



A Survey of Proposed Standards and Regulations for Artificial Intelligence

November 2023

Changing the World's Energy Future

Pepo Mena, Gunnar Babicz, Sindi Banda, Leslie Kerby, Christopher M Spirito



INL is a U.S. Department of Energy National Laboratory operated by Battelle Energy Alliance, LLC

DISCLAIMER

This information was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness, of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commercial product, process, or service by trade name, trade mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

A Survey of Proposed Standards and Regulations for Artificial Intelligence

Pepo Mena, Gunnar Babicz, Sindi Banda, Leslie Kerby, Christopher M Spirito

November 2023

**Idaho National Laboratory
Idaho Falls, Idaho 83415**

<http://www.inl.gov>

**Prepared for the
U.S. Department of Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517**

A Survey of Proposed Standards and Regulations for Artificial Intelligence

Pedro Mena*, Gunnar Babicz*, Sindi Banda*, Chris Spirito[†] and Leslie Kerby*[‡]

**Idaho State University College of Science & Engineering, CEADS Research Group, 921 S 8th Ave, Pocatello, ID 83209
gunnarbabicz@isu.edu, sindibanda@isu.edu, menapedr@isu.edu, & lesliekerby@isu.edu*

[†]Idaho National Laboratory 1955 N Fremont Ave, Idaho Falls, ID 83415 christopher.spirito@inl.gov

[‡]Center for Advanced Energy Studies 995 MK Simpson Blvd, Idaho Falls, ID 83401

INTRODUCTION

Over the last decade, the interest in machine learning and artificial intelligence (A.I) has grown exponentially. Researchers across several industries have explored and proposed uses of machine learning that have the potential to exponentially improve automation, efficiency, etc. The nuclear industry is no exception, with several studies exploring applications in almost every aspect of the field.

However, these developments have brought forth a number of concerns. Government officials, industry leaders and many others have called for new regulations to govern the development and implementation of A.I systems. In order for machine learning and A.I to be viable for nuclear applications, there will need to be great regulatory oversight. The purpose of this study is to examine current and proposed regulations that relate to machine learning and provide recommendations for future regulations, specifically for nuclear applications.

BACKGROUND

Benefits of A.I

A.I systems are used to gain insight and make predictions based on data. The advantage is that an A.I can be used to find relationships and trends in the data that would be difficult and time consuming for a person. These A.I models serve as the basis for automated systems. Advocates for A.I have long claimed that the technology could have great benefits to society. One study published in the Academy of Marketing Studies Journal found that businesses believe that A.I could improve five areas: decision making, customer engagement, human resources, inventory management and cyber security[1]. Alphabet's Google has claimed that A.I has improved the quality of many of its services, including navigation, image improvement, search results, etc. Today, many firms and industries utilize some form of A.I with the belief that A.I will allow for savings through increased automation, improved quality in operations and quicker responses. It is because of these benefits that fields like the nuclear industry have spent years researching ways to implement the technology.

Concerns with A.I

Along with the adaptation of new technology, it is common that groups affected by the technology, governments and the general public would have concerns on how the technology is used. A.I is no exception. In the short time that A.I has become front and center, there have been many concerns relating to the ethical use of A.I, the effect A.I will have on

labor, education, etc., transparency of the models used and the safety of such systems.

Many of these concerns are relevant for the nuclear industry namely, transparency, the effect on labor and safety. In terms of safety, the major concern is ensuring that an A.I system is accurate and reliable. In a report from the Alan Turing Institute, the United Kingdom's National Institute for Data Science, highlighted the need for developers to consider safety when implementing these systems[2]. According to this report, safety needs to extend beyond just the performance of the system as it is also important to consider the security of such a system. Without the proper attention to detail, it is possible these systems can cause harm to the general public and undermine trust in the technology.

Transparency is also a major consideration when implementing an A.I system. Today, advancements in computational technology allows for the development of highly complex predictive models, such as recurrent neural networks for time series applications and convolutional neural networks for photo classification. These models while extremely powerful, even compared to other models, are much more difficult to explain. Some have described these models as "Black Boxes", as complexity is so great that even the developers may have issues describing why a model made a decision. In a collaboration between the Swedish Department of Technology and Society and Linköping University, researchers concluded that issues with transparency in A.I will vary based on the context and application of the technology[3]. This is because different groups will have different needs, concerns, etc. for their purposes and at the context for each application must be considered. The researchers also argue that without appropriate transparency, it will be difficult for governments and the public to accept A.I.

