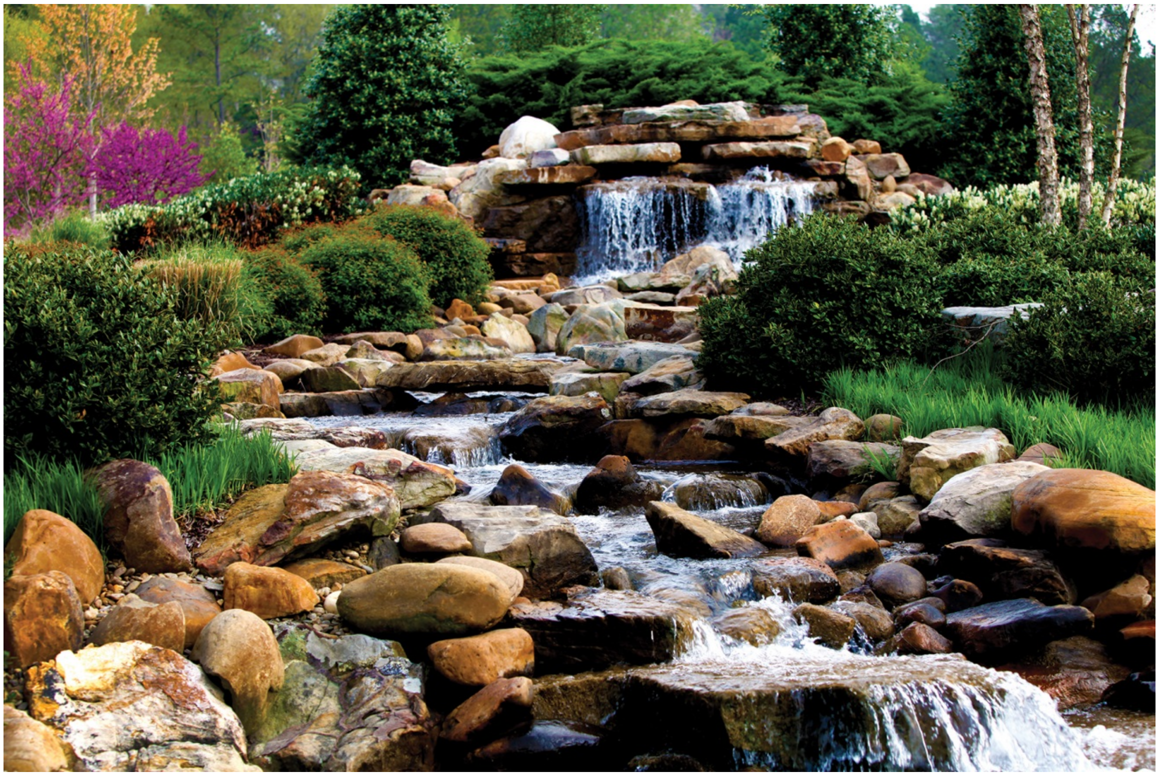
ORNL/SPR-2023/3072

Status Report on Regulatory Criteria Applicable to the Use of Artificial Intelligence (AI) and Machine Learning (ML)



M. D. Muhlheim
P. Ramuhalli
A. Huning
A. Guler
A. Saxena

September 2023

OAK RIDGE National Laboratory

ORNL IS MANAGED BY UT-BATTELLE LLC FOR THE US DEPARTMENT OF ENERGY

ORNL/SPR-2023/3072

Nuclear Energy and Fuel Cycle Division

STATUS REPORT ON REGULATORY CRITERIA APPLICABLE TO THE USE OF ARTIFICIAL INTLLIGENCE (AI) AND MACHINE LEARNING (ML)

M. D. Muhlheim
P. Ramuhalli
A. Huning
A. Guler Yigitoglu
A. Saxena

September 2023

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, TN 37831
managed by
UT-BATTELLE LLC
for the
US DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

CONTENTS

ACKNOWLEDGMENTS

The information, data, or work presented herein was funded in part by the Advanced Research Projects Agency-Energy, US Department of Energy, under Award Number DE-AR0001290. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

ABBREVIATIONS

AEA	Atomic Energy Act
AI	artificial intelligence
APRA-E	Advanced Research Projects Agency-Energy
AR	advanced reactor
BWR	boiling water reactor
CFR	Code of Federal Regulations
COL	combined operating license
CRDM	control rod drive mechanism
DC	design certification
DL	deep learning
DOE	US Department of Energy
DOT	US Department of Transportation
DT	digital twin
EASA	European Union Aviation Safety Agency
EPRI	Electric Power Research Institute
EU	European Union
FAA	Federal Aviation Administration
FAC	flow-accelerated corrosion
GDC	General Design Criterion
GE	General Electric
I&C	instrumentation and control
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
INL	Idaho National Laboratory
ISO	International Standards Organization
ITAAC	inspection, test, analysis, and acceptance criteria
LPRM	low-power range monitor
LWR	light-water reactor
MCO	moisture carryover
ML	machine learning
NASA	National Aeronautics and Space Administration
NDE	nondestructive examination
NHTSA	National Highway Traffic Safety Administration
NLP	natural language processing
NN	neural network
NPP	nuclear power plant
NRC	US Nuclear Regulatory Commission
NSTC	National Science and Technology Council
NUREG	NRC publication
O&M	operation and maintenance
ORNL	Oak Ridge National Laboratory
PHM	prognostic health management
PMx	preventive maintenance
PRA	probabilistic risk assessment
RG	regulatory guide
RPS	reactor protection system
RTS	reactor trip system
RUL	remaining useful life

SMR	small modular reactor
SSC	structures, systems, and components
TS	technical specifications
UAS	unmanned aircraft system
US	United States
V&V	validation and verification

ABSTRACT

Although the interest in the use of artificial intelligence (AI) and machine learning (ML) in nuclear energy is increasing rapidly, at present their implementation is limited. This rapid increase in interest is not surprising considering that implementing AI and ML technology would allow for continuous monitoring, facilitate the implementation of predictive maintenance with optimized staffing plans, enable automation and autonomy opportunities that could drastically reduce fixed operation and maintenance costs, and provide training for operations and maintenance. Other industries are using AI for construction, and in the nuclear arena AI could provide great benefit in decommissioning activities. The ability of AI and ML to operate in real time vastly increases their potential impact.

Before AI can be used in design, operations, or as a regulatory tool, the specifics on the regulations applicable to the use of AI for nuclear power applications need to be established. The difficulty is that the specific use cases will dictate the applicability of regulations. For example, even within the application domain associated with operations, the regulations might vary if the AI is used to create a virtual reference for plant operations or is used for training, optimization of maintenance intervals, prioritization of maintenance activities, etc. Different still is if the AI is to be used for design or setting technical specifications, which will introduce additional requirements.

US Nuclear Regulatory Commission (NRC) licensing reviews are based on an applicant's design meeting its performance assessment based on (1) safety goals and objectives, (2) deterministic and/or probabilistic analysis of accident scenarios, and (3) quantitative assessment of design alternatives against the safety goals and objectives using accepted engineering tools, methodologies, and performance criteria. The current regulatory framework does not explicitly address AI or autonomous control. However, as implementing AI technology will require the use of a digital platform, it must meet the requirements of an instrumentation and control (I&C) system.

The regulatory requirements for AI, which will be incorporated into the I&C system, will be very dependent on how it is used (i.e., its functionality, safety classification, etc.). The licensing process is primarily risk-based with the identification of components and systems as nonsafety, important to safety, or safety related. A risk-informed approach allows further gradation of components and systems based on risk metrics such as core damage frequency or large early release fractions. Thus, the use cases and the risk categorization of impacted systems and components will determine the regulatory requirements. Regardless of how AI is used it presents new opportunities for risk-informing operating, maintenance, and regulatory decisions. Trustworthiness, transparency, and the ability to validate and verify the results will be paramount in showing that the systems and plant still meet their performance requirements.

This report describes the results of research to identify regulatory implications of AI technologies and their uses. Specifically, this report reviews current regulatory guidance relevant to the application of AI for design (including design changes or new designs including advanced reactors), construction, operations, training, maintenance, research, testing, and as a regulatory tool. AI can be automated at different levels from purely informative purposes to autonomous controls. The focus of this review included determination of constraints on the application of AI technology, identification of any regulatory gaps or uncertainties, and clarification of anticipated technical basis information likely to be important for regulatory acceptance of these technologies.

Currently, any use of AI at nuclear power plants is focused on nonsafety-related applications. The NRC and other regulatory bodies are evaluating providing guidance to address gaps rather than create new regulations to address the use of AI and ML. This approach seems to be the best to encourage AI development without adding regulatory uncertainty.

1. INTRODUCTION

It is important to assess and understand how the use of artificial intelligence/machine learning (AI/ML) tools could affect the regulatory certainty of being approved for use at nuclear facilities. This task identifies the potential regulatory pathways and associated requirements for incorporating the use of AI for design, construction, operations, training, maintenance, research, testing, and as a regulatory tool. AI can be automated at different levels from purely informative to autonomous controls. Note, because ML and deep learning, and other techniques are subsets of AI, this report refers to the overarching AI term, though ML and AI are sometimes used interchangeably.

As some technical aspects of AI are still evolving, and the regulatory framework from nuclear regulators is not yet established, this document focuses on AI applications in nuclear facilities that are non-safety related yet recognizes that uses in safety-related applications will require further guidance from the regulator. Thus, this approach does not exclude the applicability of AI technologies in safety applications where the technology itself and the related regulatory framework support such potentials.

Incorporating prognostic health maintenance (PHM) and other predictive tools into an AI will likely use AI-enabled predictive maintenance and diagnostics. It is important to assess and understand how the use of these tools could affect the regulatory uncertainty of being approved for use. This is important because if it cannot be used it is of no help.

AI is transforming many fields, and the nuclear industry is interested in expanding its use and capabilities. Although recent advances in AI have enabled many more potential uses, how it is used, its autonomy, and needed regulatory perspective on its uses will slow its migration into the nuclear arena. This technical report overviews AI technologies from a nuclear perspective and explores how potential AI applications can be used in nuclear facilities.

This report provides an update on the progress made on the Advanced Research Projects Agency-Energy (ARPA-E), US Department of Energy (DOE), Milestone 5.2, “Regulatory implications for AI-enabled predictive maintenance assessed and documented,” for the GEMINA project. The objective of Milestone 5.2 is to identify the potential regulatory pathways and associated requirements for incorporation of the use of AI for operations and maintenance (O&M) of an advanced reactor (AR). The review was completed in two parts. The first part provided a review of the current regulatory guidance relevant to application of preventative maintenance (PMx) for digital twins (DTs) and the different levels of automation [1]. This task performs a similar assessment for the use of AI.

This document overviews the fundamentals of AI as it could be applied within nuclear facilities and identifies proven or potential applications within the nuclear sector and industry, with the objective to foster better understanding and adoption of AI technologies within nuclear facilities. The outcomes of this review include determination of constraints on the application of such technology, identification of any regulatory gaps or uncertainties, and clarification of anticipated technical basis information likely to be important for regulatory acceptance of PMx, AIs, and automation.

1.1 NEED FOR REGULATORY CERTAINTY

Building, maintaining, and operating a nuclear plant is expensive. In current and future plants, AI could be used to reduce O&M costs. By automating some tasks with the use of AI, nuclear facilities could reduce their operating expenses in the future.

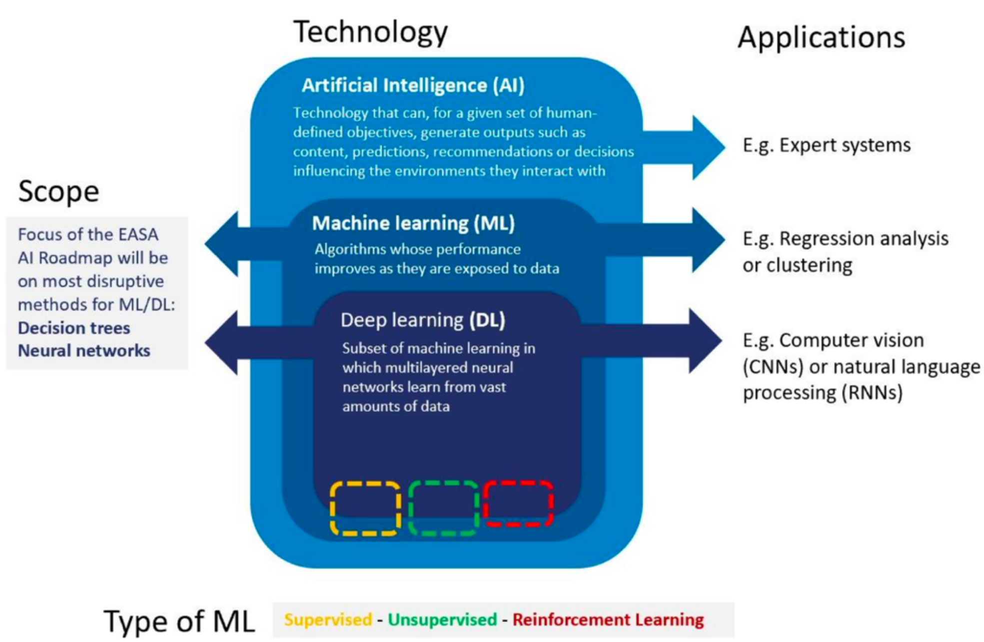
Machine-based systems using AI can go beyond defined results and scenarios and could have the ability to emulate human-like perception, cognition, planning, learning, communication, or physical action. For a given set of human-defined objectives, AI can make predictions, recommendations, or decisions influencing real or virtual environments. Thus, AI can be used in design, construction, plant operations, maintenance, R&D, testing and control. The result is increased safety, efficiency, knowledge, and savings.

Regulations must be identified so that the next essential step to develop AI-based systems can be taken. The regulations will differ depending on whether the AI has a safety, important-to-safety, or nonsafety function. Its level of autonomy will also dictate the regulatory landscape. It might be possible to identify whether existing regulations are sufficient and could be adapted or considered to accommodate AI for ARs, modifications to operating reactors, or whether new regulations and guidance are needed. However, changes to existing regulations or the creation of new regulations might not be needed; the solution could be developing guidance documents that include AI. Nevertheless, the overarching regulations that any guidance would support need to be clearly identified.

1.2 OVERVIEW

Before determining what regulations (and guidance) would be applicable to a system that uses AI to make informed decisions, it must be understood what is meant by AI and its potential uses.

AI is a catch-all term that refers to a wide set of computational techniques that allow computers and robots to solve complex, seemingly abstract problems that had previously yielded only to human cognition. AI is the term used for many types of learning that include—in addition to AI—ML, natural language processing (NLP), deep learning (DL), and data science (Figure 1) [2]. The European Union Aviation Safety Agency's (EASA's) definition for AI is "any technology that appears to emulate the performance of a human" [2]. ML is the ability of computer systems to improve their performance by exposure to data without the need to follow explicitly programmed instructions. DL is a subset of ML that emerged in recent years with deeper neural networks (NNs).



1. Types of AI that can generate outputs for making decisions [2].

The usefulness of AI is its ability to find patterns and recommend actions based on inputs from diagnostics, virtual sensors, intelligent control, aging management, preventive maintenance, anomaly detection, etc.

The workflow for an AI system is very straightforward and is comprised of three major components:

1. Input selection, curation, and data wrangling,
2. AI processing, and
3. Output post-processing.

As an example, if an AI system were used for identifying nuclear power plant (NPP) transients, it would take inputs from plant instrumentation and control (I&C) systems, perform intelligent analysis and inference, and then generate outputs that predict the type of transients the plant is experiencing.

