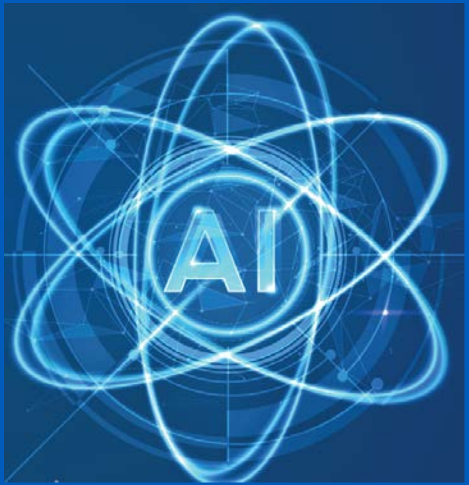
World Institute for Nuclear Security



ARTIFICIAL INTELLIGENCE TECHNOLOGIES IN NUCLEAR OPERATING SECTOR

WINS – 7th FEBRUARY 2024

Introduction to the Role of Artificial Intelligence in Strengthening the Security of Nuclear Facilities

Vienna, Austria. 06 – 08 February 2024

AI CONTEXT ON NUCLEAR FACILITIES

- **The Emergence of Generative AI Technologies with ChatGPT**
  - Innovative approaches that rely on AI technologies
  - Development of digital technology (strong managerial desire to rely on new technologies)
- **Questioning by Nuclear Safety Authorities**
  - Technical discussions on future challenges and issues that may impact the use of AI in the field of nuclear safety.
  - Potential use of AI in systems impacting the safety of operating facilities, as well as the safety demonstration and its expertise.
  - Identification of areas, systems, methods or projects that may impact the demonstration of safety in the medium and long term. The objective is to anticipate technical challenges and to prepare to accompany the position paper in a demonstration of safety and its expertise.
  - Clarification of the technical requirements that will have to be considered according to the issues and purposes, in particular with regard to explainability, interpretability, quality, integrity, security, robustness, resilience, auditability and traceability throughout the life cycle of an AI system. These aspects are also covered by the regulatory framework of the AI Act.
  - Building on current or upcoming projects related to various AI approaches such as expert systems based on Bayesian networks, learning models, surrogate models, AGI, digital twin, etc.

APPROACH

- Establishment of a multi-year roadmap to frame the development of AI by controlling all associated risks
- Joint reflections with the Authorities on the challenges of AI vis-à-vis nuclear safety
  - Workshops planned to work together on certain themes
  - First workshop decided: classification of the different types of AI (see classification proposed by NRC)
- Exchanges between French Safety Authorities and US NRC
- Exchange within the international community of operators for cyber risks related to the use of AI
  - Cyber Nuclear Forum NTI on 16 January 2024 (virtual)
  - Cyber Nuclear Forum NTI planned on 26 and 27 June (in person)

US-NRC AI STRATEGIC PLAN

The U.S. Nuclear Regulatory Commission (NRC) recognizes that interest in artificial intelligence (AI) is growing rapidly in both the public and private sectors and anticipates increased use of AI in NRC-regulated activities.