The final area of concern in nuclear for is the effect A.I will have on the workforce such as reactor operators. Over reliance on automation, has in some cases resulted in the erosion of labor skills, sometimes referred to as "deskilling". For example, there have been a number of accidents in aerospace attributed to deskilling. Other accidents have been attributed to issues relating to the interaction between operators and the automated system. One such event occurred with American Airlines flight 695, which crashed during landing. One of the major factors noted in the accident report was the failure of the operators to revert back to manual radio navigation, when the automated system failed[4]. In the medical industry, researchers have recently began citing the need for deskilling to be examined. In a 2021 study, researchers highlighted the need to consider human factors such as deskilling, when developing a framework for medical A.I[5].

METHODOLOGY

The first step in performing this study was to determine which types of regulations would be examined. For this study, it was decided that current government regulations or those enacted by entities that have authority to enact regulations on countries would be examined first. Proposed regulations from relevant organizations, as well as governments, were also examined. Finally, accepted & proposed standards from respected professional organizations were used for this study. Using these sources, it was then possible to develop a list of recommendations that would be applicable to nuclear industry.

ENACTED REGULATIONS AND STANDARDS

United States

In the United States as of 2023, there has been no major legislation passed in the area of A.I regulations. However, there have been two major government actions toward regulating A.I/machine learning in recent times. The first is Executive Order 13960[6]. Signed in December of 2020, this order directed agencies of the federal government to follow 9 principles when developing and implementing A.I. These include: that systems are accurate, reliable, and effective, safe, secure, and resilient, as well as understandable and accountable.

The second major action the U.S government has taken in regards to A.I is the NAIIA (National AI Initiative Act of 2020). The National AI Initiative Act of 2020. while not enacting regulations on A.I, directed the Department of Energy (DOE) to take on the role of advising the president in matters related to A.I, as well as incorporating A.I into achieving it's mission[7]. The legislation also directed the National Institute of Standards and Technology to develop voluntary standards for A.I.

Europe

In June of 2023, the European Parliament, passed the EU Artificial Intelligence Act or EU AI Act. This act has been promoted as providing comprehensive regulations for all areas of A.I[8]. This includes requiring that A.I applications are assigned a risk rating. There are four categories of risk: Unacceptable risk, High risk, AI with specific transparency obligations and Minimal or no risk. Applications assigned unacceptable risk would be prohibited in the EU, while high risk applications, such as medical devices, would require oversight. Also included in this act is a requirement for transparency. This would require that the end-user be informed if they are interacting with an A.I system. It should be noted, that while this has passed the European Parliament, full implementation of this act is not expected until the end of 2023.

A.I Standards

Although A.I related laws and regulations have been slow to be enacted, professional organizations have been active in issuing standards for the use of A.I. This is critical, as many industries, including nuclear, have adopted standards from such organizations in their regulations. For example, the U.S

NRC (Nuclear Regulatory Commission) has adopted IEEE (Institute of Electrical and Electronics Engineers) Std 830 as part of it's requirements on digital software systems[9]. As such, these standards will likely play a significant role in future regulations.

IEEE

Institute of Electrical and Electronics Engineers has and is in the process of publishing several standards that relate to A.I systems. One of the relevant standards for A.I adoption is Institute of Electrical and Electronics Engineers P2247, which is intended to deal with adaptive instructional system. This standard has four components: definitions, interoperability, best practices and ethics[10]. Other standards that are currently in the process of being published include: A.I terminology, adaptive instructional systems and standards for model fairness.

In 2020, Institute of Electrical and Electronics Engineers published one of the first standards from Institute of Electrical and Electronics Engineers 2941-2021 on model representation, distribution and management[11]. This standard, 2941.1-2022, is designed to provide guidelines for the operator/A.I interface. Institute of Electrical and Electronics Engineers has also published a standard for medical A.I terminology, 2802-2022, and shared machine learning, 2830-2021.

ISO

The ISO (International Standards Organization) has also published standards relating to the use of A.I. One of the most recent standards published by International Standards Organization is ISO/IEC 23894:2023. This standard was designed to provide guidance in risk management, as it relates to A.I[12]. This standard will be part of the overall International Standards Organization standard on risk management, ISO 31000:2018. International Standards Organization has also published a standard on frameworks for A.I systems, ISO/IEC 23053:2022 and a standard on concepts and terminology, ISO/IEC 22989:2022.

PROPOSED REGULATIONS

As the interest in A.I increases year by year and the concerns with the growing technology, it is surprising that there are only a few formal regulations and laws that have been enacted with regards to A.I. That said, the number of proposed guidelines, frameworks and regulations have grown in the last ten years. As it is likely that some of the proposals will form the basis for future regulations, it is important to examine and consider these when developing standards for the nuclear industry. It should be noted that these are only proposals and are non-binding agreements.

