policy and Society*, 44(1), 85–97. <https://doi.org/10.1093/polsoc/puae020>
- Karim MSS, Hasan MN, Haque MT, Hasan KMRA and Hossain MA. (2026). Fuzzy-based intelligent approach for enhancing stakeholder engagement and strategic decision-making in nuclear programs. *Aust. J. Eng. Innov. Technol.*, 8(2), 283-296. <https://doi.org/10.34104/ajeit.026.02830296>
- Khalid, F., Hanif, M. A., Rehman, S., & Shafique, M. (2018). *Security for Machine Learning-based Systems: Attacks and Challenges during Training and Inference*.
- Khamaj, A., Ali, A. M., Saminathan, R., & M, S. (2024). Human factors engineering simulated analysis in administrative, operational and maintenance loops of nuclear reactor control unit using artificial intelligence and machine learning techniques. *Heliyon*, 10(10), e30866. <https://doi.org/10.1016/j.heliyon.2024.e30866>
- Klebanov, L. R., & Lizikova, M. S. (2025). Artificial Intelligence in Nuclear Energy: Legal Challenges and International Cooperation in the Search of Approaches to Regulation. *RUDN Journal of Law*, 29(2), 494–508. <https://doi.org/10.22363/2313-2337-2025-29-2-494-508>
- Klevtsov, A., Symonov, A., & Trubchaninov, S. (2016). Computer Security of NPP Instrumentation and Control Systems: Categorization. *Nuclear and Radiation Safety*, (4(72)), 65–70. [https://doi.org/10.32918/nrs.2016.4\(72\).10](https://doi.org/10.32918/nrs.2016.4(72).10)
- Kropaczek, D. J., Badalassi, V., Jain, P. K., Ramuhalli, P., & Pointer, W. D. (2023). Digital Twins for Nuclear Power Plants and Facilities. In *The Digital Twin* (pp. 971–1022). Springer International Publishing. https://doi.org/10.1007/978-3-031-21343-4_31
- Kumari, A., Dubey, R., & Sharma, I. (2023). ShNP: Shielding Nuclear Plants from Cyber Attacks Using Artificial Intelligence Techniques. *2023 Annual International Conference on Emerging Research Areas: International Conference on Intelligent Systems (AICERA/ICIS)*, 1–6. <https://doi.org/10.1109/AICERA/ICIS59538.2023.10420386>
- Law, H., & Ho, L. (2023). Can a dual mandate be a model for the global governance of AI? *Nature Reviews Physics*, 5(12), 706–707. <https://doi.org/10.1038/s42254-023-00670-4>
- Lee, C., & Lee, S. (2023). Evaluating the Vulnerability of YOLOv5 to Adversarial Attacks for Enhanced Cybersecurity in MASS. *Journal of Marine Science and Engineering*, 11(5), 947. <https://doi.org/10.3390/jmse11050947>
- Lezzi, M., Martino, L., Damiani, E., & Yeun, C. Y. (2025). A Systematic Literature Review on AI-Based Cybersecurity in Nuclear Power Plants. *Journal of Cybersecurity and Privacy*, 5(4), 79. <https://doi.org/10.3390/jcp5040079>
- Li, J. (2026). Governing High-Risk Technologies in a Fragmented World: Geopolitical Tensions, Regulatory Gaps, and Institutional Barriers to Global Cooperation. *Fudan Journal of the Humanities and Social Sciences*, 19(1), 113–137. <https://doi.org/10.1007/s40647-025-00445-4>
- Mena, P., Borrelli, R. A., & Kerby, L. (2024). Detecting Anomalies in Simulated Nuclear Data Using Autoencoders. *Nuclear Technology*, 210 (1), 112–125. <https://doi.org/10.1080/00295450.2023.2214257>
- Mendoza, H., Yadav, V., Alam, S. B., Eskins, D., & Iyengar, R. (2026). Advances in digital twins and AI/ML for condition monitoring in nuclear applications. *Frontiers in Energy Research*, 14. <https://doi.org/10.3389/fenrg.2026.1716514>
- Minh, D., Wang, H. X., Li, Y. F., & Nguyen, T. N. (2022). Explainable artificial intelligence: a comprehensive review. *Artificial Intelligence Review*, 55(5), 3503–3568. <https://doi.org/10.1007/s10462-021-10088-y>
- Mondal, K., Martinez, O., & Jain, P. (2024). Advanced manufacturing and digital twin technology for nuclear energy*. *Frontiers in Energy Research*, 12. <https://doi.org/10.3389/fenrg.2024.1339836>
- Moskalenko, V., Kharchenko, V., Moskalenko, A., & Kuzikov, B. (2023). Resilience and Resilient Systems of Artificial Intelligence: Taxonomy, Models and Methods. *Algorithms*, 16(3), 165. <https://doi.org/10.3390/a16030165>

- Najar, M., & Wang, H. (2024). Establishing operator trust in machine learning for enhanced reliability and safety in nuclear Power Plants. *Progress in Nuclear Energy*, 173, 105280.