Behind the workflow for the AI system is an AI model that has been designed and developed by human developers. Essentially there are three major approaches to build the model:

1. Encode human knowledge (e.g., expert experience) into the AI model (i.e., symbolic AI).

2. "Educate" a model by enabling it to "learn" from data gathered for this specific purpose (i.e., continuous learning, sub-symbolic AI).

3. Use both symbolic and sub-symbolic AI to create a hybrid model.

Symbolic AI makes use of strings that represent real-world entities or concepts. These strings are then stored manually or incrementally and made available to the interfacing human being/machine and when requested, they are also used to make intelligent conclusions and decisions based on the memorized facts and rules put together by propositional logic or first-order predicate calculus techniques. Sub-Symbolic AI involves providing raw data to the machine and leaving the AI to recognize patterns and create its own complex, high-dimensionality representations of the raw sensory data being provided to it.

Before assessing the regulatory requirements for an AI system, this report evaluates what is meant by AI and the challenges presented using AI to make decisions. Potential and current applications of AI are identified with examples of several use cases.

2. WHAT IS AI

AI is an umbrella term that refers to any computer algorithm that makes decisions intended to mimic or replace those made by a human. ML is a subset of AI and is a collection of computer algorithms that learn to make decisions based on the observation of training data. However, DL relies less on human feature engineering to preprocess the data of interest. Therefore, ML and DL are classes of AI. ML and AI are often used interchangeably because ML is the leading class of AI algorithms used today in many industries to perform or assist decision-making by relying on digital data that are difficult or inefficient for a human to process. Assisting decision-making has been the focus within the nuclear industry because risk-informed decision-making is used throughout many operational and design challenges; AI decision-making can be used to focus efforts on those with the greatest risk impact or importance. Because AI algorithms should always be used in tandem with a human domain expert for the performance of safety critical tasks, the term *augmented intelligence* is more appropriate when describing AI use in nuclear applications. However, this term has not yet reached sufficient awareness to be commonplace. Nevertheless, it is critical for the industry and regulator to understand the human role in any AI nuclear application.

AI can have a varying degree of influence over the decision-making process. With higher order influence comes higher cost and complexity of the algorithms. For example, the relative value and complexity of AI is its use to determine and assist in the design of additively manufactured components. At the low end is computer vision, which can be used for anomaly detection. At the top, AI can be used to predict part properties (e.g., fracture toughness) based on in situ data, process parameter information, and part geometry. Finally, a prescriptive AI would autonomously modify a part design to improve the predicted performance.

The usefulness of AI is in its ability to access more information faster, process that information to find patterns, and recommend actions based on inputs from diagnostics, virtual sensors, intelligent control, aging management, preventive maintenance, anomaly detection, etc. Advances in computing have led to massive amounts of data, and that data can be unstructured. AI can be used to structure that data.

AI can provide machines with the ability to examine examples and create models based on the inputs and desired outputs. This can be accomplished in different ways such as supervised learning, unsupervised learning, and reinforcement learning.

2.1 A BRIEF HISTORY OF AI

The history of AI dates back to the 1940s when Warren McCulloch and Walter Pitts suggested that connected neuron networks could learn [3]. Five years later, Alan Turing proposed the Turing test in his 1950s article "Computing Machinery and Intelligence" [4]. Turing suggested that humans use available information as well as reason to solve problems and make decisions and wondered why machines could not do the same thing. This was the logical framework of his 1950 paper in which he discussed how to build intelligent machines and how to test their intelligence.

The term "artificial intelligence" was first used in the mid-1950s by John McCarthy when he held the first academic conference on the subject [5]. At the conference, proof of concept was initialized through Allen Newell, Cliff Shaw, and Herbert Simon, who introduced a program designed to mimic the problem-solving skills of a human.

At the time, the biggest obstacle for advancing AI was the lack of computational power; computers simply could not store enough information or process it fast enough. The fundamental limit of computer storage and processing speed is no longer a problem, which is leading to further advances in AI.

The nuclear industry's effort in pursuing the application of AI techniques is driven by three factors that built on each other: the availability of big data, improved approaches and algorithms, and the capabilities of more powerful computers [6].

2.2 TYPES OF AI

Historically, AI has been approached from several different perspectives and accordingly diverse definitions have been established. Some definitions of AI have defined intelligence in terms of fidelity to human performance, while others prefer an abstract, formal definition of intelligence called rationality. Thinking humanly refers to the way that humans naturally think and process information, which often involves emotions, biases, and subjective reasoning. This type of thinking can be influenced by personal experiences and cultural norms. Thinking rationally refers to a more systematic and logical approach to problem-solving and decision-making. It involves using facts and evidence to make logical conclusions, rather than relying on emotions or personal biases. This type of thinking is often associated with scientific or mathematical reasoning. In summary, thinking humanly is more subjective and personal, while thinking rationally is more objective and logical.

In NUREG-2261 [7], the US Nuclear Regulatory Commission (NRC) states that "AI refers to machine-based system that can go beyond defined results and scenarios and has the ability to emulate human-like perception, cognition, planning, learning, communication, or physical action. For a given set of human-defined objectives, AI can make predictions, recommendations, or decisions influencing real or virtual environments." Encyclopedia Britannica defines AI as "the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings [8]." The Air Force Research Lab defines AI as "the ability of machines to perform tasks that normally require human intelligence – for example, recognizing patterns, learning from experience, drawing conclusions, making predictions, or taking action – whether digitally or as the smart software behind autonomous physical systems" [9]. What these definitions have in common is that AI is a technology that enables machines to think or act like humans.

A rational AI system acts to achieve the best outcome, or best expected outcome when there is uncertainty. That is, AI is a technology that enables machines to think or act rationally.

Moreover, the AI model itself can autonomously evolve through a process known as continuous learning. Technically, these two different ways for achieving intelligence are known as symbolic and sub-symbolic AI. A combination of the two approaches is also possible by leveraging to various extents the data and human knowledge.

2.2.1 Symbolic and Sub-Symbolic AI

The AI research field is divided into two camps—symbolic and sub-symbolic—that follow different paths towards building an intelligent system [10]. Symbolists firmly believe in developing an intelligent system based on rules and knowledge and whose actions are interpretable, while the sub-symbolic approach strives to build a computational system inspired by the human brain. With respect to output, symbolic AI produces declarative outputs, whereas sub-symbolic AI is based on statistical approaches and produces outputs with a given probability of error.

AI is used to make predictions, recommendations, or decisions. Figure 2 shows how symbolic and sub-symbolic AI fits in to the computational process of reading the data and providing a prediction or action.

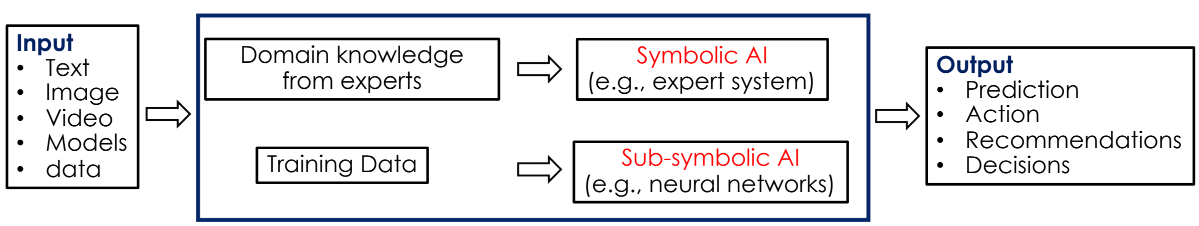


Figure 2. Symbolic and sub-symbolic systems.

Symbolic AI is based on techniques and models that manipulate symbols and structures according to explicitly defined rules to obtain inferences. Symbolic AI focuses on attempting to express human knowledge clearly in symbols, that is, facts and rules.

The traditional symbolic approach, introduced by Newell and Simon in 1976, describes AI as the development of models using symbolic manipulation [10]. In AI applications, computers process symbols rather than numbers or letters. Symbols can be arranged in structures such as lists, hierarchies, or networks and these structures show how symbols relate to each other.

For a given set of human-defined objectives, AI can make predictions, recommendations, or decisions influencing real or virtual environments.

The representations are written in a human-level understandable language.

Symbolic AI is based on techniques and models that manipulate symbols and structures according to explicitly defined rules to obtain inferences**.**

Symbolic AI is a reasoning-oriented field that relies on classical logic (usually monotonic) and assumes that logic makes machines intelligent.

The symbolic approach works best with static problems and is not a natural fit for real-time dynamic issues. It focuses on a narrow definition of intelligence as abstract reasoning, while artificial neural networks focus on the ability to recognize patterns. For example, NLP systems that use grammar to parse language are based on symbolic AI systems.

Symbolic AI refers to the fact that all steps are based on symbolic human readable representations of the problems that use logic and search to solve problem.

A key advantage of symbolic AI is that the reasoning process can be easily understood—a symbolic AI program can easily explain why a certain conclusion is reached and what the reasoning steps had been.

Sub-symbolic AI is based on techniques and models that use an implicit encoding of information that can be derived from experience or raw data. The origins of sub-symbolic AI come from the attempt to mimic a human brain and its complex network of interconnected neurons. It is based on statistical approaches and produces outputs with a given probability of error. Thus, instead of clearly defined human-readable relations, sub-symbolic AI is based on less explainable mathematical equations to solve problems and performs calculations according to some principles that have demonstrated to be able to solve problems.

Whereas symbolic AI produces declarative outputs, sub-symbolic AI is based on statistical approaches and produces outputs with a given probability of error.

Sub-symbolic AI (like DL algorithms) requires huge amounts of data to be able to learn any representation effectively. It also creates representations that are too mathematically abstract or complex to be viewed and understood.

The main assumption of the sub-symbolic paradigm is that the ability to extract a good model with limited experience makes a model successful. Here, instead of clearly defined human-readable relations, sub-symbolic AI is designed with less explainable mathematical equations to solve problems.

Sub-symbolic AI systems do not manipulate a symbolic representation to find solutions to problems. Instead, they perform calculations according to some principles that have demonstrated the ability to solve problems without exactly understanding how to arrive at the solution. Examples of sub-symbolic AI include [10]

- genetic algorithms,
- neural networks, and
- DL.

Reinforcement learning is the learning of an optimal sequence of actions to maximize a reward through interaction with an environment.

- Unsupervised AI is AI that makes use of only unlabeled data during training.
- Supervised AI is AI that makes use of labeled data during training.
- Semi-supervised AI is AI that makes use of both labeled and unlabeled data during training.

Sub-symbolic AI is also known as “connectionist AI,” and the current applications are based on this approach—from Google’s automatic transition system (that looks for patterns) to IBM’s Watson, Facebook’s face recognition algorithm, and self-driving car technology.

Neural networks, ensemble models, regression models, decision trees, and support vector machines are some of the most popular sub-symbolic AI models.

A key disadvantage of sub-symbolic AI is that it is difficult to understand how the system came to a conclusion. Another disadvantage of symbolic AI is that for the learning process the rules and knowledge have to be hand coded, which is difficult.

2.2.2 Types of Learning in AI

Based on strength, breath, and application, AI can be described in different ways. Learning in AI can fall under the types “narrow,” “general,” and “super” (Table 1).

Table 1. Types of AI

Weak or Narrow AI	Strong AI or Generalized AI	Super AI or Conscious AI
AI that is applied to a specific domain: for example, language translators, virtual assistants, self-driving cars, AI-powered web searches, recommendation engines, and intelligent spam filters. Applied AI can perform specific tasks, but not learn new ones, and can make decisions based on programmed algorithms and training data. Weak AI focuses on performing a specific task, such as detecting an anomaly in a reactor neutron monitor or optimizing a refueling plan.	AI that can interact and operate a wide variety of independent and unrelated tasks. It can learn new tasks to solve new problems, and it does this by teaching itself new strategies generalized intelligence is the combination of many AI strategies that learn from experience and can perform at a human level of intelligence. Strong AI would possess an intelligence equal to humans, and it would have a self-aware consciousness with the ability to solve problems, learn, and plan for the future.	AI with human-level consciousness, which would require it to be self-aware. Because we are not yet able to adequately define what consciousness is, it is unlikely that we will be able to create a conscious AI in the near future.

These categories demonstrate AI’s capabilities as it evolves—performing narrowly defined sets of tasks, performing the same ability to think like humans (general), and performing beyond human capability.

- Applied AI can perform specific tasks, but not learn new ones, making decisions based on programmed algorithms, and training data.
- Strong AI or generalized AI is AI that can interact and operate a wide variety of independent and unrelated tasks. It can learn new tasks to solve new problems, and it does this by teaching itself new strategies. Generalized intelligence is the combination of many AI strategies that learn from experience and can perform at a human level of intelligence.

- Super AI or conscious AI is AI with human-level consciousness, which would require it to be self-aware. It is unlikely that we will be able to create a conscious AI in the near future.

Weak or narrow AI is AI that is applied to a specific domain and that is focused on defined tasks to address a specific problem. Recognizable examples of narrow AI include specific application areas such as playing strategic games, language translation, self-driving vehicles, image recognition, virtual assistants, self-driving cars, AI-powered web searches, recommendation engines, and intelligent spam filters.

Narrow AI is not a single technical approach but rather a set of discrete problems whose solutions rely on a toolkit of AI methods along with some problem-specific algorithms. The diversity of narrow AI problems and solutions, and the apparent need to develop specific methods for each narrow AI application, has made it infeasible to "generalize" a single narrow AI solution to produce intelligent behavior of general applicability.

It is worthwhile to note the difference between "strong AI" and "weak AI," which describes the levels of intelligence within an AI system. Weak AI focuses on performing a specific task, such as detecting an anomaly in a reactor neutron monitor or optimizing a refueling plan. In contrast, strong AI would possess an intelligence equal to humans and would have a self-aware consciousness with the ability to solve problems, learn, and plan for the future. As far as the state of the art is concerned, "we are still far from building "strong AI" systems for nuclear applications" [11].

General AI (sometimes called artificial general intelligence) refers to a notional future AI system that exhibits apparently intelligent behavior at least as advanced as a person across the full range of cognitive tasks. A broad chasm seems to separate today's narrow AI from the much more difficult challenge of general AI. Attempts to reach general AI by expanding narrow AI solutions have made little headway over many decades of research. The current consensus of the private-sector expert community, with which the President's National Science and Technology Council (NSTC) Committee on Technology concurs, is that general AI will not be achieved for at least decades.

2.3 CHALLENGES

The use of AI applications will bring unique challenges, not only in its use but from the regulatory perspective.

OMB EO 13960 [12] recognizes that the ongoing adoption and acceptance of AI will depend significantly on public trust. Although the principles provided in the EO are applicable to Federal Government agencies when designing, developing, acquiring, and using AI, the principles are useful to those in industry. AI applied to nuclear operations falls within the category of digital I&C systems. Because of the uniqueness of AI characteristics in comparison with typical digital I&C systems in NPPs, specific considerations need to be taken for AI [11, [12], [13]]:

- **Autonomy**
  - Higher autonomy levels indicate less reliance on human intervention or oversight and, therefore, might require greater regulatory scrutiny of the AI system.
- **Cybersecurity and data security**
  - As more data is generated, the incentive for attempting to steal or modify that data equally increases. The AI models themselves may need to be secured against unauthorized access and modification. This includes updating the models and incorporating change control procedures.

- Users must ensure the safety, security, and resiliency of their AI applications, including resilience when confronted with systematic vulnerabilities, adversarial manipulation, and other malicious exploitation. Open source tools would necessitate steps to ensure their safety and security.