UNESCO Recommendation on AI Ethics

In November of 2021, the UNESCO (United Nations Educational, Scientific and Cultural Organization) published recommendations for the ethical use and adaptation of A.I. Although this proposal is non-binding, all 193 member states of the United Nations have adopted this resolution[13]. The purpose of these recommendations is to ensure that A.I is used

in a manner that maintains and promotes human rights.

This proposal establishes ten key principles in utilizing A.I. This includes "Do No Harm" and establishes guidelines for the appropriate use of A.I:

1. The AI method chosen should be appropriate and proportional to achieve a given legitimate aim.
2. The AI method chosen should not infringe upon the foundational values captured in this document, in particular, its use must not violate or abuse human rights.
3. AI methods should be appropriate to the context and should be based on rigorous scientific foundation.

Another important principle from this proposal is human oversight and protection. This states that an A.I system can never replace a human in terms of responsibility and accountability. It is stated that life and death decisions should never be ceded to an A.I. Finally, the UN proposal discusses the need for transparency and explainability with A.I. According to the UN, efforts should be made to develop A.I systems that provide justification for the decisions and results of the model.

A.I Bill of Rights (United States)

In late 2022, the executive branch of the U.S government submitted a proposal for A.I regulations called the "AI Bill of Rights". This white paper defined several areas that would be relevant to the nuclear industry. For example, one area of emphasis from this proposal is that the general public should be protected from unsafe and ineffective automated systems[14]. The proposal highlights the need for rigorous testing, risk assessment, mitigation strategies and continuous monitoring.

Another major area addressed with AI Bill of Rights, is the role humans play when managing decisions made by an A.I system. The proposal advocates for systems to have the ability to fall back on human decision making, without issue. This is an important issue, as monitoring by human operators will be necessary when an A.I system is in use. Also, it is likely that more training takes place to ensure the transition from automation to human control is done in a timely and reliable fashion. Finally, the proposal also includes a number of key definitions, such as what constitutes an automated system, what is considered sensitive data in the A.I context, etc.

Other Polices and Proposals

It should be noted that the standards, proposals, etc. examined in this study are only a small number of the polices being examined. Most developed nations are exploring different approaches to regulate A.I for a number of applications. Figure 1 shows the distribution of polices being explored by each of the world's top six economies (The United States, China, Japan, India, Germany and the United Kingdom)[15].

CONCLUSIONS & RECOMMENDATIONS

Definitions

Perhaps one of the key requirements nuclear regulators and researchers are going to need to implement A.I is a set

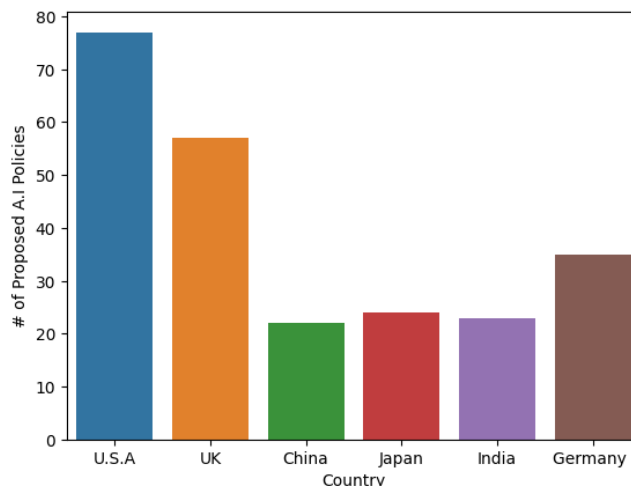


Fig. 1. Distribution of A.I Policies Among the World's Largest Economies

standard of definitions. Without these, it will be impossible for regulators to approve these systems for actual use. Developers and researchers also stand to benefit, as it will provide guidelines on what to produce. This will increase the probability of new proposals for nuclear applications being more quickly evaluated and implemented. Regulatory agencies, such as the NRC, should begin evaluating set standards, such as ISO/IEC 22989:2022 and proposed standards, such as Institute of Electrical and Electronics Engineers P3123 to see if adaptation is appropriate for nuclear applications.

System Requirements

Another key area that regulation and standard must address for the successful implementation of A.I in nuclear systems is performance requirements. Over the years, there have been many studies that have shown great evaluation results, such as accuracy. However, if these results are adequate for a nuclear system is still in question. Furthermore, there is little to no guidance in how transparent and explainable the system have to be in order to be implemented. If regulations look to address these issue, developers would have solid targets they could look to achieve. Of course, it is likely that different applications will have different requirements. Regulators should not only be looking to derive benchmark performance metrics for A.I in nuclear system, but also develop a set of applications that will be considered for approval.

