IEC TR 63468

Edition 1.0 2023-05

TECHNICAL REPORT

colour inside

Nuclear facilities – Instrumentation and control, and electrical power systems – Artificial Intelligence applications

IEC TR 63468:2023-05(en)

IEC Secretariat
3, rue de Varembé
CH-1211 Geneva 20
Switzerland

Tel.: +41 22 919 02 11
info@iec.ch
www.iec.ch

About the IEC
The International Electrotechnical Commission (IEC) is the leading global organization that prepares and publishes International Standards for all electrical, electronic and related technologies.

About IEC publications
The technical content of IEC publications is kept under constant review by the IEC. Please make sure that you have the latest edition, a corrigendum or an amendment might have been published.

IEC publications search - webstore.iec.ch/advsearchform
The advanced search enables to find IEC publications by a variety of criteria (reference number, text, technical committee, …). It also gives information on projects, replaced and withdrawn publications.

IEC Just Published - webstore.iec.ch/justpublished
Stay up to date on all new IEC publications. Just Published details all new publications released. Available online and once a month by email.

IEC Customer Service Centre - webstore.iec.ch/csc
If you wish to give us your feedback on this publication or need further assistance, please contact the Customer Service Centre: sales@iec.ch.

IEC Products & Services Portal - products.iec.ch
Discover our powerful search engine and read freely all the publications previews. With a subscription you will always have access to up to date content tailored to your needs.

Electropedia - www.electropedia.org
The world's leading online dictionary on electrotechnology, containing more than 22 300 terminological entries in English and French, with equivalent terms in 19 additional languages. Also known as the International Electrotechnical Vocabulary (IEV) online.

IEC TR 63468

Edition 1.0 2023-05

TECHNICAL REPORT

colour inside

Nuclear facilities – Instrumentation and control, and electrical power systems – Artificial Intelligence applications

INTERNATIONAL ELECTROTECHNICAL COMMISSION

ICS 27.120.20; ICS 35.020

ISBN 978-2-8322-6901-5

Warning! Make sure that you obtained this publication from an authorized distributor.

CONTENTS

INTERNATIONAL ELECTROTECHNICAL COMMISSION

____

NUCLEAR FACILITIES – INSTRUMENTATION AND CONTROL, AND ELECTRICAL POWER SYSTEMS – ARTIFICIAL INTELLIGENCE APPLICATIONS

FOREWORD

IEC TR 63468 has been prepared by subcommittee 45A: Instrumentation, control and electrical power systems of nuclear facilities, of IEC technical committee 45: Nuclear instrumentation. It is a Technical Report.

The text of this Technical Report is based on the following documents:

Draft	Report on voting
45A/1458/DTR	45A/1472/RVDTR

Full information on the voting for its approval can be found in the report on voting indicated in the above table.

The language used for the development of this Technical Report is English.

This document was drafted in accordance with ISO/IEC Directives, Part 2, and developed in accordance with ISO/IEC Directives, Part 1 and ISO/IEC Directives, IEC Supplement, available at www.iec.ch/members_experts/refdocs. The main document types developed by IEC are described in greater detail at www.iec.ch/publications.

The committee has decided that the contents of this document will remain unchanged until the stability date indicated on the IEC website under webstore.iec.ch in the data related to the specific document. At this date, the document will be

- reconfirmed,
- withdrawn,
- replaced by a revised edition, or
- amended.

IMPORTANT – The "colour inside" logo on the cover page of this document indicates that it contains colours which are considered to be useful for the correct understanding of its contents. Users should therefore print this document using a colour printer.

INTRODUCTION

a) Technical background, main issues and organization of the technical report

Artificial intelligence(AI) is transforming many fields including nuclear industry drastically. It has been explored and deployed for many years in the nuclear industry and recent advances in AI have enabled many more potentials. Wide adoption of AI calls for standardization efforts to minimize the risks and optimize the efficiency in developing and deploying AI applications. Due to its nature as an enabling technology, the topic of AI applications will cross-cut with almost all working groups within SC 45A, which entails discussions on the setting up of a new working group to dedicate to this new technical field.

This technical report overviews AI technologies from a nuclear perspective, and summaries potential AI application scenarios in nuclear facilities. Based on these inputs, a three-tiered structure for nuclear AI standards within the framework of SC 45A is proposed, and development priorities are discussed. This document then moves on from technical discussions to the organizational challenges in SC 45A. It analyses the cross-cutting nature of AI applications in nuclear facilities and makes the recommendation the setting-up a new working group, whose title and scope are also proposed. Possibility of SC 45A liaison with other technical subcommittees is explored and recommendation is given accordingly.

b) Situation of the current technical report in the structure of the IEC SC 45A standard series

The technical report IEC TR 63468 is a fourth level IEC SC 45A document. This document overviews the fundamentals of artificial intelligence (AI) and its potential applications in nuclear facilities to foster better understanding and adoption of AI technologies within such facilities. It also proposes a structure for future SC 45A standard series on nuclear AI applications.