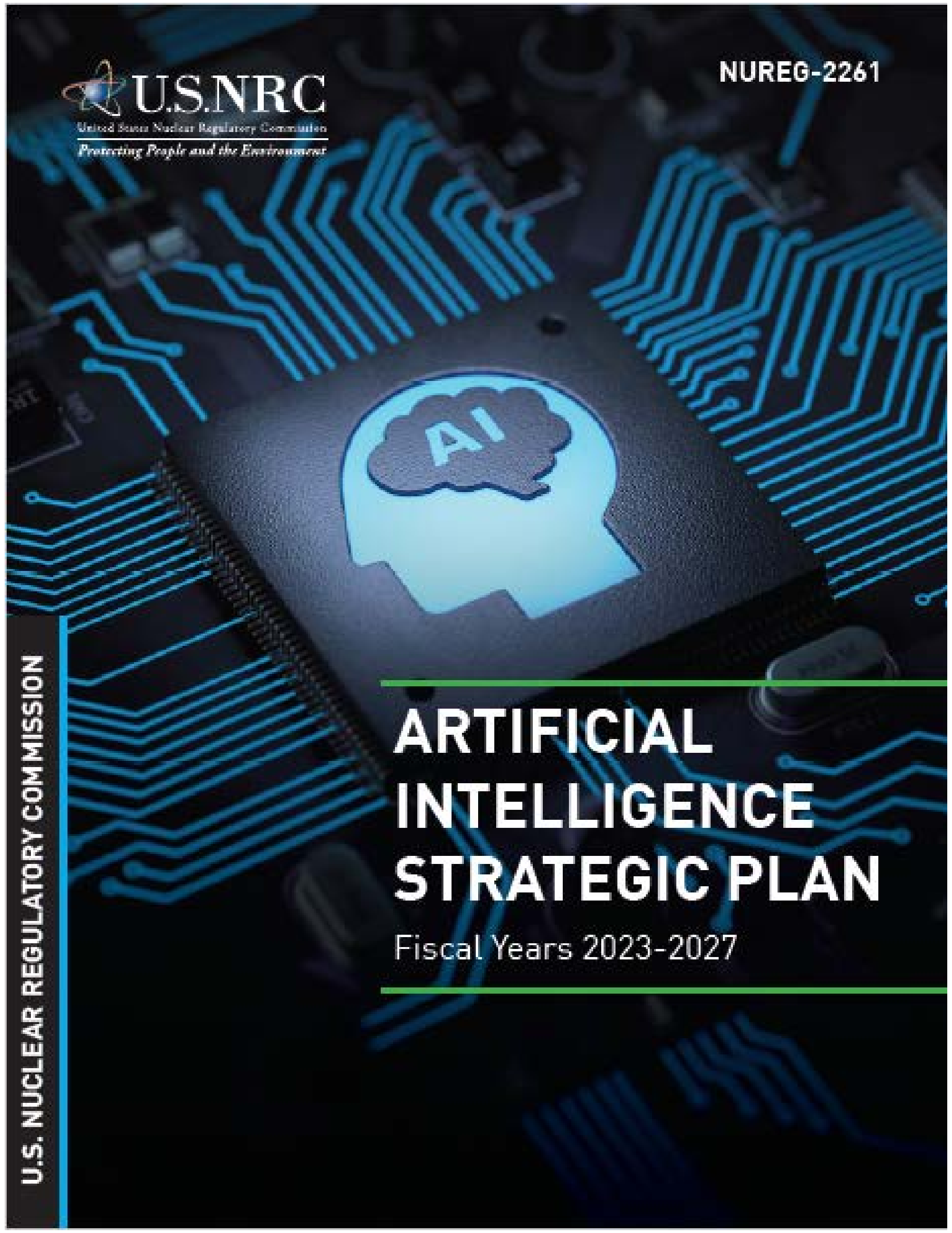
This strategic plan focuses on a broad spectrum of AI sub-specialties (e.g., natural language processing, machine learning, deep learning, etc.) which could encompass various algorithms and application examples which the NRC has not previously reviewed and evaluated

Anticipating the industry’s potential application of AI to NRC-regulated activities**, the NRC has developed this strategic plan to ensure the agency’s readiness to review such uses.**

The strategic plan includes five goals:

(1) ensure NRC readiness for regulatory decision-making
(2) establish an organizational framework to review AI applications
(3) strengthen and expand AI partnerships
(4) cultivate an AI-proficient workforce
(5) pursue use cases to build an AI foundation across the NRC.

The overall goal of this strategic plan is to ensure continued staff readiness to review and evaluate AI applications effectively and efficiently.