- **Trustworthiness**
  - AI applications have raised some challenges because of their high complexity and low level of reliability. To address these, licensees must demonstrate the trustworthiness of their AI.
  - The benefits of AI should significantly outweigh the risks, and the risks can be assessed and managed and that the application of AI is consistent with the use cases for which that AI was trained, and such use is accurate, reliable, and effective.
- **Transparency**
  - AI, particularly one that uses deep neural networks, operates like a black box. The input and the output of such a system are observable, but the computational process leading from one to the other is difficult for humans to understand. It is particularly difficult for humans to understand what such a system has learned and hence how it might react to input data that is different from that used during the training phase. The lack of transparency and explainability of these systems in turn creates a fundamental problem of predictability. An AI-based system might fail in ways that were unthinkable to humans because the engineers do not have a full understanding of its inner working. The lack of transparency can be problematic from a regulatory standpoint if it makes identifying the source of a problem and attributing responsibility when something goes wrong untraceable. This type of AI would likely have an increased amount of regulatory uncertainty.
- **AI verification and validation (V&V)**
  - As in the application of digital I&C systems for the nuclear industry, V&V will play a central role in enabling the use of AI systems. However, because of the black-box nature of some AI systems, there are special considerations in their V&V processes, particularly in scenarios where only a subset of the AI system can be verified, validated, or both. Concerns associated with V&V include:
    - Some systems can be partially verified and partially validated (e.g., at least one system component can be individually verified, and the remaining system components can be validated, as can the complete system).
    - Some systems cannot be verified but can be validated (e.g., no system component can be verified, but all system components can be validated, as can the complete system).
    - Some systems cannot be verified and only partially validated (e.g., no system component can be verified, but at least one system component can be individually validated).
    - Some systems cannot be verified and cannot be validated (e.g., no system component can be either verified or validated).
- **Human factors**
  - The human factors for AI needs to account for the specific human factors needs linked with the introduction of AI. Among other aspects, AI operational explainability deals with the capability to provide the human end users with understandable, reliable and relevant information with the appropriate level of details and with appropriate timing on how an AI application produces its results. As with other human-machine interfaces, there must be adequate cooperation or collaboration between the human end users and the AI-based systems.

2.3.1 Data Processing and Specifications

Data are central to any data-driven AI system's ability to learn. Data can come in structured form (e.g., relational databases in NPP data historians) or unstructured form (e.g., event reports, specification documents, nondestructive testing images and files).

The biggest hurdle in data science is always the data. Models are only ever as good as the data. Most often, the data infrastructure for currently operating NPPs was built before AI was even a thought. As a result, the data from operating plants is siloed, incomplete, and filled with errors that must be dealt with before an AI model can be developed. Historically, about 45% of the time is spent collecting, cleaning, and organizing the data [14]. Often data goes through multiple processing steps before it can be used in an AI system. These efforts seek to improve the transfer of useful (and insightful) data to operators in an easy to understand format.

Data are a key aspect of AI systems, and they go through processes (as defined in International Standards Organization [ISO]/International Electrotechnical Commission [IEC] 22989 [15] including the following [11]:

- Data acquisition. Data may be obtained from one or more sources. Thus, the suitability of the data needs to be assessed. For example, the data could be biased in some ways, or it may not be broad enough to be representative of the expected operational data input. Although the nuclear industry has been accumulating data over decades of operation, most of the data is available in separate tools and databases and may be more representative of islands of data even for a single NPP. Because the variation in data across the industry appears to be consistent, the data are considered broad enough to apply AI methods, provided they can be mapped to a standard set of data models across the industry.

- Normalization. Normalization of the data is where the data is adjusted to a common scale so that they are mathematically comparable. A significant part of data that exists in the nuclear industry uses different definition and ranges; therefore, this is a key process to enable mapping the data sets from the different plants. For example, one nuclear facility could prioritize events alphabetically, while others might use numeric prioritization. Within a process, the process data units could also differ (e.g., gallons per minute vs m^3/s).

- Data quality checking. The data needs to be examined for completeness, bias, and other factors that affect its usefulness for the AI system. For example, a sensor's partial failure (e.g., sensor drifting) or complete loss could need to be rectified before the data are used.

- Data labeling. To use supervised learning techniques, samples should be labeled. The nuclear industry data range from fully unlabeled data (e.g., process sensors missing labels on equipment condition) to fully labeled data (e.g., an event assigned a priority by a human and logged into a database).

Depending on the use case, data in an AI system can be involved in several ways [11]:

- Training data. In the context of AI, training data serve as the raw data from which the AI learning algorithm modifies its model to address the given task. These data can be of any form such as a time series, image, video, text, or acoustic. Because the current fleet of nuclear facilities are data abundant, these data usually exist. However, for newly developed ARs, these data are usually not available and a simulation of a model that represents the system is performed to generate synthetic data for use.

- Validated data. The data used by the developer needs to be validated to make or validate some algorithmic choices (hyperparameter search, rule design, etc.). This set of data is usually extracted from the same data set from which the training data are extracted.

- Test data. Test data are usually used for supervised AI methods. Given that the unsupervised AI methods lack labeled data, dedicated methods exist to evaluate the performance of the unsupervised AI algorithm, which often involve human evaluations of the results. One example of the methods used to test unsupervised AI methods involves benchmarking the findings of one model or method against others.

- Production data. Production data are processed by the AI system in the operation phase for nuclear facility applications.

AI technology allows for continuous monitoring, facilitates implementation of predictive maintenance with optimized staffing plans, and enables automation and autonomy opportunities that can drastically reduce the fixed O&M costs in advanced nuclear power systems. However, technology development and demonstration must incorporate regulatory considerations such as built-in safeguards and objectives that must be achieved for compliance. It is equally important to understand the safety and security drivers for these policies, the data needs for satisfying them, and, where appropriate, the defining technical bases that will inform future regulatory policy decisions.

2.3.2 Bias in Decision-Making

Bias in the data occurs when the available data is not representative of the population or phenomenon of study [2]. Bias in the AI model is where an error results from erroneous assumptions in the learning process. Thus, bias can cause a learning algorithm to miss the relevant relations between attributes and target outputs.

ISO/IEC TR 24027:2021, *Information technology — Artificial intelligence (AI) — Bias in AI systems and AI aided decision-making* addresses bias in relation to AI systems, especially with regard to AI-aided decision-making. Measurement techniques and methods for assessing bias are described with the aim to address and treat bias-related vulnerabilities. All AI system life cycle phases are in scope, including but not limited to data collection, training, continual learning, design, testing, evaluation, and use.

2.3.3 Autonomy and Automation

AI is often applied to systems that can control physical actuators or trigger online actions. When AI comes into contact with the everyday world, issues of autonomy, automation, and human-machine teaming arise. Autonomy refers to the ability of a system to operate and adapt to changing circumstances with reduced or without human control. For example, an autonomous car could drive itself to its destination. Despite the focus in much of the literature on cars and aircraft, autonomy is a much broader concept that includes scenarios such as automated financial trading and automated content curation systems. Autonomy also includes systems that can diagnose and repair faults in their own operation, such as identifying and fixing security vulnerabilities.

Automation occurs when a machine does work that might previously have been done by a person.[1] The term relates to both physical work and mental or cognitive work that might be replaced by AI.

---

[1] Different definitions of "automation" are used in different settings. The definition used in the main text, involving the substitution of machine labor for human labor, is commonly used in economics. Another definition is used in the

Automation, and its impact on employment, has been a significant social and economic phenomenon since at least the Industrial Revolution. It is widely accepted that AI will automate some jobs, but there is more debate about whether this is just the next chapter in the history of automation or whether AI will affect the economy differently than past waves of automation have previously.

There are different levels of autonomy [11]:

- Algorithmically based autonomy is where the AI is purely an assistance system without autonomous decision-making authority (i.e., a human analysis decides what needs to be done).
- Algorithmically informed autonomy is where the actual decision-making scope of humans and consequently their possibilities for self-determination shrink (e.g., AI-based predictive maintenance where AI estimates how long equipment has before it needs to be maintained and the human makes the final decision on that prediction).
- Algorithmically determined autonomy is where decisions are made independently by the AI and thus exhibit a high degree of autonomy (e.g., a human does not review automated decisions).

Algorithmically based AI applications work as pure assistance systems without autonomous decision-making authority. That is, AI is used to flag anomalies for a human to analyze and decide what needs to be done. Thus, the actual decision-making is made by humans and the possibilities for AI making self-determinations shrink. AI-based predictive maintenance is an example of such an approach, where AI estimates how long equipment has before it needs to be maintained and the human makes the final decision on that prediction. However, an algorithmically based AI application with a high degree of autonomy could make decisions independently. Incorporating the AI into an autonomous system means that there may no longer be a human decision in individual cases, depending upon the level of autonomy.

When considering the AI system autonomy level for nuclear facility applications, at least the following criteria should be taken into account [11]:

- The level of external supervision, either by a human operator ("human-in-the-loop") or by another automated system. Instead of substituting for human work, in some cases the machine will complement human work. Overall, the presence of accountable supervision during operation can assist in ensuring that the AI system works as intended and avoids unwanted impacts.
- The system's degree of situated understanding, including the completeness and operational ability of the system's model of the states of its environment and the certainty with which the system can reason and act in its environment.
- The degree of reactivity or responsiveness, including whether the system can adapt to internal or external changes, necessities or drives, and react to changes and whether it can stipulate future changes.

---

systems analysis setting in the US Department of Defense: Automation means that the system functions with little or no human operator involvement. However, the system performance is limited to the specific preprogrammed actions it has been designed to execute. Once the system is initiated by a human operator, it executes its task according to those instructions and subroutines, which have been tested and validated. Typically, these are well-defined tasks that have predetermined responses (i.e., rule-based responses in reasonably well known and structured environments).

- The ability to evaluate its own performance or fitness, including assessments against preset goals and the ability to decide and plan proactively in respect to system goals, motivations, and drives.

A more discretized classification system typically used for autonomous vehicles is from the SAE defines six levels of automation compared to the three levels of autonomy, from no autonomy, to informative to fully autonomous systems described in IEC TR 63468 [11]. The six levels of automation range from 0 (fully manual) to five (fully autonomous). It has been widely adopted by autonomous driving system companies worldwide.

In 1992, Sheridan defined three global levels of automation: manual control (all control is accomplished by humans), supervisory control (some or all of the control loop is closed by the computer, but the human supervisor can assert control), and fully automatic control (all control is automatic, and the human cannot vary the process except perhaps to terminate it). As technology has evolved, Sheridan [16] offered more fine-grained distinctions between these levels of automation, bringing the levels of automation to eight.

Originally, the NRC had eight levels of automation. NUREG-0700 [17] , NUREG-0711 [18], and NUREG-2261 [7] have five levels of autonomy. The NRC Strategic Plan (NUREG-2261) has four levels of automation and identifies no automation as Level 0 (not included in Levels 1–4), which maintains the five levels of automation.

Table 2 provides this in the context of operations in an NPP with five levels of automation ranging from 1 (fully manual) to five (fully autonomous). This table also shows that all of these different levels of automation already exist in operating NPPs.

Table 2. Levels of automation for NPP applications

Level	Automation functions	Human functions	NPP example
1. Manual operation	No automation	Operators manually perform all functions and tasks	Demineralized water system
2. Shared operation	Automatic performance of some functions/tasks	Operators perform some functions/tasks manually	Suppression pool cooling mode of residual heat removal service water system
3. Operation by consent	Automatic performance when directed by operators to do so under close monitoring and supervision	Operators monitor closely, approve actions, and can intervene with supervisory commands that automation follows	Advanced boiling water reactor (BWR) startup process
4. Operation by exception	Essentially autonomous operation unless specific situations or circumstances are encountered	Operators must approve of critical decisions and can intervene	Automatic depressurization system/safety relief valve system BWR automatic depressurization system AP1000 passive containment cooling
5. Autonomous operation	Fully autonomous operation. System or function not normally able to be disabled, but might be manually started	Operators monitor performance and perform backup if necessary, feasible, and permitted	Reactor scram on trip setpoint violation

The biggest concern for autonomy is a purely autonomous system that takes action without operator input. For the existing fleet of operating LWRs, the response time for safety actions to some types of events is short. For example, for the design basis loss-of-coolant accident (LOCA), emergency core cooling systems must operate immediately to prevent core damage, so most RPS actions are automated. In fact, typically no regulatory credit can be taken for a manual action that would be required within 30 minutes of the start of the event [19]. Nevertheless, for the current generation of reactors, the operating staff are a key element to ensure that plants are operated safely. Although safety actions are typically automated, they are always performed under the eye of the operating crew. The key question regarding control is, what is the appropriate level of automation? This is different than taking a protective action. However, "enabling highly or fully autonomous AI use in nuclear facilities is still not practical because of both technical and regulatory concerns" [11].

Although AI enables autonomy, not all uses of AI are autonomous.[2] For example, many AI capabilities can be used to augment human decision-making rather than replace it. Higher autonomy levels indicate less reliance on human intervention or oversight and, therefore, might require greater regulatory scrutiny of the AI system.

3. POTENTIAL APPLICATIONS OF AI

Implementation of AI technology in the nuclear arena is currently limited, and how it is used and its interfaces with plant systems will determine the regulatory criteria. Industry and regulators must develop agreed-upon guidance and frameworks for acceptance of AI applications that are consistent, explicit, and enable the use of AIs.

The regulatory requirements and guidance on the use of AI will depend on how it will be used. Figure 3 shows eight different application areas for the use of AI.

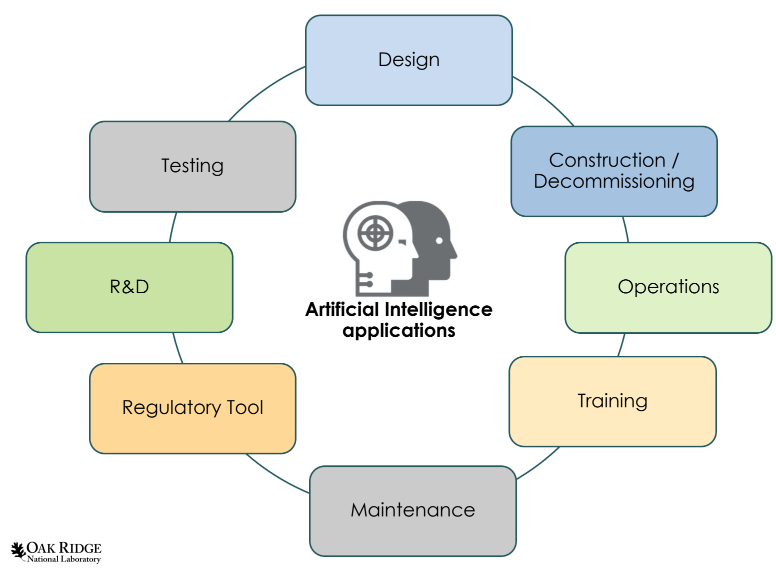


Figure 3. AI applications.

[2] Automation is considered to be a system that automatically takes action on a specific task according to predefined, prescriptive rules. For example, reactor protection systems are automatically actuated when process parameters exceed certain defined limits. In an autonomous system, both the point at which action is taken and the action that is taken are the result of training an algorithm on data collected about the system.

Applications of AIs can be mapped to two distinct lifecycle phases: (1) AR design, construction, and commissioning (2) and O&M. Several use cases were identified as described subsequently.

Many uses of AI for public good rely on the availability of data that can be used to train AI models and test the performance of AI systems. Agencies and organizations with data that can be released without implicating personal privacy or trade secrets can help to enable the development of AI by making those data available to researchers. Standardizing data schemas and formats can reduce the cost and difficulty of making new data sets useful. Implementation of AI technology in NPPs is currently limited.