Human Factors & Training

The nuclear industry has been a leader in Human Factors research and operator training for some time. A pressing question for those involved with A.I is how people will interact with the new technology. As more and more research is being performed looking to implement A.I, it is vital that the industry begin focusing on how operators will interact with the technology. Efforts should be performed to help develop regulations on how operators will interact with these A.I systems, how operators will be trained on these systems and how

operators' readiness will be maintained once these systems are implemented. This will ensure that the nuclear industry not only maintains the high level of operator performance it enjoys today, but also improves and evolves with the new advancements in technology.

Security

Security has long been an area of emphasis within the nuclear industry. The introduction of digital systems to reactor instrumentation, while extremely beneficial, does introduce new security risk. The same applies with A.I systems. Currently, the NRC in 2010, issued regulatory guide 5.71 to address cyber security with nuclear systems[16]. Since then, the threat to digital systems has only increased, with bad actors looking to target and disrupt key infrastructure. The NRC should look to expand their cyber security programs to include A.I systems. Doing this will allow for the proactive evaluations and management of potential cyber risk to proposed systems that utilize A.I.

Final Thoughts

The utilization of A.I in nuclear systems looks as it could aid an industry that will be critical to ensure energy demands are met in a sustainable and reliable manner for years to come. However, there are many steps that need to take place prior to this coming to pass. While regulations are required and these recommendations can help guide the industry into the A.I era, more work will have to be done. The nuclear industry must be proactive about its interest in A.I. For example, stakeholders in the U.S and EU should be advocating against the place of nuclear applications in the unacceptable risk category with the new EU regulations. It is important to remember that nuclear technology is highly sensitive and there are many groups who advocate against the industry. As such, it is vital that the industry begin advocating now for the use of A.I. Doing so will help ensure the continued growth and success of the nuclear field.

FUTURE WORK

The next step in this effort to help develop regulations for A.I in nuclear is to evaluate current nuclear system requirements and see which adopted and proposed regulations could be easily integrated into these requirements.

ACKNOWLEDGEMENT

Funding for this work was provided by the U.S Department of Energy.

REFERENCES

- [1] Kuldeep Bhalerao et al. "A Study of Barriers and Benefits of Artificial Intelligence Adoption in Small and Medium Enterprise". In: *Academy of Marketing Studies Journal* 26 (2022), pp. 1–6.
- [2] David Leslie. "Understanding artificial intelligence ethics and safety". In: *CoRR* abs/1906.05684 (2019). arXiv: 1906.05684. URL: <http://arxiv.org/abs/1906.05684>.
- [3] Stefan Larsson and Fredrik Heintz. "Transparency in artificial intelligence". In: *Internet Policy Review* 9.2 (2020).
- [4] Peter Ladkin. *AA965 Cali Accident Report*. (1996). URL: <http://sunnyday.mit.edu/accidents/calirep.html>.
- [5] Leslye Denisse Dias Duran. "Deskilling of medical professionals: an unintended consequence of AI implementation?" In: *Giornale di Filosofia* 2.2 (2021).
- [6] Office of the President. *Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government*. Executive Order 13960. (2020).
- [7] 116th Congress. *H.R.6216 - National Artificial Intelligence Initiative Act of 2020*. (2020).
- [8] 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*. (2021).
- [9] NRC. *Digital Systems Software Requirements Guidelines*. NUREG/CR-6734, Vol. 1. (2001).
- [10] IEEE. *IEEE portfolio of AIS technology and impact standards and standards projects*. (2015). URL: <https://standards.ieee.org/initiatives/autonomous-intelligence-systems/standards/>.
- [11] IEEE. *IEEE 2941-2021 IEEE Standard for Artificial Intelligence (AI) Model Representation, Compression, Distribution, and Management*. (2022).
- [12] ISO. *ISO/IEC 23894:2023 Information technology — Artificial intelligence — Guidance on risk management*. (2023).
- [13] UNESCO. *Recommendation on the Ethics of Artificial Intelligence*. (2021).
- [14] The White House. *BLUEPRINT FOR AN AI BILL OF RIGHTS*. (2022).
- [15] OECD AI Policy Observatory. *Policies, data and analysis for trustworthy artificial intelligence*. (2023). URL: <https://oecd.ai/en/>.
- [16] NRC. *REGULATORY GUIDE 5.71, CYBER SECURITY PROGRAMS FOR NUCLEAR FACILITIES*. (2010).