U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

NUREG-2261

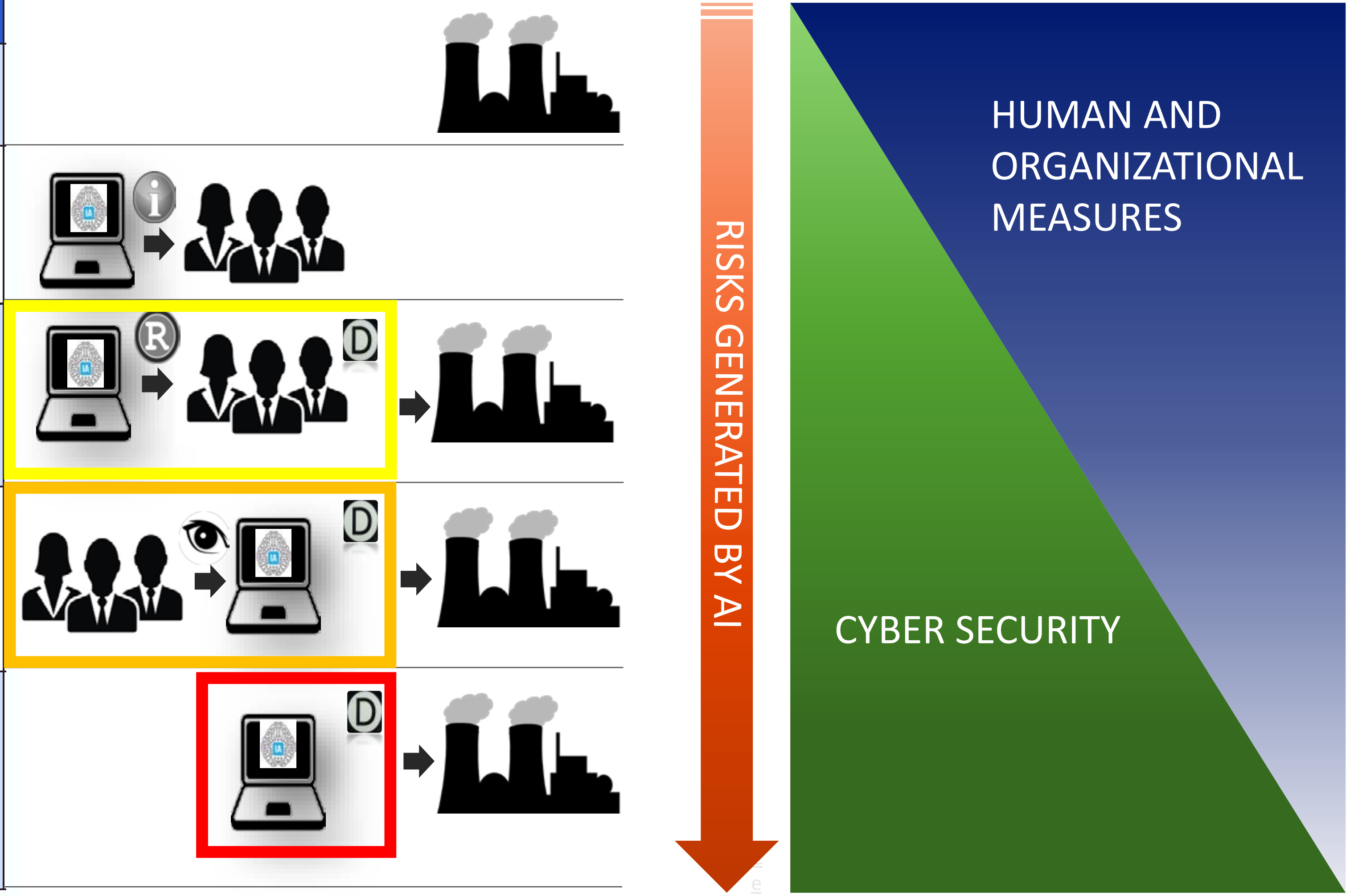
ARTIFICIAL INTELLIGENCE STRATEGIC PLAN

Fiscal Years 2023-2027

U.S. NUCLEAR REGULATORY COMMISSION

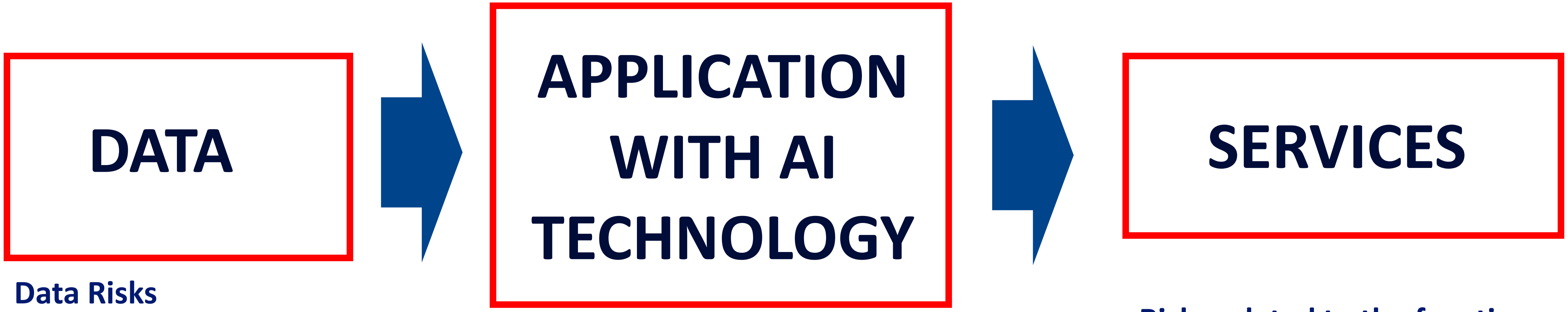
AI AUTONOMY LEVELS AND CYBER SECURITY MEASURES