3.1 DESIGN

AI can provide valuable early insights into equipment or system problems that can then be factored into component design and installation, maintenance routines, operating methodologies, and other areas [11].

Of the various prospective ***uses for AIs*** identified in this report, the primary focus of the current research involves support ***for design and O&M***. Regarding design support, there is current experience using engineering simulators as the basis for design and testing of control systems and full-scope simulators to validate operational procedures and constraints. If the AI use fits one of these categories, then the level of review would be similar to that covered in regulatory treatment of a control system or review of operational procedures. Key evidence would involve documentation of the quality and fidelity of the tool used for the designated purpose (control algorithm development, operational procedure determination, or validation of similar outcomes). The use of AIs to support design iterations for a plant or system could involve regulatory review of the design products if used in a safety-related application except for failures of a nonsafety-related AI affecting the operation of a safety system. It is also noted that if the AI is used during design to provide insights into technical specification (TS) limits, it could be subject to review to show compliance with Title 10 of the *Code of Federal Regulations*, Part 50.36 (10 CFR 50.36), *Technical Specifications*.

Optimized design will impact fuel use, reuse, and maintenance [20]. Fuel operation optimization will reduce new fuel purchases and costs from fuel maintenance operations. For example, optimizing staff productivity will of course provide benefit, but since preventing a single outage per year can be equivalent to four or more full-time equivalents and reusing a single fuel assembly during refueling is equivalent to eight or more full-time equivalents, the AI focus should be on design, maintenance, and fuel.

Safety concepts are built directly into the reactor and plant design. AI models that can determine the current plant's state within the context of probabilistic risk assessment (PRA) models would be very beneficial. One step further, AI models that could predict future states [21].

Data from the existing fleet on operating NPPs can be used to not only improve the safety/performance of the plants in the existing fleet but also to inform and guide plans for AR technologies and next-generation reactor designs. For example, the current sensor sets installed in the existing fleet impose limitations on the degree that AI can support online monitoring, structures, systems, and components (SSC) performance monitoring, condition-based maintenance programs, and so on [22]. A full understanding of the current capabilities and limitations can inform designs for future I&C systems, OLM systems, etc. This data can also be used for modeling and simulation of advanced/next-generation reactor designs to inform the design process.

AI tools have been developed to predict moisture carryover (MCO) in boiling water reactors (BWRs) [22]. Excessive MCO causes both corrosive wear and increased exposure if allowed to get too high during the fuel cycle. New advanced fuel strategies have saved significant fuel costs but, as a side effect, have exacerbated the MCO. The tool develop by BlueWave AI Labs is used by core designers and cycle managers to assess the impact of cycle-management actions, such as control-rod sequencing, on MCO. This allows core designers to assess the impact of alternative fuel configurations on future fuel cycles as part of the fuel planning process. In addition to its predictive capabilities, the neural networks model provides technical insights into operational and core design conditions that might exacerbate MCO levels. This model provides insightful information to engineers to reduce MCO levels through improved decision-making. The use of this AI has been integrated into the utilities core design and cycle management processes.

X-Energy notes that AI can be used to optimize designs (e.g., core geometry, sensor placement) [21]. AI can identify patterns unnoticed by humans. A significant advantage of AI is that the AI models can sometimes run thousands or even millions of simulations in the time it takes a human to conduct one analysis. The use of AI in design can be used to suggest control strategies, including novel approaches not thought of beforehand. During operations, AI can be used to simultaneously tune multiple controllers in control loops.

3.2 CONSTRUCTION/DECOMMISSIONING

AI tools can be incorporated into helping work order planning, with future work management applications expected in work screening, scheduling, and prioritization. Procurement and supply chains can use AI tools for material and parts ordering optimization.

One licensee is using AI in a limited business process, not directly related to nuclear plant design, O&M, or decommissioning [23]. The business process of classifying condition reports uses IBM Watson (an AI product) to make the initial classification. There is still a human review and approval of the classifications. This licensee is exploring using the same product to automate condition report classification in the near future. This licensee has also discussed using AI in simple work order creation, work order planning, scheduling, and predictive maintenance but does not have any immediate plans or approved projects to start those initiatives.

3.3 OPERATIONS

Responses to an NRC request on the use of AI in NPPs [24] was primarily in the area of O&M. This included the use of virtual sensors, intelligent control, aging management, preventive maintenance, anomaly detection, cybersecurity, and diagnostics.

Regarding the ***use of AIs for operations support***, the safety classification of the functions provided or supported by the AI and the nature of its role in operations would be key factors in determining the regulations that apply, as well as the scope and rigor of a regulatory review. Safety (e.g., reactor trip system [RTS], engineered safety features actuation system [ESFAS] and safety-related (control systems, communications systems) designations would lead to an assessment against the applicable regulations to the degree associated with the safety significance of the functions performed. As noted, regulatory reviews are not as onerous for control systems as for protection systems. It is anticipated that AIs will not be implemented as an integral part of any independent line of defense within a plant's I&C architecture (i.e., no protection functions will rely on or be incorporated into an AI) in the near future. Thus, it is likely that classification of AIs employed for operations support might range from safety to safety significant, nonsafety related, and nonsafety. Nevertheless, even a system that does not involve a safety or safety-related function must be evaluated to show that its failure would not compromise a safety function.

IEC TR-63468 [11] identifies these same issues with a different perspective.

3.3.1 Diagnostics

Diagnostics and prognostics provide the technical means for enhancing affordability and safe operation of ARs over their lifetime by enabling lifetime management of significant passive components and reactor internals [25. Here, *diagnostics* refers to the ability to determine the presence and cause of any specific condition or quantity of interest, whereas *prognostics* refers to the prediction of the expected level of change in this condition over time [26]. All systems for diagnostics and prognostics (often referred to as *prognostic health management* [PHM] systems) have several technical elements, including [25] "(1) sensors for performing measurements of both process parameters, as well as indicators of degradation; (2) diagnostic algorithms that use the sensor measurements to estimate the condition of the component; (3) prognostics algorithms to calculate the RUL [remaining useful life] of the component with degradation; and (4) interfaces to decision and control systems that are used to make O&M decisions." Although the use of AIs for diagnostics and prognostics is a relatively recent phenomenon, these systems have always required robust models that enable diagnostic and prognostic algorithms to perform their function. Fundamentally, these models (traditional regression-based models, physics models, or AIs) and the associated algorithms are required to accomplish the following [1]:

- Provide early warning of potential degradation, especially in difficult-to-access components leading to failure in AR environments.
- Provide enhanced situational awareness of plant equipment and component conditions and margins to failure, particularly for conditions in which knowledge of physics-of-failure in the AR environment is limited.
- Enable lifetime management of significant components operating in harsh environments (high-temperature, fast flux, and corrosive coolant chemistry).
- Relieve the cost and labor burden of currently required periodic inspections.
- Support a science-based justification for extended plant lifetime by ensuring reliable component operation.

Typically, these technologies have been proposed for nonsafety-significant components, and as such, they have required limited regulatory review. To date, the methods and models used have been largely data driven and statistical in nature, although some physics-based models (probabilistic fracture mechanics) have been applied for passive component diagnostics and prognostics. The technology development focus has largely been on validating and assessing the risk to plant safety associated with deploying these techniques—specifically, the risk due to a missed detection or misdiagnosis of a fault condition.

Safety evaluations by the regulator on similar technologies (for instance, online monitoring for drift/fault detection) have also tended to focus on these issues, with the need to demonstrate proper performance, quantify associated uncertainty bounds, and incorporate alternative monitoring technologies to address concerns of missed detection.

3.3.2 Virtual Sensors

AI can use data from the existing fleet to inform and guide plans for AR technologies and next-generation reactor designs [22]. For example: the current sensor sets installed in the existing fleet impose limitations on the degree that AI can support online monitoring, SSC performance monitoring, condition-based maintenance programs, and so on. A full understanding of the current capabilities and limitations can inform designs for future I&C systems, on-line maintenance (OLM) systems, and the like.

AI can be used to create virtual sensors and virtual calibration tools. They are currently being developed for low-power range monitors (LPRMs) [22]. These virtual sensors can be used to estimate values for process variables that cannot be directly measured but that are of significant value to plant O&M. In these instances, a high-fidelity multiphysics model can be developed such that values of virtual sensors and other measurable variables are generated [11]. Then, a supervised learning algorithm can be explored to find the hidden correlation. A supervised learning algorithm learns the underlying model (inferred function) between the input and the output using the training set, and the requirement is that the model should be able to generalize from the training set to unseen data samples. The trained model is then deployed to operate as the virtual sensor.

3.3.3 Preventative Maintenance

Preventative maintenance of I&C and electrical equipment is essential to ensure operational safety and reduce maintenance costs of NPPs. Particularly, it can lead to a reduction in unnecessary maintenance, thus reducing costs associated with parts, labor, and unnecessary planned, forced, or extended outages [11]. Plant performance monitoring and oversight is an active area of AI use to measure and analyze performance change across multiple functional areas.

AI and ML are often associated with AIs because the operational data are often used to detect and diagnose potential issues in the physical system through simulation and comparing against the digital system. Additionally, significant research is being performed to predict future physical system issues, also referred to as *predictive maintenance,* with DT which is also applicable to AI and ML tools [[27, [28, 29].

Preventative maintenance trending will benefit from AI-based models in several ways [22]:

1. RUL models can be developed using historical data. These models will allow condition-based replacement and maintenance as opposed to the present time-based methods. Time-based methods often replace equipment based on a time schedule. Models estimating the RUL can be used to prevent premature failures. AI can play a prominent role in development of these models.

2. These same models allow the creation of virtual sensors and virtual calibration tools. They are currently being developed for LPRMs.

There are many examples of AI are used within the nuclear industry for applications beyond AI. For example, AI is being used to predict a departure from nucleate boiling ratio, which is critical for light-water reactor (LWR) thermal hydraulic design [30]. AI is also being used to support nuclear data [31]. For more examples, see the comprehensive report Idaho National Laboratory (INL) prepared for the NRC, which consists of an industry survey and overview of AI and ML for nuclear applications [32].

Thus far, experience with AI-based tools has prevented unplanned shutdowns and saved tens of fresh fuel bundles per reload on three refueling programs [22]. The savings have been in the tens of millions and development costs in the one-half to one-million-dollar range. Similarly, AI-trained monitoring software could replace some portions of human surveillance. These systems could monitor data and written reports and detect problems/compliance issues before they occur. In addition, computer vision and AI enabled drones can be used to recognize features of their surrounding environment, eliminating or reducing the need for the human walk-arounds [11].

NLP-based tools are used on textual reports to identify critical issues and trends [23]. Some engineering groups are actively developing AI applications focused on equipment condition-based maintenance and monitoring tools, including RUL. AI tools can be incorporated into helping work order planning, with future work management applications expected in work screening, scheduling, and prioritization.

AI can use system and component data to analyze for trends and correlations [23]. Alarms can be generated with parameter drift outside of established trends. Neural network–based analytics can use NLP to enhance the review of thousands of equipment issue reports to help facilitate the identification of potential maintenance rule functional failures. Also, neural network–based analytics that use NLP and a scoring system can be used to review work orders as found condition codes, completion remarks, cost data, and maintenance frequencies to suggest changes in maintenance strategies.

Another example of a preventive maintenance task that can be improved using AI techniques is the periodic inspections of control rod drive mechanisms (CRDMs). Detecting problems in the coils of CRDMs is difficult and often involves evaluation of thousands of coil current measurements that can take hundreds of work-hours to perform. Supervised AI has been used to automate the detection of CRDM coil current issues in near real-time and predict the maintenance needs [11].

3.3.4 Operational Decision Support

In an operating NPP, a typical work process comprises multiple steps, where human-based decisions need to be made. This kind of human-based decision is not only time consuming but also prone to error. If every major and minor component in the plant has diagnostic sensors attached that feed data to the plant historian, and if there are AI models trained to detect anomalies and degradation in those components, then maintenance schedules become more efficient and there are less inspections needed [21].

Being able to detect anomalies and degradation of components using AI offers an opportunity to safely and efficiently operate both the existing fleet and plants that will be developed and built. These tools will enhance operators' abilities to plan and prepare for plant transients [33].

Electric Power Research Institute (EPRI) is using data science technologies to leverage the accumulated thousands of years of operating experience in NPPs to identify new best practices and to better inform future decisions [34]. These new trends and observations being developed are identified below:

- An exploration of the Work Order Database is assessing how different preventive maintenance strategies affect costs [35].
- A project looking at Corrective Action Program records aims to process entries automatically and obtain insights from the Corrective Action Program database.
- NLP fueled by an industry-specific dictionary provides the backbone for the aforementioned activities. A recently published Quick Insight brief on the Power Industry Dictionary for Text-Mining [36] showcases the value of this dictionary for one case study. EPRI continues to develop this dictionary to better support this and other efforts.
- Prognostics can be used in anticipating the future health of components and systems.
- Data science approaches can be leveraged to predict events, including failures, and assess current asset conditions such as RUL. If utilities are prepared, they can plan their maintenance and outage strategies, reducing unexpected downtime and minimizing periodic inspections.
- EPRI is leveraging ML techniques to inform flow-accelerated corrosion (FAC) inspection programs based on industry FAC databases [37]. A well-informed FAC inspection program could help assess needs and timing, eliminating costly and unnecessary inspections.

- Sensors installed at nuclear plants generate a wealth of data that is largely underused. EPRI has begun to mine these data for insights into plant operation and to investigate the potential trends and prognostics that can be exploited.

3.3.5 Anomaly Detection

Currently, numerous NPPs rely on alarming systems to detect abnormal situations. The objective of anomaly detection is to identify events or samples deviating from what it could be considered "ordinary" or "normal" behavior within an application domain. Informing operators after an anomaly is significant enough to actuate an alarm places the operator (and plant) in a reactive state. The use of AI, noting the changes in state, would allow operators to be proactive.

Largely, the AI tools can be broadly grouped into two categories: Anomaly detection and ML. Anomaly detection monitors live data or computed results to identify instances of data that are not consistent with the previously defined statistical norm. Such tools can give operators and designers warnings of anomalies otherwise not perceptible by a human observer (i.e., something is not quite right and might require repair). In practice, anomalies are often "spikes" in the data that show significant deviations from the expected values [38].

AI is, in principle, akin to human learning: teach the software (human) to find a pattern. Give an algorithm data and let it train itself to find patterns in the data. There are many algorithms used for AI. Westinghouse has recently developed a tool that evaluates over ten regression-based AI algorithms to find trends in the data and then select the optimal algorithm based on data-driven modeling validation metrics [38]. These functional patterns can then augment anomaly detection, as well as prognosticate future behavior, based on the historical (or simulated) data. These types of predictive capabilities are very useful in determining the RUL of a component or structure, long-term behavior of a system (maintenance related), and system optimization.

Relating to the benefits of predictive maintenance and condition monitoring, several methods are currently in use [33]. These include clustering algorithms for anomaly detection and Gaussian approaches for correcting instrument error based on adjacent or physically redundant data. In this same scope, Metroscope employs a Bayesian method to find root causes leveraging physics models, expert knowledge, and operating experience. In addition, autocorrelation methods are also being explored for nondestructive examination (NDE) applications.

DL models from image detection and NLP to generative adversarial networks for anomaly detections are being used to solve various plant problems [23].

Under the Transformational Challenge Reactor program at Oak Ridge National Laboratory (ORNL) [39], AI and ML are being used extensively to detect material defects, visualize them, and ultimately determine the material properties based on in situ data alone.