<https://doi.org/10.1016/j.pnucene.2024.105280>
- Nelson, A. J. (2021). Innovation acceleration, digitization, and the crisis of nonproliferation systems. *The Nonproliferation Review*, 28(1–3), 177–197.
<https://doi.org/10.1080/10736700.2021.2022898>
- Nuclear Energy Agency. (2019). Nuclear Law Bulletin No. 103.
<https://doi.org/https://doi.org/10.1787/a13d1788-en>
- Ogunmolu, A. M. (2025). Enhancing Data Security in Artificial Intelligence Systems: A Cybersecurity and Information Governance Approach. *Journal of Engineering Research and Reports*, 27(5), 154–172.
<https://doi.org/10.9734/jerr/2025/v27i51500>
- Onwuatuwegwu IN., and Nwagu NB. (2021). The inevitable consequences of nuclear armament: an urgent call for nuclear disarmament, *Br. J. Arts Humanit.*, 3(4), 103-105.
<https://doi.org/10.34104/bjah.02101030105>
- Paperiello, C. J. (2011). Essential infrastructure: national nuclear regulation. *Health Physics*, 100(1), 77–85.
<https://doi.org/10.1097/HP.0b013e3181f6551a>
- Park, J., Kim, T., & Koo, S. (2023). Verification strategy for artificial intelligence components in nuclear plant instrumentation and control systems. *Progress in Nuclear Energy*, 164, 104842.
<https://doi.org/10.1016/j.pnucene.2023.104842>
- Poresky, C., Alivisatos, C., Kendrick, J., Peterson, P. F., Lew, R., Ulrich, T., & Boring, R. L. (2023). Advanced Reactor Control and Operations (ARCO): A University Research Facility for Developing Optimized Digital Control Rooms. *Nuclear Technology*, 209(3), 354–365.
<https://doi.org/10.1080/00295450.2022.2092366>
- Rakitin, I. D., & Chebyshov, S. B. (2022). Systems of Regulatory Documents Governing the Development and Implementation of Digital Monitoring and Control Systems for Nuclear Plants. *Physics of Atomic Nuclei*, 85(9), 1462–1466.
<https://doi.org/10.1134/S1063778822090277>
- Ramos, A., Carrasco, A., Fontanet, J., Herranz, L. E., Ramos, D., Díaz, M., Zazo, J. M., Cabellos, O., & Moraleda, J. (2024). Artificial intelligence and machine learning applications in the Spanish nuclear field. *Nuclear Engineering and Design*, 417, 112842.
<https://doi.org/10.1016/j.nucengdes.2023.112842>
- Rathod, R. J., K Patel, H., & Rathod, P. J. (2024). Application of Artificial Intelligence and Machine Learning in Robotic Systems for Nuclear Micro Reactor Development and Manufacturing. *Journal of Artificial Intelligence & Cloud Computing*, 3(2), 1.
[https://doi.org/10.47363/JAICC/2024\(3\)281](https://doi.org/10.47363/JAICC/2024(3)281)
- Russell, Q. S., & Linzi, N. P. (2023). AI-Driven approach for enhancing nuclear reactor safety predictive anomaly detection and risk assessment. *Vertex*, 12(2), 70–79.
<https://doi.org/10.35335/eh0bph05>
- Salehpour, A., & Al-Anbagi, I. (2024). Digital Substations: Cyberattack detection system for small modular reactor-based power plants. *IEEE Electrification Magazine*, 12(4), 57–67.
<https://doi.org/10.1109/MELE.2024.3473329>
- Sample, C., & Eggers, S. (2020). *Artificial Intelligence for Digital Security and Protections*.
<https://doi.org/10.2172/1847906>
- Sandhu, H. K., Bodda, S. S., & Gupta, A. (2023). A Future with Machine Learning: Review of Condition Assessment of Structures and Mechanical Systems in Nuclear Facilities. *Energies*, 16(6), 2628.