Notional AI and Autonomy Levels	Potential Uses of AI and Autonomy in Commercial Nuclear Activities
Level 0: AI Not Used	No AI or autonomy integration in systems or processes
Level 1: Insight (Human decision-making assisted by a machine)	AI integration in systems is used for optimization, operational guidance, or business process automation that would not affect plant safety/security and control
Level 2: Collaboration (Human decision-making augmented by a machine)	AI integration in systems where algorithms make recommendations that could affect plant safety/security and control are vetted and carried out by a human decisionmaker
Level 3: Operation (Machine decision-making supervised by a human)	AI and autonomy integration in systems where algorithms make decisions and conduct operations with human oversight that could affect plant safety/security and control
Level 4: Fully Autonomous (Machine decision-making with no human intervention)	Fully autonomous AI in systems where the algorithm is responsible for operation, control, and intelligent adaptation without reliance on human intervention or oversight that could affect plant safety/security and control



Source: NRC

RISK CONTROL



DATA

APPLICATION WITH AI TECHNOLOGY

SERVICES

Data Risks

The risks relate to the availability, integrity and confidentiality of the data.

Risks related to the application itself

The risks relate to the security of the application itself in terms of availability and integrity. These are the risks intrinsic to the computational system itself.

Risks related to the functions provided by the application

The risks are identified and controlled through a nuclear safety analysis in case of deny of service for instance

INFORMATION SYSTEM SECURITY

To prevent attacks on data and services by ensuring fundamental properties such as confidentiality, integrity, and availability

NEW ISSUES WITH AI

- AI relies on very large amount of data that need to be available and reliable. A robust **data management is also paramount.**

- AI/ML relies on predictive methods based on likelihoods and statistics; results are not always correct or accurate enough (hallucination risk,…); training and testing of AI applications are periodically needed to improve the reliability of the results. This leads to a **large attack surface distributed over time**, as attackers can target these systems in the design phase, during initial and ongoing training or testing, and when deployed on operational systems.

- AI/ML products could have a **limited transparency** with difficulties to be interpreted and assessed. **Failures of such models could be difficult to understand.** AI relies on computational programs and codes with libraries; these programs need sufficient computational power that could be provided by cloud services (GCP, Amazon, Microsoft, etc.), but **trustworthiness in public clouds should be guaranteed. Sovereignty with private cloud could be needed for nuclear sector.**

- AI applications increase attack surface; **specific risks analysis should be carried out and specific security measures should be implemented to reduce risks** as low as acceptable. **Cyber security should be considered for models (potential backdoors,…), data (confidentiality,…), libraries (integrity,..)**

- **Existing standards should be adapted**; maybe, new standards should be developed **to control new AI technologies** (operators and regulators). Establishing fully requirements for AI/ML is a real challenge.

CYBER SECURITY CHALLENGES

- How to secure AI technologies over time? Which complementary security measures should be implemented to mitigate fresh risks generated by AI technologies?

  - Setting a framework and a policy for the use of AI-type resources for the nuclear sector: security in terms of sources that can be used to have AI (URL, data library, etc.), countermeasures to be put in place (verification, validation) when the contributions of resources are likely to change over time,
  - Applying the rules of the art in terms of balancing training data and monitoring results, with regular retraining phases if necessary,
  - Ensuring monitoring over time of the ratio of objects created by the AI vs. the total number of objects to control the looping of the AI on itself where generated data could be used to train future versions of the model,
  - Setting up a structured control system in order to have the ability to objectify the quality of the various products generated,
  - Carrying out nuclear safety and human&social analysis taking into account the associated business and safety issues.

CYBER SECURITY CHALLENGES

- Commitment of senior management with the developing of AI technologies:
  - Risk awareness
  - Investment needed
  - Balance between innovation and nuclear safety risks
  - …
- Which framework should be established by regulator and/or operator?
  - AI Classification
  - AI vs standard technologies
  - Competencies,
  - Regulatory framework
  - …
- Is our data management mature enough for supporting AI technologies?
  - Data availability and integrity
  - Amount and quality of data
  - Data governance
  - Infrastructure (private/public cloud, …)
  - …

CONCLUSION

- AI is undoubtedly necessary to improve operational performance (production, nuclear safety, and other NPP stakes like radioprotection or environment)

- AI technologies and applications are growing rapidly on nuclear facilities (**Automation,** to increase reliability, and reduce time of common operations**. Optimization,** to increase efficiencies and design of complex operations. **Analytics,** to increase the quality of current models and understanding of the used systems. **Prediction and prognostics,** to better inform maintenance activities. I**nsights,** to extract lessons from experiments and operating experience.

- However, AI could also be harmful for nuclear safety if not sufficiently controled

- It's also time to reflect collectively (operators, Authorities, stakeholders, contractors,etc.) to establish a common framework with a multi-year roadmap

THANKS FOR YOUR ATTENTION