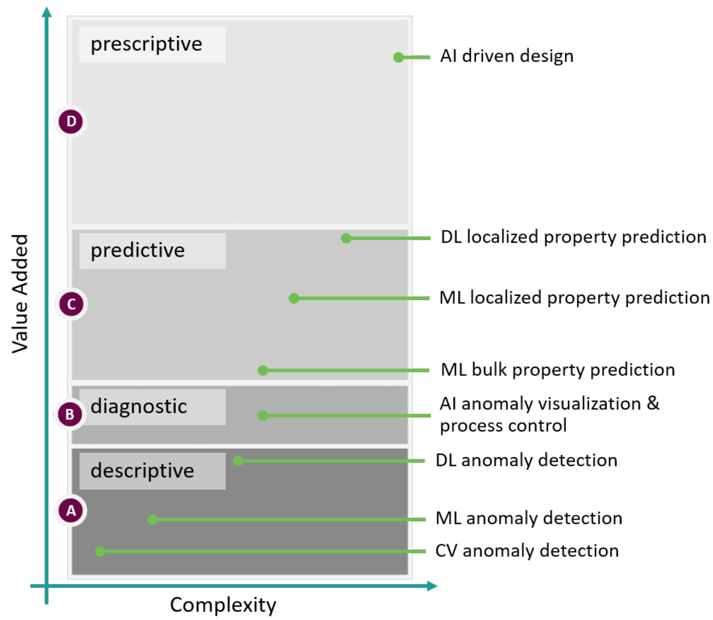


Figure 4. Relative value and complexity of different types of AI algorithms [39].

3.3.6 Cybersecurity

With the increased use of digital systems in nuclear industry, maintaining effective cybersecurity has become an ever-growing challenge. Prevention, detection, and reaction are key elements of a cybersecurity program. AI has been investigated for the detection of a cyberattack in nuclear applications [11]. The workflow of applying AI techniques to cyberattack detection consists of four steps, namely, data acquisition, feature extraction, cyber anomaly detection, and attack identification. In addition to network traffic data, process-specific features such as measurements, control commands, and setpoints are also considered.

With respect to cybersecurity, tools exist to assist in network threat detection and diagnostics. Many router security packages, for example, use AI to automatically detect and defeat denial of service attacks. X-energy plans to build its IT and data infrastructure from a cybersecurity-centric approach [21].

AI typically necessitates the creation, transfer, and evaluation of very large amounts of data. Cybersecurity and data security is a major concern. As more and more data are generated, the incentive for attempting to steal that data equally increases. Because AR companies do not have decades old IT infrastructure in place, they are more flexible to adopting state-of-the-art cybersecurity tools. Moreover, the IT infrastructure can be built with cybersecurity in the forefront of design considerations [21].

3.3.7 Human Factors Engineering

Human factors engineering has been instrumental in increasing the safety and performance of the nuclear energy industry. AI has the potential to further mitigate human errors to improve safety and performance [11].

In contrast to automation, where a machine substitutes for human work, in some cases a machine will complement human work. This might happen as a side-effect of AI development, or a system might be developed specifically with the goal of creating a human-machine team.

Various common activities in nuclear plants place staff into high-pressure or demanding situations, increasing the chance of human factor errors and personal safety risk. Many of these issues can be reduced by leveraging data science technologies for automation [34]. Even activity time can be diminished, reducing factors such as radiation exposure or critical downtime.

Responsible AI principals includes everything from hardware reliability to careful data and model management to explainable AI to human factors such as misinterpretation, automation bias, or alert fatigue [20].

NDE (pre-service and in-service) inspections of components and piping systems (including primary systems) can significantly benefit from Al/ML methods, not only from a cost perspective (efficient data analysis) but also to support/improve defense-in-depth strategies to ensure structural integrity of systems and components. As AI/ML tools are not subject to human factors [33], these advanced processes provide a complementary tool for the NDE analyst for rapid screening and focused attention on suspect regions of interest. AI/ML implementation can further provide benefit to industry by providing efficient and accurate means to detect and monitor service-induced degradation by providing a conservative risk assessment of the component when predictive analytics are incorporated to govern planned repair/replacement activities, minimizing emergent repair needs. The result could be a more robust and efficient protocol for maximizing the NDE data streams and providing actionable data in a planned fashion.

3.3.8 Intelligent Control

A fully autonomous control system is an example of intelligent control. In theory, a fully autonomous control system would be able to handle all modes of operation from startup to full power to shutdown, including load following capabilities.

For reactor operation and control, Framatome has begun to develop Al/ML I&C products such as the Operator Assistance Predictive System, which will reduce operator burden and enhance plant safety while optimizing the operation of the reactor [33]. The software employed is Artificial Narrow Intelligence, which is primarily goal seeking and operations optimizing. All of this is done with the anticipated need of greater plant flexibility because of changing demand on the grid.

X Energy, LLC, (X-energy) is exploring a fully autonomous control system in its small modular high-temperature gas reactors [21]. As such, these comments will focus on possible use cases for the AR community. The AR community is fully embracing AI technologies for a wide array of applications including autonomous control systems, performance monitoring, predictive maintenance, additive manufacturing, design optimization, etc. X-Energy also notes that autonomous control systems could support load following capabilities, which are not typical for large LWRs.

Incorporating a dynamic PRA would incorporate PRA models that provide results dynamically to provide feedback regarding possible end states, given current plant conditions and trends. It would potentially also be possible to determine the plant's state on an event tree within the PRA model.

Tools around the industry are being developed to improve control systems and identify the most efficient response to anomalies or method of performing a state transition.

3.3.9 Aging Management

Over the years, the nuclear industry has accumulated a large history of NDE inspections across many inspection programs. AI offers an opportunity to assess this vast amount of operational experience holistically for trends or other insights relating to component aging and condition.

Degradation models that are used to predict failures or the RUL of a component typically rely heavily on physics and expertise, which might have limited capacity to learn from massive measured or simulated data. In other words, when new data are available, it is not easy to improve the model [11].

3.4 TRAINING

AI can be used to expand the use of operator training using a simulator. For example, the Xe-100 plant simulator is being used to create synthetic data that are used to build ML models [21]. Using synthetic data from the Xe-100 simulator it would theoretically be possible to build a surrogate ML model that could mimic the simulator behavior with limited increase in uncertainty.

X-energy noted that Fire Brigades use a lot of plant staff resources, especially for regular training and extra shift duties. Replacing some of that burden with AI models could be beneficial [21].

3.5 MAINTENANCE

Future AI applications will be used to form a unified system or plant AI to provide for increased communication and control capabilities and for supporting scheduling maintenance activities. AI implementation will require that gaps for licensing be identified, and licensees will require assurance of an AI regulatory infrastructure.

Complete benefits analysis is beyond the scope of this document, but the three areas outlined previously are poised to have significant benefit to the commercial nuclear power industry. AI optimization of design and operations will have potentially massive cumulative benefits, possibly reducing overall operational costs by several cumulative percentage points for a plant's lifetime [20]. Facility maintenance optimization also has the potential to dramatically improve plant productivity by decreasing the number and length of offline events, increasing annual power generation revenue. Staff productivity will also have an impact but not as much. Labor costs in the commercial nuclear power industry are dwarfed by big ticket items such as fuel purchases and outages. Saving a week of reactor engineer time will have an impact 400 times less significant than obviating the purchase of one fuel assembly.

3.6 REGULATORY TOOL

Similar to the use of DTs, how AI will be used will determine the regulatory criteria. To minimize requirements, the data connection between the physical asset and the digital model are currently exclusively one-directional and not always in real time [40]. In the broader scope, data connections are to support the interconnection of all components of the AI, including the connection between physical entities and virtual models, between physical entities and data, between physical entities and services, between virtual models and data, between virtual models and services, and between services and data.

Al/ML does have the potential to improve the efficiency and/or effectiveness of nuclear regulatory oversight or otherwise affect regulatory costs associated with safety oversight if diagnostics can reasonably expand in coordination with risk-informed categorization. That is, oversight might be improved by simplification of inspections and standardized rationale for maintenance deferral. This is especially true when regulator Al/ML competencies increase with operating sites [33].

The benefit is that AI/ML/automation will provide the ability to assess compliance on a continual basis. This will provide efficiency for both the licensee and external stakeholders and will enhance open communication paths.

The NRC might benefit from using AI to review plant documentation to identify trends in performance or documentation to more efficiently analyze plant performance and issues. The added efficiency for the NRC would also benefit the nuclear industry by reducing the inspection burden.

3.7 R&D

Optimized sensor placement: (ongoing) Rick Vilim at Argonne National Laboratory gave a talk detailing how they are using ML combined with physics-based models to optimize the placement of sensors in power plant systems [21]. Xenergy is interested in pursuing similar research.

3.8 TESTING

Some of the focus areas where AIs will be applied in the nuclear industry are design and licensing, plant construction, training simulators, predictive O&M, autonomous operation and control, failure and degradation prediction, obtaining insights from historical plant data, and safety and reliability analysis.

4. REGULATORY REQUIREMENTS FOR AI

In terms of regulating AI applications in the US, the NRC is still mostly in an information collection phase, seeking input from stakeholders and holding workshops to better understand the technology and its potential nuclear applications [[41, [42]. However, a report developed by INL and ORNL [40] outlines the technical challenges and gaps in AI-enabling technologies for nuclear reactor applications. When considering AI as applied to NPP O&M, the report summarizes three specific challenges related to licensing and regulatory activities:

1. Quality/optimum input data
2. Identification and selection of appropriate AI algorithms
3. Explainability of the AI algorithms

Regarding the input data, the quality of the data is paramount for nuclear safety applications. Poor data might lead to false output relationships or large uncertainties.

A wide array of different AI algorithms is available to select. The specific application will reduce this array to a short list of suitable algorithms. At this point, the selection of the appropriate algorithm will depend on many factors, such as the desired performance, size, and complexity of the training data, scalability of the algorithm, and deployment and business case criteria (software implementation, legacy solutions, cost, etc.) [1]. Eliminating bias—or proving that this approach is systematic and fits the specific application—is a significant challenge.

Finally, explainability will be critical for regulatory and public acceptance. At every step, the process will require textual and graphical explanations. "Black-box" types of arguments might not adequately explain the underlying processes. The algorithm must generate an explanation in human natural language of (a) what it did and (b) the rationale for what it did. Explaining the performance of the algorithm is equally important. From the output, the analyst or regulator should be able to clearly recognize if the model succeeds better than chance or if a human could achieve a better result.

Similar needs were identified in other ML research activities related to nuclear energy regulatory questions [43]. Sun et al. [43] examined the state of the art in ML for NDE applications in nuclear energy in-service inspection. Key needs identified in this review included data, algorithm selection, and validation. In this context, the term *data needs* refers to the need for relevant and representative reference data sets that might be used in developing and validating ML performance. The needs are not only for sufficient data but also for quality data. The wide range of algorithms available requires a mechanism for identifying the appropriate technique and demonstrating that the selected algorithm is appropriate for the application. Validation of the data used, algorithm selection, and algorithm performance is essential to building the necessary level of confidence for regulatory and public acceptance. Similar studies [44] have led to the development of a methodology for qualifying AI for nuclear energy use that provides guidance on best practices for addressing these various needs.

4.1 REGULATORY FRAMEWORK OUTSIDE THE NUCLEAR INDUSTRY

The consensus seems to be that new regulations might not be needed, and that the solution could be developing guidance documents that include AIs.

Both the US [45] and EU [46] have identified the need for ensuring better conditions for development and use of AI. While the EU has proposed a regulatory framework for this purpose, the US has set out policy considerations that should guide regulatory and non-regulatory approaches to AI developed and deployed outside the federal government. Of relevance to the nuclear power regulations is the guidance from the FAA and NHTSA in that they are working to establish nimble and flexible frameworks that ensure safety while encouraging innovation.

On a national scale, the President's NSTC Committee on Technology's assessment is that long-term concerns about super intelligent general AI should have little impact on current policy [2016] [6]. The NSTC assessed policy requirements for self-driving vehicles, drones, etc. The consensus of the commenters, as part of the White House Future of Artificial Intelligence Initiative, was that broad regulation of AI research or practice would be inadvisable at this time. Instead, commenters said that the goals and structure of existing regulations were sufficient, and commenters called for existing regulation to be adapted as necessary to account for the effects of AI. Commenters suggested that motor vehicle regulation should evolve to account for the anticipated arrival of autonomous vehicles and that the necessary evolution could be carried out within the current structure of vehicle safety regulation. In doing so, agencies must remain mindful of the fundamental purposes and goals of regulation to safeguard the public good, while creating space for innovation and growth in AI.

As part of its digital strategy, the European Union (EU) wants to regulate AI to ensure better conditions for the development and use of this innovative technology [47]. In April 2021, the European Commission proposed the first EU regulatory framework for AI [46]. It says that AI systems that can be used in different applications are analyzed and classified according to the risk they pose to users. The different risk levels will mean more or less regulation. Once approved, these will be the world's first rules on AI.

AI has applications in many products, such as cars and aircraft, which are subject to regulation designed to protect the public from harm and ensure fairness in economic competition. In general, the approach to regulation of AI-enabled products to protect public safety should be informed by assessment of the aspects of risk that the addition of AI may reduce, alongside the aspects of risk that it may increase. It should be recognized that how an AI may fail could be different than how an operator may take an incorrect action. An assessment of the risk would evaluate if the risk for those scenarios are tolerable.

The NSTC recommends that if a risk falls within the bounds of an existing regulatory regime, the policy discussion should start by considering whether the existing regulations already adequately address the risk or whether they need to be adapted to the addition of AI. Also, where regulatory responses to the addition of AI threaten to increase the cost of compliance or slow the development or adoption of beneficial innovations, policymakers should consider how those responses could be adjusted to lower costs and barriers to innovation without adversely impacting safety or market fairness [6].

A relevant example of the regulatory challenges associated with an agency updating legacy regulations to account for new AI-based products is the work of the US Department of Transportation (DOT) on automated vehicles and unmanned aircraft systems (UAS, or "drones") [6]. Within DOT, automated cars are regulated by the National Highway Traffic Safety Administration (NHTSA) and aircraft are regulated by the Federal Aviation Administration (FAA). Realizing the potential benefits of these automated technologies requires these agencies take steps to ensure the safety of the airspace and roads, while continuing to foster a culture of innovation and growth.

Applying techniques of AI in safety-critical environments raises several challenges. First among these is the need to translate human responsibilities while driving or flying into software. Unlike in some other successful applications of narrow AI, there are no concise descriptions for the task of operating ground or air vehicles. While it might seem straightforward to require the AI to obey all traffic laws, a skilled human driver might cross a double-yellow road boundary to avoid an accident or move past a double-parked vehicle. Though such situations might be rare, they cannot be ignored—simple arithmetic dictates that for failures to occur at least as infrequently as they do with human drivers, a system must handle many such rare cases without failure.

4.1.1 Activities Related to AI

Standards Development Organizations and industry are moving forward with guidelines for AI. Many of the efforts have focused on the ethics of AI, and many companies and governments are working to create definitions, policies, and regulations around the ethics. The following is an overview of some of the activities related to AI that might be of interest to the nuclear community:

- The European Commission published *Ethics Guidelines for Trustworthy AI* in April 2019 [48]. The guideline provides general guidelines to achieve a trustworthy AI, two of which are privacy and data governance.

- The European Union Aviation Safety Agency (EASA) published EASA Concept Paper: First usable guidance for Level 1 machine learning applications [3]. This concept paper presents a first set of objectives for Level 1 AI ("assistance to human"), in order to anticipate future EASA guidance and requirements for safety-related ML applications.

- SAE International AIR6988, *Artificial Intelligence in Aeronautical Systems: Statement of Concerns* (2021-04) assesses whether its current standards are compatible with a typical AI and ML development approach [13].