<https://doi.org/10.3390/en16062628>
- Sandhu, H. K., Bodda, S. S., Yan, E., Sabharwall, P., & Gupta, A. (2024). A comparative study on deep learning models for condition monitoring of advanced reactor piping systems. *Mechanical Systems and Signal Processing*, 209, 111091.
<https://doi.org/10.1016/j.ymsp.2023.111091>
- Sethu, M., Kotla, B., Russell, D., Madadi, M., Titu, N. A., Coble, J. B., Boring, R. L., Blache, K., Agarwal, V., Yadav, V., & Khojandi, A. (2023). Application of Artificial Intelligence in Detection and Mitigation of Human Factor Errors in Nuclear Power Plants: A Review. *Nuclear Technology*, 209(3), 276–294.
<https://doi.org/10.1080/00295450.2022.2067461>

- Shah, R., Irpan, A., Turner, A. M., Wang, A., Conmy, A., Lindner, D., Brown-Cohen, J., Ho, L., Nanda, N., Popa, R. A., Jain, R., Greig, R., Albanie, S., Emmons, S., Farquhar, S., Krier, S., Rajamanoharan, S., Bridgers, S., Ijitoeye, T., ... Deepmind, G. (2025). *An Approach to Technical AGI Safety and Security Work done while at Google DeepMind,* Core contributor, alphabetical order.*
- Sharikov, P. (2018). Artificial intelligence, cyber-attack, and nuclear weapons- A dangerous combination. *Bulletin of the Atomic Scientists*, 74(6), 368–373.
<https://doi.org/10.1080/00963402.2018.1533185>
- Shen, J., Tencent, M. X., Security, T., & Dpt, P. (2020). *AI Data poisoning attack: Manipulating game AI of Go.*
- Shubayr, N. (2024). Nuclear security measures: A review of selected emerging technologies and strategies. *Journal of Radiation Research and Applied Sciences*, 17(1), 100814.
<https://doi.org/10.1016/j.jrras.2023.100814>
- Shumaker, B., Brewer, S., Hashemian, A., & Kettle, R. (2023). Developments in Online Monitoring Technologies for Autonomous Microreactor Operations. *Nuclear Technology*, 209(3), 390–400.
<https://doi.org/10.1080/00295450.2022.2067460>
- Silva, M. V., Schirru, R., de Stefani, G. L., Nicolau, A. S., de Lima, A. M. M., Pereira, C. M. N. A., Guimaraes, J. V. S. A., & Gonçalves, D. M. E. (2024). Optimized modular nuclear reactor project utilizing artificial intelligence: Seed-blanket concept. *Nuclear Engineering and Design*, 423, 113187.
<https://doi.org/10.1016/j.nucengdes.2024.113187>
- Slattery, P., Saeri, A. K., Grundy, E. A. C., Graham, J., Noetel, M., Uuk, R., Dao, J., Pour, S., Casper, S., & Thompson, N. (2024). *The AI Risk Repository: A Comprehensive Meta-Review, Database, and Taxonomy of Risks From Artificial Intelligence.*
- Slattery, P., Saeri, A. K., Grundy, E. A. C., Graham, J., Noetel, M., Uuk, R., Dao, J., Pour, S., Casper, S., & Thompson, N. (2025). *The AI Risk Repository: A Comprehensive Meta-Review, Database, and Taxonomy of Risks From Artificial Intelligence.*
<https://doi.org/10.1016/j.patter.2026.101517>
- Spirito, C., Kerby, L., & Mena, P. (2023). *Complete Evaluation on Advanced Reactor Machine Learning Subversion Attacks (Final).*
<https://doi.org/10.2172/2331401>
- Su, F., Wan, W., Saalman, L., & Chernavskikh, V. (2025). *Pragmatic Approaches to Governance at the Artificial Intelligence–Nuclear Nexus.*
<https://doi.org/10.55163/BCHM4674>
- Trout, C. (2025). *Liability and Insurance for Catastrophic Losses: the Nuclear Power Precedent and Lessons for AI.*
<http://arxiv.org/abs/2409.06673>
- Vassilev, A., Oprea, A., Fordyce, A., Anderson, H., Davies, X., & Hamin, M. (2025). *Adversarial machine learning: A taxonomy and terminology of attacks and mitigations.* NIST AI 100-2e2025. doi:10.6028/NIST.AI.100-2e2025.