- SAE International AIR6987, *Artificial Intelligence in Aeronautical Systems: Taxonomy*, establishes a comprehensive taxonomy of AI in aviation (currently a work in progress) [49].

- The Institute of Electrical and Electronics Engineers (IEEE) has a global initiative for ethically aligned design of autonomous and intelligent systems to define standards and guidelines for ethical AI [50]. These guidelines cover issues such as human rights, accountability, transparency, and awareness of misuse. The proposed guidelines address system design, transparency of autonomous systems, data privacy, and others.

- The ISO is working on creating standards and addressing issues such as trustworthy AI, safety, and AI applications.

- The IEC is developing a technical report that provides an overview of AI technologies from a nuclear perspective and summaries of potential AI application scenarios in nuclear facilities. The objective of this document is to support future standard development work in this technical area.

OMB M-21-06 [Nov. 2020] [45] sets out policy considerations that should guide regulatory and nonregulatory approaches to AI applications developed and deployed outside of the federal government. Although federal agencies currently use AI in many ways to perform their missions, government use of AI is outside the scope of this memorandum. Although this memorandum uses the definition of AI recently codified in statute, it focuses on "narrow" (also known as "weak") AI, which goes beyond advanced conventional computing to learn and perform domain-specific or specialized tasks by extracting information from data sets or other structured or unstructured sources of information. The memorandum states that

> "When developing regulatory and non-regulatory approaches, agencies should pursue performance-based and flexible approaches that are technology neutral and that do not impose mandates on companies that would harm innovation. Rigid, design-based regulations that attempt to prescribe the technical specifications of AI applications will in most cases be impractical and ineffective."

4.1.2 Examples of Use Cases

AIs are built using software tools and are only now beginning to be used in the design, monitoring, and control in nuclear plants. However, this technology is already being explored or used in process industries, UASs, ocean-going supertankers, oil derricks and ocean drilling platforms, telemedicine and remote patient monitoring, and robotic surgery. Although this list is not exhaustive, it does provide a wide variety of AI definitions and use cases currently being pursued across the nuclear industry. Based on perceived importance from original equipment manufacturers and utility operators, the most highly ranked use cases were determined to be (1) construction sequence simulation, (2) real-time construction sequence optimization, and (3) predictive maintenance. Predictive maintenance use cases were identified as the most important in terms of new value added for sustainable low-cost O&M and for maximizing component life.

The Executive Office of the President's NSTC's Subcommittee on Machine Learning and Artificial Intelligence [6] surveys the current state of AI, its existing and potential applications, and the questions that are raised for society and public policy by progress in AI. The subcommittee's report notes that one area of great optimism about AI is its potential to improve people's lives by helping to solve some of the world's greatest challenges and inefficiencies.

Advances in AI are already providing benefits in the public and private sectors of health care, transportation, the environment, criminal justice, and economic inclusion. For example, at the Walter Reed Medical Center, the Department of Veteran Affairs is using AI to better predict medical complications and improve treatment of severe combat wounds, leading to better patient outcomes, faster healing, and lower costs [6]. The same general approach—predicting complications to enable preventive treatment—has also reduced hospital-acquired infections at Johns Hopkins University [6]. Given the current transition to electronic health records, predictive analysis of health data might play a key role across many health domains like precision medicine and cancer research. In transportation, AI-enabled smarter traffic management applications are reducing wait times, energy use, and emissions by as much as 25% in some places [6]. Cities are now beginning to leverage the type of responsive dispatching and routing used by ride-hailing services and linking it with scheduling and tracking software for public transportation to provide just-in-time access to public transportation that can often be faster, cheaper, and, in many cases, more accessible to the public. The NSTC Committee on Technology's also makes recommendations for specific further actions by federal agencies and other actors [6].

Autonomy is unlikely to have a major transformative impact on nuclear command-and-control systems [5]. There are two reasons for this. First, command-and-control systems already rely on a great degree of automation. Second, the types of algorithms underlying AI-driven applications and complex autonomous systems remain too unpredictable because of the problems of transparency and explainability. As noted by Boulanin [5], "nuclear command-and-control systems are too safety-critical to be left to algorithms that engineers and operators cannot fully understand." Moreover, relatively more traditional rule-based algorithms would be sufficient to further automate command and control. However, even if they are not transformative, advances in AI and autonomous systems could bring some qualitative improvements in the nuclear command-and-control architecture. AI could be used to enhance protection against cyberattacks and jamming attacks. AI could also help planners to more efficiently manage their forces, including their human resources. Similarly, autonomous systems could be used to enhance the resilience of the communications architecture. Long-endurance unmanned aerial vehicles (UAVs) could, for instance, be used to replace signal rockets in forming an alternative airborne communications network in situations where satellite communication is impossible.

For the interested reader, ISO/IEC TR 24030:2021[51] provides a collection of use cases of AI applications in a variety of domains. In total, 132 AI use cases were submitted by experts between July 2018 and the end of November 2019. In this document, the term "use cases" means "collection of submitted use cases."

4.2 REGULATORY FRAMEWORK IN THE NUCLEAR INDUSTRY

The current regulatory framework for nuclear energy in the United States is focused on the use of defense-in-depth measures that provide a reasonable assurance of safety. Generally, such measures include periodic inspection and testing of safety-significant components, with preventive maintenance performed on a time-based schedule to ensure component operability. AIs present new opportunities to risk-inform all aspects of safety and security.

The deployment of AI technology might require additional considerations and requirements such as specific documentary evidence of performance and specific forms of technical data prior to acceptance of safety-significant components. Deployment of AI technologies on balance-of-plant components or on components that are not considered to be safety significant likely will have little or no regulatory restrictions or necessary approvals unless it can affect the operation of a safety system.

As noted, the NRC has issued its strategic plan for AI in NUREG-2261 [7]. This 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, and (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.

As some technical aspects of AI are still evolving, the regulatory framework from nuclear regulators in the US is not yet established. The US is not alone; an IEC working group on AI states that "the regulatory framework from nuclear regulators is not yet established" and recognizes that uses in safety-related applications will require further guidance from a regulatory perspective [11].

The NRC continues to assess whether any regulatory guidance or inspection procedures need to be updated or created to clarify the process and procedure for the licensing and oversight of AI in NRC-regulated activities. The regulatory aspects and guidance will depend on how the AI is used. Although NPP operation was one of the main drivers of the nuclear industry leveraging AI, the broader nuclear scientific and professional community rapidly adopted AI too. AI is now used in reactor design, fuel optimization, intelligent control, preventive maintenance, ageing management, nondestructive testing, physical protection, cybersecurity, and many other related fields.

Effective regulation of technologies such as AI requires agencies to have in-house technical expertise to help guide regulatory decision-making [6]. The need for senior-level expert participation exists at regulating departments and agencies and at all stages of the regulatory process. A range of personnel assignment and exchange models (e.g., hiring authorities) can be used to develop a federal workforce with more diverse perspectives on the current state of technological development.

Current regulations are based on providing reasonable assurance of adequate protection and do not presume a particular technology. As described subsequently, the current regulatory environment regarding the use of I&C and information technologies is based on a heritage of predominately isolated, independent, hardwired systems based on analog technology that provide protection, control, and monitoring functions. Over the past four decades, the application of digital technologies and the integration of data and functionality has proceeded slowly for the nuclear power industry. For existing plants, implementation of digital technology has progressed through piecemeal upgrades of standalone systems. In response, the NRC developed additional guidelines and endorsed acceptable practices to address the unique behavior characteristics of software-based digital systems. Meanwhile, the nuclear power industry has developed and begun construction of ALWRs incorporating more comprehensive use of digital technologies to provide highly reliable integrated command and (data) communications capabilities. Additionally, ARs, small modular reactors (SMRs), and microreactors have been and continue to be designed based on greater degrees of automation and optimal asset management (e.g., PMx) to optimize human resource use and minimize O&M cost profiles. The use of AIs and the potential application of AI techniques follow from this trend.

The regulatory treatment of AIs will depend not only on whether they are to be used to support safety systems but also on their embedded functionality. In this report, the functional roles for AIs are grouped into categories that are generally consistent with NRC characterizations of levels of automation. The main categories are identified as (1) noncontrol functionality (advisory), (2) control functionality (shared), and (3) communications functionality (generally advisory).

- Regarding the noncontrol functionality category, the role of the AI would be to provide information or advice to the operator but not to directly affect the plant or its operation. The main issues involve quality, correctness, and fidelity of the information provided to the operator (i.e., related to the trustworthiness of the information), and the potential risk to plant safety arising from deployment and reliance on these techniques. Thus, review might address whether proper performance, determination of uncertainty, and transparency of the basis for the information have been demonstrated. Consequently, the depth of the review would depend on the impact of erroneous or uncertain performance on safety and human reliability.

- Regarding the control functionality, the consideration is the degree to which responsibility for actions is shared between the operator and AI. This could range from the boundaries of manual control with the AI advising of expected responses to potential control actions to autonomous control using an embedded AI for predictive control and/or automatic adaptation. The extreme end of autonomy is not anticipated as a near-term application of AIs. The range of control functionality introduced into an AI would lead to more rigorous regulatory review and a greater level of evidence on the safety impact of the system. Plants with high degrees of passive safety would seem to be good candidates for implementing AI.

- Regarding the communications functionality, the level of regulatory review would depend on the safety significance of the data being transmitted. If vital communication of safety-related data is involved, then the communications functionality provided by the AI would necessarily be subject to safety or safety-related review. This would include independence, isolation, reliability, fault-tolerance/accommodation, and so forth. If the communications functionality is solely advisory or nonvital, then the review would be similar to that of other noncontrol functionality.

Another characteristic to consider in anticipating the regulatory considerations for use of an AI involves complexity, which directly relates to the capability to predict its behavior under normal and faulted conditions and to its vulnerability to change by accident or incursion. The significance of these characteristics is tied to the functionality implemented in the AI and its potential impact on safety.

A risk-informed/graded approach based on the functionality of the AI would evaluate the complexity of the AI with its function to set the level of review sufficient to reach a safety conclusion. Simple monitoring, the lowest complexity level, minimal consequences of failure, and other such aspects would not require guidance on software tools or type of digital device. At the other end of the graded approach, existing guidance will apply to control those AIs that perform a control function, while more scrutiny would result for components that perform safety functions. Thus, an understanding of the AI's functionality can be applied to facilitate a graded approach to qualification, testing, and inspections.

4.2.1 General Requirements for I&C

The regulatory framework for the NRC is to protect public health and safety. The NRC's mission is to ensure the safe use of radioactive materials for beneficial civilian purposes while protecting people and the environment. The NRC regulates commercial NPPs and other uses of nuclear materials through licensing, inspection, and enforcement of its requirements.

NRC licensing reviews are based on an applicant's design meeting its performance assessment based on (1) safety goals and objectives, (2) deterministic and/or probabilistic analysis of accident scenarios, and (3) quantitative assessment of design alternatives against the safety goals and objectives using accepted engineering tools, methodologies, and performance criteria. Because implementing AI technology will require the use of a digital platform, it must meet the requirements of an instrumentation and control (I&C) system.

At a high level, a malfunction or failure of any system cannot prevent/block a safety action or initiate a challenge to that system.[3] The fundamental design principles for an I&C system at an NPP are as follows:

- Redundancy (single failure criteria)
- Quality (especially software quality assurance)
- System integrity (determinism)
- Independence (physical, electrical, communication)
- Diversity and defense-in-depth

[3] This is an example of a control system failure that, although it did not prevent/block a safety action, resulted in a challenge to a safety system and a reactor scram. On May 10, 1996, Browns Ferry Unit 2 experienced an automatic reactor scram on low reactor water level from full power (LER 260-96-005-00). When software parameter changes were made active (saved) in the control system, a reinitialization sequence occurred within the control software block, which drove the feed pump speed demand signal to zero for a few seconds. Plant personnel were unaware that entering these new software parameters would cause the feedwater control system to reinitialize. The cause of the event was attributed to inadequate design of the control system software. A design weakness existed in the installed system, in that making software parameter changes in certain software blocks would cause the control system to automatically reinitialize to zero output. This characteristic of the software design was not known to plant personnel.

- Environmental qualification
- Reliability
- Simplicity
- Control of access

In addition to the aforementioned fundamental design principles, operators must have a diverse means of seeing current and reliable values of essential reactor operating parameters. That is, operators must be able to trust the information provided and to understand the current state of the plant.

The objectives of the control systems are to maintain the controlled variables within prescribed operating ranges, and the effects of operation or failure of these control systems are bounded by the accident analyses in Chapter 15 of the safety analysis report.

AI is an emerging, analytical tool, which, if used properly, shows promise in its ability to improve reactor safety yet offer economic savings. The NRC received 10 responses to its request for comments [24] on 11 issues to enhance the NRC's understanding of the short- and long-term applications of AI and ML in nuclear power industry operations and management, as well as potential pitfalls and challenges associated with their application. Because the list of issues and responses were useful in identifying potential applications of AI in NPPs, they are repeated here:

1. What is the status of the commercial nuclear power industry development or use of AI tools to improve aspects of nuclear plant design, operations or maintenance or decommissioning? What tools are being used or developed? When are the tools currently under development expected to be put into use?

2. What areas of commercial nuclear reactor operation and management will benefit the most, and the least, from the implementation of AI? Possible examples include, but are not limited to, inspection support, incident response, power generation, cybersecurity, predictive maintenance, safety/risk assessment, system and component performance monitoring, operational/maintenance efficiency and shutdown management.

3. What are the potential benefits to commercial nuclear power operations of incorporating AI in terms of (a) design or operational automation, (b) preventive maintenance trending, and (c) improved reactor operations staff productivity?

4. What AI methods are either currently being used or will be in the near future in commercial nuclear plant management and operations? Example of possible AI methods include, but are not limited to, artificial neural networks, decision trees, random forests, support vector machines, clustering algorithms, dimensionality reduction algorithms, data mining and content analytics tools, gaussian processes, Bayesian methods, natural language processing, and image digitization.

5. What are the advantages or disadvantages of a high-level, top-down strategic goal for developing and implementing AI across a wide spectrum of general applications versus an ad-hoc, case-by-case targeted approach?

6. With respect to AI, what phase of technology adoption is the commercial nuclear power industry currently experiencing and why? The current technology adoption model characterizes phases into categories such as: the innovator phase, the early adopter phase, the early majority phase, the late majority phase, and the laggard phase.

7. What challenges are involved in balancing the costs associated with the development and application of AI, against plant operational and engineering benefits when integrating AI applications into operational decision-making and workflow management?

8. What is the general level of AI expertise in the commercial nuclear power industry (e.g. expert, well-versed/skilled, or beginner)?

9. How will AI effect the commercial nuclear power industry in terms of efficiency, costs, and competitive positioning in comparison to other power generation sources?

10. Does AI have the potential to improve the efficiency and/or effectiveness of nuclear regulatory oversight or otherwise affect regulatory costs associated with safety oversight? If so, in what ways?

11. AI typically necessitates the creation, transfer and evaluation of very large amounts of data. What concerns, if any, exist regarding data security in relation to proprietary nuclear plant operating experience and design information that may be stored in remote, offsite networks?

This report describes potential common frameworks and AI applications as a starting point for meeting that objective to foster better understanding and adoption of AI technologies within nuclear facilities. Regulatory requirements for AI begin with the Atomic Energy Act (AEA) as amended (Figure 55), which licensees must meet. The AEA remains the primary authority for the NRC's implementing of regulations.

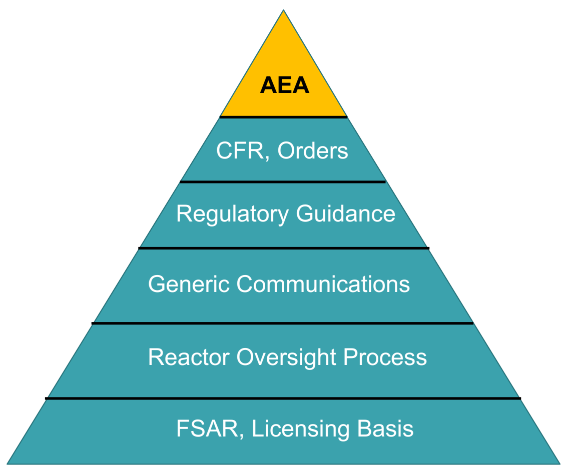


Figure 5. Hierarchy of regulations for AI.

The AEA remains the primary authority for the NRC's implementing regulations. 10 CFR and commission orders are rules for NRC staff to follow to implement the law. 10 CFR Parts 50 and 52 provide the current rules for plant licensing; Part 53 is under development, as detailed subsequently.

Part 50, "Domestic Licensing of Production and Utilization Facilities," is a two-step licensing process including: (1) issuance of licenses and construction permits and (2) issuance of an operating license. Part 50 is for the licensing of production and utilization facilities.

Part 52, "Licenses, Certifications, and Approvals for Nuclear Power Plants," governs the issuance of early site permits, standard design certifications, combined licenses, standard design approvals, and manufacturing licenses for nuclear power facilities. This is a *combined operating license* (COL) that provides for approval of a combined construction permit and operating license. Many sections of Part 52 require compliance with sections of Part 50, including compliance with all Appendices in 10 CFR 50, including the General Design Criterion (GDC) in Appendix A.

Part 53, “Risk Informed, Technology Inclusive Regulatory Framework for Advanced Reactors,” is under development.

Guidance on satisfying the rules is provided in the Standard Review Plan (NUREG-0800) and in NRC regulatory guides (RGs). Once a design is approved, a plant’s licensing basis is the plant’s legal authorization for design, construction, maintenance, and operation.

Relevant regulatory criteria for an I&C system using AI will vary depending on its use. Applicable parts of the CFR for I&C systems at NPPs are provided here:

- 10 CFR 50.34—Includes information describing the facility, presents the design bases and the limits on operation, and presents a safety analysis of the SSCs and of the facility as a whole.
  - The NRC generally invokes three categories of consequences: two are in regulations 10 CFR 20 and 10 CFR 50.34.
  - 10 CFR 50.34(c) and 10 CFR 52.79(a) specify the content of license applications under varying licensing paths; both specify that a cybersecurity plan is required as set forth in 10 CFR 73.54.
  - This part is applicable to simulators and cybersecurity.
- 10 CFR 50.49—Requires each licensee to establish a program for qualifying all “electric equipment important to safety.”
  - Electric equipment important to safety includes safety-related electric equipment and nonsafety-related electric equipment whose failure under postulated environmental conditions could prevent satisfactory accomplishment of safety functions.
- 10 CFR 50.54(jj) and 50.55(i)—SSCs subject to the codes and standards in 10 CFR 50.55a must be designed, fabricated, erected, constructed, tested, and inspected to quality standards commensurate with the importance of the safety function to be performed.
  - The applicant must provide sufficient information and calculations (if applicable) that the design of the RTS and RTS initiation of auxiliary supporting features and other auxiliary features are acceptable.
  - This part is applicable to classification and quality standards.
- 10 CFR 50.55a(h)(2)—Requires compliance with IEEE 603-1991 and the correction sheet dated January 30, 1995.
  - This part is applicable to classification based on the system functions’ ability to ensure safety, ML, and AI.
- 10 CFR 50.65—The scope of the monitoring program will include safety and some nonsafety SSCs.
  - Some NPP licensees are in the process of demonstrating new approaches (e.g., NEI 18-10) for meeting regulatory requirements in 10 CFR 50.65.
  - The new approach in NEI 18-10 is a departure from the current preventative maintenance assessment paradigm (e.g., establishing SSC performance criteria) and is intended to allow for a more dynamic assessment of maintenance effectiveness based on the use of data and risk trending analytics.
  - This part is applicable to monitoring.
- 10 CFR 50.69—Risk-informed classification.
  - This part allows a risk-informed approach to safety classification.
- CFR Part 50, Appendix A, General Design Criterion (GDC)—Establishes the minimum requirements for the principal design criteria.
  - GDC 1, “Quality Standards and Records”

    - This part is applicable to the classification of components and controls and that they be designed, fabricated, erected, and tested to quality standards commensurate with the importance of the safety functions to be performed.

  - GDC 10, “Reactor Design”
  - GDC 13, “Instrumentation and Control”
    - This part is applicable to controls.

  - GDC 15, “Reactor Coolant System Design”
  - GDC 19, “Control Room”
  - GDC 21, “Protection System Reliability and Testability”
    - This part is applicable to classification of components.

  - GDC 24, “Separation of Protection and Control Systems”
    - AI cannot cause failure of the protection system.

  - GDC 28, “Reactivity Limits”
  - GDC 29, “Protection against Anticipated Operational Occurrences”
  - GDC 44, “Cooling Water”
- 10 CFR 50, Appendix B, Quality Assurance Criteria—Those actions necessary to provide adequate confidence that an SSC will perform satisfactorily.
  - The criteria in Appendix B apply to all activities—such as design, purchase, installation, testing, operation, maintenance, or modification—that affect the safety-related functions of such SSCs.
  - This part is applicable for classification and AI.
- 10 CFR 52.47(b)(1)—A design certification (DC) application must contain the proposed inspection, test, analysis, and acceptance criteria (ITAAC) that are necessary and sufficient to provide reasonable assurance that a plant that incorporates the design certification is built and will operate in accordance with the design certification, the provisions of the AEA, and NRC regulations.
  - The requirements regarding ITAAC for DC, manufacturing license, and COL applications are set forth in 10 CFR 52.47, 52.80, and 52.158.
- 10 CFR 52.80(a)—A COL application must contain the proposed ITAAC and that if the inspections, tests, and analyses are performed and the acceptance criteria are met, then the facility has been constructed and will operate in conformity with the COL, the provisions of the AEA, and NRC regulations.
  - The requirements regarding ITAAC for DC, manufacturing license, and COL applications are set forth in 10 CFR 52.47, 52.80, and 52.158.

At the highest conceptual level, I&C systems in an NPP can be categorized by safety (protection) and nonsafety (control) systems. The RPS includes the RTS and the ESFs. The RTS is designed to automatically initiate the reactivity control system (control rods) to ensure that specified acceptable fuel design limits are not exceeded. Because the safety systems are important to protecting public health and safety, they have strict requirements on their design, construction, and operation, and they receive the NRC’s most stringent review.

If the AI is to provide a protection system function, then it must meet the requirements of a safety system. Relevant acceptance criteria are based on meeting the relevant requirements for an I&C *protection system*.

4.2.2 Regulatory Requirements for AI

The licensing process is a performance-based approach that requires licensees to show that they meet the safety and security requirements for their facility. Thus, the addition of AI must still meet these performance-based criteria. The licensing process is also primarily risk-based with the identification of components and systems as nonsafety, important to safety, or safety related. A risk-informed approach allows further gradation of components and systems based on risk metrics such as core damage frequency or large early release fractions. Thus, the use cases will determine the regulatory requirements. For example, the use of AI in a nonsafety system or solely for data processing could be outside the licensing basis unless its failure (or faulty information) can affect the operation of a safety system.

The regulatory requirements for AI, which will be incorporated into the I&C system, will be very dependent on how it is used (i.e., its functionality, safety classification, etc.). The licensing process is primarily risk-based with the identification of components and systems as nonsafety, important to safety, or safety related. A risk-informed approach allows further gradation of components and systems based on risk metrics such as core damage frequency or large early release fractions. Thus, the use cases will determine the regulatory requirements. Regardless of how the AI is used it presents new opportunities for risk-informing operating, maintenance, and regulatory decisions. Trustworthiness, transparency, and the ability to validate and verify the results are paramount in showing that the systems and plant still meet their performance requirements.

Industry and regulators must develop agreed-upon guidance and frameworks for acceptance of AI applications that are consistent, explicit, and that enable the use of AIs as an additional avenue to meet the intent of existing regulations.

AI applied to nuclear operations falls within the category of digital I&C because such applications involve digital computer hardware and custom-designed software that input plant data, execute complex software algorithms, and output the results to a system or licensed human operator to potentially provoke an action. Because of the uniqueness of AI characteristics in comparison with typical digital I&C systems, specific considerations need to be taken for AI to fit a regulatory-controlled nuclear action.

The conceptual features of AIs are varied. However, the identification of the regulations applicable to the specific uses of AIs have not been established. Regulations must be identified so that the next essential step to develop AIs can be taken. The regulations will differ if the AI has a safety, important-to-safety, or nonsafety function. For example, if the AI is used as a construction tool and is not part of the licensing basis it would not be subject to regulatory requirements. By identifying the regulations that might apply to an AI (which will be based on its application), it might be possible to identify whether existing regulations are sufficient and could be adapted or considered to accommodate AIs for ARs or whether new regulations and guidance are needed.

The NRC recognizes that interest in AI is growing rapidly in both the public and private sectors and anticipates increased use of AI in NRC-regulated activities. NRC’s strategic plan focuses on a broad spectrum of AI subspecialties (e.g., NLP, ML, DL) that could encompass various algorithms and application examples that 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. NRC’s strategic plan focuses on AI and data science. NRC’s guidance will have to be different for each use case.

NUREG-2261 [7] provides the NRC’s strategic plan to ensure the agency’s readiness to review uses of AI. This strategic plan focuses on a broad spectrum of AI subspecialties (e.g., NLP, ML, DL) that could encompass various algorithms and application examples that the NRC has not previously reviewed and evaluated. As part of its strategic plan, the NRC will assess whether any regulatory guidance (e.g., RGs or standard review plan sections) or inspection procedures need to be updated or created to clarify the process and procedure for the licensing and oversight of AI in NRC-regulated activities.

At a higher level, the intended use of AI will determine its requirements and regulatory criteria. A summary of regulations based on functionality is provided here:

- Regarding the ***use of AIs for operations support***, the safety classification of the functions provided or supported by the AI and the nature of its role in operations would be key factors in determining the regulations that apply (e.g., safety, safety-related).

- If no protection functions will rely on or be incorporated into an AI, it is likely that classification of AIs employed for operations support might be nonsafety. Nevertheless, even a system that does not involve a safety or safety-related function must be evaluated to show that its failure would not compromise a safety function.

- If used in a noncontrol function (e.g., monitor and display), the role of the AI would be to provide information or advice to the operator but not to directly affect the plant or its operation. The main issues involve quality, correctness, and fidelity of the information provided to the operator (i.e., related to the trustworthiness of the information) and the potential risk to plant safety arising from deployment and reliance on these techniques.

- If used in a control function, the regulatory consideration is the degree to which responsibility for actions is shared between the operator and AI. This could range from the boundaries of manual control with the AI advising of expected responses to potential control actions to autonomous control using an embedded AI for predictive control and/or automatic adaptation.

- If communications functions are added to the AI, the level of regulatory review would depend on the safety significance of the data being transmitted and whether two-way communication capabilities are present.

Any AI must be assessed against licensing requirements. Presentations at the various workshops recognize that the licensing requirements for AIs must be met but do not provide any specifics rather than to indicate that software, cybersecurity, and requirements in general are needed. In summarizing the working group's work to participants, Roland and Guerra stated that the "IAEA in collaboration with other organizations should develop [a] SMR [small modular reactor] design roadmap to safety/security requirements with respect to regulatory requirements. This should include differences in requirements, such as dose thresholds (e.g., variations of URC [unacceptable radiological consequence] between countries)" [52].

Specifics on the regulations applicable to the use of AIs are needed and, again, will vary based on use.

- **10 CFR 50.34**—Includes information describing the facility, presents the design bases and the limits on operation, and presents a safety analysis of the SSCs and of the facility as a whole.
  - The consequences resulting in the action or inaction of an AI on plant systems and the consequences of the failure on those systems must be assessed.
  - This part is applicable to simulators, cybersecurity

- **10 CFR 50.36 and 52(ii)(30)**—Applies if an AI is used during the design phase to provide insights into TS limits.
  - As noted in the regulation, the "TS will be derived from the analyses and evaluation included in the safety analysis report."
  - Highly autonomous reactor designs must consider design safety limits, limiting safety settings, and limiting control settings based on these analyses.
  - This part is applicable to technical specifications and changes to equipment functionality.

- **10 CFR 50.46**—The cooling performance of an AR must be calculated in accordance with an acceptable evaluation model. It must include sufficient supporting justification to show that the analytical technique realistically describes the behavior, and uncertainty must be accounted for to show that there is a high level of probability that the criteria would not be exceeded.
  - If the AI is used to support analysis used in a safety analysis/calculation, then it is governed by 10 CFR 50.46, 10 CFR 50, Appendix K, and some form of software V&V.
  - This part is applicable to AI, technical specifications.

- **10 CFR 50.47**—The use of an AI could support reducing the number of operators and other O&M staffing levels.
  - Current regulations, as written, do not provide for reducing the number of licensed operating staff for cases with or without AIs.
  - This part is applicable to impact on staffing.

- **10 CFR 50.54**—An AI with autonomous control will not affect the number of licensed operators required based on current regulations.
  - The requirements of 10 CFR 50.54(m) tend to prohibit reducing the operating staff by taking credit for a reactor design with a highly autonomous AI.
  - Regulations and guidance appear to lag the state-of-the-art use of Ais.
  - This part is applicable to impact on staffing, manipulation of controls, event notification, and technical specifications.

- **10 CFR 50.55a(h)(2)** (IEEE 603-1991)—Independence must be maintained between safety systems and other systems.
  - AI is an umbrella term referring to any computer algorithm that makes decisions intended to mimic or replace those made by a human.
  - ML is a subset of AI and is a collection of computer algorithms that learn to make decisions based on the observation of training data.
  - This part is applicable to classification, quality, AI, control of access, common-cause failures, and plant integration (ITAAC).

- **10 CFR 50.59**—Licensees could backfit an AI into the control system.

- **10 CFR 50.65**—An AI could be used to monitor the effectiveness of maintenance.
  - This part is applicable to monitoring and maintenance effectiveness.

- **10 CFR 50.90**—If an AI cannot be installed under 50.59, then a licensee must submit a license amendment request.

- **10 CFR 50, Appendix B**—The quality assurance criteria are provided to ensure that those actions necessary to provide adequate confidence that an SSC will perform satisfactorily are taken.
  - The criteria in Appendix B apply to all activities—such as design, purchase, installation, testing, operation, maintenance, or modification—that affect the safety-related functions of such SSCs.
  - If the AI will be used for a safety function, then it must meet the requirements of a safety-related I&C system.
  - If the AI is classified as safety related or is important to safety, it must follow the guidance for safety-related software.
  - This part is applicable for classification and AI.

- **10 CFR 50, Appendix K**—Part II sets forth the documentation requirements for each evaluation model.
    - For a new model that has not been previously reviewed, the NRC reviewers will initiate an evaluation of the entire analytical models and computer codes used by licensees to analyze accident and transient behavior.
    - This part is applicable to AI and technical specifications.

- **10 CFR 52**—DC and COL applications must contain proposed ITAAC.
    - This part is applicable to cybersecurity and plant integration (ITAAC).

- **10 CFR 55**—Licensed operators must be fully aware of any manipulation of reactor controls.
    - Regulations require that licensed operators be continuously present at the controls.
    - The requirement for prior licensed operator knowledge and consent of reactivity and power level changes via a highly autonomous AI must be explored in detail; this approach magnifies the importance of the operators having trust in the AI.