- Verma, A., & Williams, E. (2025a). *Towards an AI Observatory for the Nuclear Sector: A tool for anticipatory governance.*
<https://doi.org/https://doi.org/10.48550/ARXIV.2504.12358>
- Verma, A., & Williams, E. (2025b). *Towards an AI Observatory for the Nuclear Sector: A tool for anticipatory governance.*
- Villegas-Ch, W., & García-Ortiz, J. (2023). Toward a Comprehensive Framework for Ensuring Security and Privacy in Artificial Intelligence. *Electronics*, 12(18), 3786.
<https://doi.org/10.3390/electronics12183786>
- Wan, S., & Lei, J. (2025). Research on a Small Modular Reactor Fault Diagnosis System Based on the Attention Mechanism. *Energies*, 18(14), 3621. <https://doi.org/10.3390/en18143621>
- Wasil, A. R., Barnett, P., Gerovitch, M., Hauksson, R., Reed, T., & Miller, J. W. (2024). *Governing dual-use technologies: Case studies of international security agreements and lessons for AI governance.* <http://arxiv.org/abs/2409.02779>
- Wasil, A. R., Clymer, J., Krueger, D., Dardaman, E., Simeon, I., Saferai, C., Murphy, E. R., Governance, A. I., & Canada, S. (2024). *Affirmative safety: An approach to risk management for high-risk AI.*
- Ylönen, M., & Björkman, K. (2023). Integrated management of safety and security (IMSS) in the

- nuclear industry – Organizational culture perspective. *Safety Science*, 166, 106236.
<https://doi.org/10.1016/j.ssci.2023.106236>
- Yockey, P., Erickson, A., & Spirito, C. (2023). Cyber threat assessment of machine learning driven autonomous control systems of nuclear power plants. *Progress in Nuclear Energy*, 166, 104960.
<https://doi.org/10.1016/j.pnucene.2023.104960>
- Yu, Y., Zheng, Q., Yang, S., Yang, W., Liu, J., Lu, S., Tan, Y.-P., Lam, K.-Y., & Kot, A. (2024). *Unlearnable Examples Detection via Iterative Filtering*.
- Zbrzezny, A. M., & Grzybowski, A. E. (2023). Deceptive Tricks in Artificial Intelligence: Adversarial Attacks in Ophthalmology. *Journal of Clinical Medicine*, 12(9), 3266.
<https://doi.org/10.3390/jcm12093266>
- Zhang, F., & Kelly, K. (2023). Overview and Recommendations for Cyber Risk Assessment in Nuclear Power Plants. *Nuclear Technology*, 209 (3), 488–502.
<https://doi.org/10.1080/00295450.2022.2092356>
- Zhou, C., Liu, Q., & Zeng, R. (2020). Novel Defense Schemes for Artificial Intelligence Deployed in Edge Computing Environment. *Wireless Communications and Mobile Computing*, 2020, 1–20.
<https://doi.org/10.1155/2020/8832697>
- Zohuri, B. (2023). Artificial Intelligence Driven Nuclear Reactor Instrumentation and Control Measurements In-core and Out-core in Nuclear Power Plants. *Journal of Material Sciences & Manufacturing Research*, 4(1), 1.
[https://doi.org/10.47363/JMSMR/2023\(4\)145](https://doi.org/10.47363/JMSMR/2023(4)145)
- Zohuri, B. (2024). Navigating the Nuclear Renaissance Economic Viability, Zero Emissions, and the Future of Nuclear Energy with Generation IV Reactors and SMRs with Artificial Intelligence Integration. *Journal of Economics & Management Research*, 5(7), 1.
[https://doi.org/10.47363/JESMR/2024\(5\)245](https://doi.org/10.47363/JESMR/2024(5)245)
- Zubair, M., & Akram, Y. (2024). Enhancement in the safety and reliability of Pressurized Water reactors using Machine Learning approach. *Annals of Nuclear Energy*, 201, 110448.
<https://doi.org/10.1016/j.anucene.2024.110448>

Citation: Azam MG, Islam KZ, and Mahbub S. (2026). Security threats of artificial intelligence in nuclear fields and the role of regulatory authority to minimize the threats. *Aust. J. Eng. Innov. Technol.*, 8(4), 321-337.
<https://doi.org/10.34104/ajeit.026.02210237>

Copyright: © The Author(s), 2026. Published by the UniversePG. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, and provided the original work is properly cited. 