    - This part is applicable to simulators, number of licensed operators, and manipulation of controls.

- **10 CFR 55.46**—An AI would have increased scope and fidelity compared with a simulator used for operator training.
    - This part is applicable to simulators, number of licensed operators, and manipulation of controls.

- **10 CFR 72**—Notification of the declaration of an emergency class and notification of many other events must be made in a timely manner.
    - A highly autonomous reactor must be designed so that (1) notifications are made automatically as required or (2) operating staff are apprised of notifications that must be made in a timely manner.
    - This part is applicable to event notification.

- **10 CFR 73**—There must be a high assurance that systems and networks are adequately protected against cyberattacks.
    - Cybersecurity will be an important consideration for any highly autonomous reactor design to demonstrate adequate protection of the health and safety.
    - Providing appropriate cybersecurity will be complicated if the design implements an off-site control room to support a remotely sited reactor.
    - This part is applicable to cybersecurity.

- **10 CFR 110 (NRC)**—An AI would most likely be subject to export control. § 110.8 provides a list of nuclear facilities and equipment under NRC export licensing authority; Appendix A provides an illustrative list of nuclear reactor equipment under NRC export licensing authority.

- **15 CFR 730-774 (DOC)**—A family of export control classification numbers related to simulators for NPPs (2A291/2D290/2E001). The "Software" "specially designed" or modified for the "development," "production," or "use" of items controlled by 2A290 or 2A291 are export controlled.
    - A full-scope simulator provides the capability to train plant operators in a replica control room that represents the control consoles, control panels, and displays in the plant control rooms.

- **10 CFR 810 (DOE)**—Restricts the transfer of technology for the development, production, or use of equipment or material especially designed or prepared for any of the activities listed in 10 CFR 810.2(b). This part does not apply to exports authorized by the NRC, US Department of State, or DOC.

The applicability of the regulations will depend on how the AI is to be used, and the specific requirements for an AI will be dependent on its functionality. All of the requirements outlined in this report are not applicable to all potential uses of AIs and will depend on its use as a control system, a protection system, or input to determining safety system settings. For example, if the AI is used during the design phase to provide insights into TS limits, then it could be subject to 10 CFR 50.36.

The hardware and development of the software for the AI, regardless of its functionality, must be of sufficient quality commensurate with the importance of the function(s) to be performed. QA become more critical for AIs that use AI and ML. It is important that this is applied throughout the software lifecycle.

5. CONCLUSIONS

AI technology is constantly evolving with new types of explainability methods and algorithms. AIs are being used in industry and will be used more in NPPs. The current regulatory framework does not explicitly address AI or autonomous control. However, because AI applied to nuclear operations falls within the category of digital I&C, such applications will have to meet the basic requirements for those systems and will depend on the use cases of AI. The regulatory requirements for AI will be directly dependent on how it is to be used (i.e., its functionality).

Regardless of how the AI is used it presents new opportunities for risk-informing operating, maintenance, and regulatory decisions. Trustworthiness, transparency, and the ability to validate and verify the results are paramount in showing that the systems and plant still meet their performance requirements.

It is expected that AI, when first applied in NPPs, will mainly play a recommendatory role and is unlikely to be directly used in important and safety-related scenarios.

ARs might be at an advantage because they can be designed with the use of AI; however, a disadvantage for ARs is that they do not have operational data and costs that can be compared.

ARs, SMRs, and microreactors have been and continue to be designed based on greater degrees of automation and optimal asset management (e.g., PMx) to optimize human resource use and minimize O&M cost profiles. The use of AIs and potential application of AI techniques follow from this trend.

Once the data that support ML models become outdated, so too do the models—a problem known as “data drift.”

The deployment of AI technology might require additional considerations and requirements such as specific documentary evidence of performance, particular forms of technical data, and so on, prior to acceptance in modes that are part of or could impact safety-significant systems and components. The information on regulations and review processes captured with this report establish the regulatory framework within which AIs will be evaluated for approval.

The development of an AI might be export-controlled information, and its development and operation should be performed within these constraints.

To summarize, the current regulatory framework does not explicitly address AIs or autonomous control. Assessment of regulations and current practices leads to the conclusion that regulatory requirements for AI will be very dependent on how it is used (i.e., its functionality). In many anticipated cases, the AI might receive limited regulatory attention. However, as the safety classification, operational functionality, or importance of the AI’s information or performance increase in significance, the regulatory burden associated with a safety justification will grow. The capability to classify the AI by safety significance, functionality, complexity, and other characteristics can provide the basis for a graded approach to implementing or licensing such technologies.

To provide a means of compliance for the certification of AI within *safety* critical aeronautical systems, SAE International reviewed existing standards and performed a gap analysis to understand how and why existing standards cannot be reliably used [13]. Among the main gaps identified, requirements traceability, mapping of ML model functions and parameters between aerospace engineering concerns, and the application or lack of verification methods suitable for data sets were the gaps that raised many concerns. The identified gaps highlight that a data-driven paradigm for AI might not be adequately addressed by existing standards. Lastly, as the field of AI in aerospace matures, the SAE committee notes that [49]

> "Extending this licensing procedure [of training and extensive testing for pilots and air traffic control] to autonomous software would lead to an analogous system of gained trust. Certification would be eventually attained through extensive, though not comprehensive, demonstration of knowledge and skill by the advanced software systems."

Any use of AI at NPPs is focused on the nonsafety related applications. The NRC and other regulatory bodies are evaluating providing guidance to address gaps rather than new regulations to address the use of AI and ML. This approach seems to be the best approach to encourage its development without adding regulatory uncertainty. Conversely, SAE International recognizes that as the field of AI in aerospace matures, there will be a need to address gaps not only in industry standards but also in regulatory standards. However, the use of AI in aviation is concerned with certification of AI within safety critical aeronautical systems.

6. REFERENCES

1. M. D. Muhlheim, P. Ramuhalli, A. Huning, A. Guler, R. Wood, and A. Saxena, *Status Report on Regulatory Criteria Applicable to the Use of Digital Twins*, ORNL/SPR-2022/2493, June 2022.
2. European Union Aviation Safety Agency (EASA), "EASA Concept Paper: First usable guidance for Level 1 machine learning applications," December 2021.
3. "The History of Artificial Intelligence History of Computing," CSEP 590A ,University of Washington, December 2006.
4. A. M. Turing, Alan. 1950. Computing Machinery and Intelligence. *Mind 49,* 433 – 460.
5. *The Impact Of Artificial Intelligence On Strategic Stability And Nuclear Risk,* Volume I, Euro-Atlantic Perspectives, edited by Vincent Boulanin, May 2019.
6. Executive Office of the President National Science and Technology Council Committee on Technology, "Preparing for The Future of Artificial Intelligence," National Science and Technology Council, October 2016.
7. NUREG-2261, *Artificial Intelligence Strategic Plan, Fiscal Years 2023-2027*, May 2023. (NRC ADAMS Accession No. ML23132A305)
8. Copeland, B.J., "artificial intelligence". Encyclopedia Britannica, Invalid Date, https://www.britannica.com/technology/artificial-intelligence. Accessed 15 September 2023.
9. AFRL, "Artificial Intelligence." https://afresearchlab.com/technology/artificial-intelligence/ [accessed September 14, 2023]
10. Richa Bhatia, "Understanding the difference between Symbolic AI & Non Symbolic AI," *Endless Origins*, December 27, 2017. https://analyticsindiamag.com/understanding-difference-symbolic-ai-non-symbolic-ai/
11. IEC TR-63468, *Nuclear facilities – Instrumentation and control, and electrical power systems – Artificial Intelligence applications*, 2023-05.

12. OMB EO 13960, "Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government," December 8. 2020.
13. AIR6988, *Artificial Intelligence in Aeronautical Systems: Statement of Concerns*, SAE International, 2021-04.
14. 2020 State of Data Science. https://www.anaconda.com/resources/whitepapers/state-of-data-science-2020/
15. ISO/IEC 22989:2022, Information technology — Artificial intelligence — Artificial intelligence concepts and terminology, 2022-07.
16. T. Sheridan, Humans and Automation: System Design and Research Issues, New York: Wiley & Sons, 2002.
17. J. M. O'Hara and S. Fleger, *Human-System Interface Design Review Guidelines*, NUREG-0700, Rev. 3, July 2020 (ADAMS Accession no. ML20162A214).
18. J. M. O'Hara, J. C. Higgins, S. A. Fleger, and P. A. Pieringer, *Human Factors Engineering Program Review Model*, NUREG-0711, Rev. 3, November 2012 (ADAMS Accession No. ML12324A013).
19. NUREG-0800, Appendix 18-A, Rev. 0, Crediting Manual Operator Actions In Diversity And Defense-In-Depth Analyses, April 2014 (ADAMS Accession No. ML13115A156).
20. Insight Enterprises, Inc., Comment On: NRC-2021-0048-0001, Role of Artificial Intelligence Tools in U.S. Commercial Nuclear Power Operations, May 7, 2021. [NRC ADAMS Accession No. ML21145A367]
21. X-energy, Comment On: NRC-2021-0048-0001, Role of Artificial Intelligence Tools in U.S. Commercial Nuclear Power Operations, May 21, 2021. [NRC ADAMS Accession No. ML21145A366]
22. Blue Wave AI Labs, Comment On: NRC-2021-0048-0001, Role of Artificial Intelligence Tools in U.S. Commercial Nuclear Power Operations, [NRC ADAMS Accession No. ML21145A364]
23. Nuclear Energy Institute, Comment On: NRC-2021-0048-0001Role of Artificial Intelligence Tools in U.S. Commercial Nuclear Power Operations, May 21, 2021. [NRC ADAMS Accession No. ML21145A369]
24. Docket ID NRC-2021-0048, "Role of Artificial Intelligence Tools in U.S. Commercial Nuclear Power Operations," April 21, 2021. [NRC DAMS Accession No. ML21085A611]
25. P. Ramuhalli et al., PNNL-24377R0, "Component-Level Prognostics Health Management Framework for Passive Components Advanced Reactor Technology Milestone," M2AT-15PN2301043, June 2015.
26. P. Ramuhalli, C. Walker, V. Agarwal, N. Lybeck, Development of Prognostic Models Using Plant Asset Data, ORNL, ORNL/TM-2020/1697, 2020.
27. A. K. Sleiti, J. S. Kapat, and L. Vesely, "Digital Twin in Energy Industry: Proposed Robust Digital Twin for Power Plant and Other Complex Capital-Intensive Large Engineering Systems," *Energy Rep.* 8, pp. 3704–3726 (2022). https://doi.org/10.1016/j.egyr.2022.02.305 [Accessed: Jun. 10, 2020].
28. Y. You, C. Chen, F. Hu, Y. Liu, and Z. Ji, "Advances of Digital Twins for Predictive Maintenance," *Procedia Comput. Sci.* 200, pp. 1471–1480 (2022). https://doi.org/10.1016/j.procs.2022.01.348 [Accessed: Jun. 10, 2020].
29. J. Browning, A. Slaughter, R. Kunz, J. Hansel, B. Rolston, K. Wilsdon, A. Pluth, and D. McCardell, "Foundations for a Fission Battery Digital Twin," *Nucl. Technol.* (2022).
30. U. Rohatgi, *Machine Learning-based Prediction of Departure from Nuclear Boiling Power for the PSBT Benchmark*, Brookhaven National Laboratory, BNL-222878-2022-COPA (June 2022).
31. P. A. Grechanuk, M. E. Rising, and T. S. Palmer, "Application of Machine Learning Algorithms to Identify Problematic Nuclear Data," *Nucl. Sci. Eng.*, 195, pp. 1265–1278 (2021). https://doi.org/10.1080/00295450.2021.2011574 [Accessed: Jun. 10, 2020].
32. Z. Ma, H. Bao, S. Zhang, M. Xian, and A. Mack, "Exploring Advanced Computational Tools and Techniques with Artificial Intelligence and Machine Learning in Operating Nuclear Plants," NUREG/CR-7294, February 2022 (ADAMS Accession No. ML22042A662).

33. Letter from Gary Peters, Framatome Inc. to US NRC, Role of Artificial Intelligence Tools in U.S. Commercial Nuclear Power Operations [NRC-2021-0048], May 21, 2021. [NRC ADAMS Accession No. ML21153A056]
34. EPRI 3002021067, *Quick Insight Brief: Leveraging Artificial Intelligence for the Nuclear Energy Sector*, April 13, 2021.
35. EPRI 3002020120, *Application of Data Analytics to Mine Nuclear Plant Maintenance Data*, March 26, 2021.
36. EPRI 3002019609, *Quick Insight — Power Industry Dictionary for Text-Mining and Natural Language Processing Application: A Proof of Concept*, September 11, 2020.
37. EPRI 3002019731, *Using Machine Learning to Predict Flow-Accelerated Corrosion Wear Rates*, February 24, 2021.
38. Letter from Gregory A. Banyay, Westinghouse to US NRC, Role of Artificial Intelligence Tools in U.S. Commercial Nuclear Power Operations, May 31, 2021. [NRC ADAMS Accession No. ML21202A180]
39. A. Huning et al., *Digital Platform-Informed Certification of Components Derived from Advanced Manufacturing Technologies*," ORNL, ORNL/TM-2021/2210 (September 2021).
40. V. Yadav et al., "The State of Technology of Application of Digital Twins," TLR/RES-DE-REB-2021-01, June 2021.
41. NRC, "Virtual Workshop on Digital Twin Applications for Advanced Nuclear Technologies," Virtual Workshop, December 1–4, 2020, Adams Accession No. ML20314A108, (December 2020).
42. NRC, "Enabling Technologies for Digital Twin Applications for Advanced Reactors and Plant Modernization," Virtual Workshop, September 14–16, 2021, Adams Accession No. ML21228A082 (September 2021).
43. H. Sun, P. Ramuhalli and R. Jacob, "Machine Learning for NDE Literature Review," Submitted to *Ultrasonics*, 2022.
44. ENIQ Recommended Practice 13, *Qualification of Non-destructive Testing Systems that Make Use of Machine Learning*, Issue 1, ENIQ Report 65, European Network for Inspection Qualification (ENIQ), June 2021. Available at https://snetp.eu/wp-content/uploads/2021/06/ENIQ_RP13_Issue1.pdf . [Accessed: Jun. 10, 2020].
45. Executive Office of The President, Office of Management and Budget, OMB M-21-06, "Guidance for Regulation of Artificial Intelligence Applications," Washington, D.C. 20503, November 17, 2020.
46. News European Parliament, "EU AI Act: first regulation on artificial intelligence," 4-06-2023 - 14:06 Created: 08-06-2023 - 11:40. https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence
47. News European Parliament, "Artificial intelligence: threats and opportunities," Updated: 20-06-2023 - 09:08 Created: 23-09-2020 - 14:36. https://www.europarl.europa.eu/news/en/headlines/society/20200918STO87404/artificial-intelligence-threats-and-opportunities
48. EU High Level Expert Group on AI, 'Ethics Guidelines for Trustworthy AI', European Commission, 2019.
49. AIR6987, *Artificial Intelligence in Aeronautical Systems: Taxonomy*, SAE International, 2020-02-05.
50. IEEE Standards Association, 'The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems', [n.d.]; and IEEE Global Initiative, *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, version 2 (Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, [Dec. 2017]).
51. ISO/IEC TR 24030:2021, Information technology — Artificial intelligence (AI) — Use cases, 2021-05.
52. Mike Rowland and Sofia Guerra-Adelard, "Working Group # 3 Report to the Technical Meeting Participants," Technical Meeting on Instrumentation and Control and Computer Security for Small Modular Reactors and Microreactors, IAEA, Vienna, Austria, 21-25 February 2022.