For more details on the structure of the SC 45A standard series, see item d) of this introduction.

c) Recommendations and limitations regarding the application of this technical report

This document is the first of its kind within SC 45A, intended to pave the road for extensive and systematic efforts in the standard development activities with regard to AI applications. It helps stakeholders to understand the main benefits and challenges of AI from a nuclear perspective. More documents are expected to follow in this direction in the coming years.

It is important to note that a technical report is entirely informative in nature, and it establishes no requirements.

d) Description of the structure of the IEC SC 45A standard series and relationships with other IEC documents and other bodies documents (IAEA, ISO)

The IEC SC 45A standard series comprises a hierarchy of four levels. The top-level documents of the IEC SC 45A standard series are IEC 61513 and IEC 63046.

IEC 61513 provides general requirements for instrumentation and control (I&C) systems and equipment that are used to perform functions important to safety in nuclear power plants (NPPs). IEC 63046 provides general requirements for electrical power systems of NPPs; it covers power supply systems including the supply systems of the I&C systems.

IEC 61513 and IEC 63046 are to be considered in conjunction and at the same level. IEC 61513 and IEC 63046 structure the IEC SC 45A standard series and shape a complete framework establishing general requirements for instrumentation, control and electrical power systems for nuclear power plants.

IEC 61513 and IEC 63046 refer directly to other IEC SC 45A standards for general requirements for specific topics, such as categorization of functions and classification of systems, qualification, separation, defence against common cause failure, control room design, electromagnetic compatibility, human factors engineering, cybersecurity, software and hardware aspects for programmable digital systems, coordination of safety and security requirements and management of ageing. The standards referenced directly at this second level should be considered together with IEC 61513 and IEC 63046 as a consistent document set.

At a third level, IEC SC 45A standards not directly referenced by IEC 61513 or by IEC 63046 are standards related to specific requirements for specific equipment, technical methods, or activities. Usually these documents, which make reference to second-level documents for general requirements, can be used on their own.

A fourth level extending the IEC SC 45 standard series, corresponds to the Technical Reports which are not normative.

The IEC SC 45A standards series consistently implements and details the safety and security principles and basic aspects provided in the relevant IAEA safety standards and in the relevant documents of the IAEA nuclear security series (NSS). In particular this includes the IAEA requirements SSR-2/1 , establishing safety requirements related to the design of nuclear power plants (NPPs), the IAEA safety guide SSG-30 dealing with the safety classification of structures, systems and components in NPPs, the IAEA safety guide SSG-39 dealing with the design of instrumentation and control systems for NPPs, the IAEA safety guide SSG-34 dealing with the design of electrical power systems for NPPs, the IAEA safety guide SSG-51 dealing with human factors engineering in the design of NPPs and the implementing guide NSS42-G for computer security at nuclear facilities. The safety and security terminology and definitions used by the SC 45A standards are consistent with those used by the IAEA.

IEC 61513 and IEC 63046 have adopted a presentation format similar to the basic safety publication IEC 61508 with an overall life-cycle framework and a system life-cycle framework. Regarding nuclear safety, IEC 61513 and IEC 63046 provide the interpretation of the general requirements of IEC 61508-1, IEC 61508-2 and IEC 61508-4, for the nuclear application sector. In this framework, IEC 60880, IEC 62138 and IEC 62566 correspond to IEC 61508-3 for the nuclear application sector.

IEC 61513 and IEC 63046 refer to ISO 9001 as well as to IAEA GSR part 2 and IAEA GS-G-3.1 and IAEA GS-G-3.5 for topics related to quality assurance (QA).

At level 2, regarding nuclear security, IEC 62645 is the entry document for the IEC/SC 45A security standards. It builds upon the valid high level principles and main concepts of the generic security standards, in particular ISO/IEC 27001 and ISO/IEC 27002; it adapts them and completes them to fit the nuclear context and coordinates with the IEC 62443 series. At level 2, IEC 60964 is the entry document for the IEC/SC 45A control rooms standards, IEC 63351 is the entry document for the human factors engineering standards and IEC 62342 is the entry document for the ageing management standards.

NOTE 1 It is assumed that for the design of I&C systems in NPPs that implement conventional safety functions (e.g. to address worker safety, asset protection, chemical hazards, process energy hazards) international or national standards would be applied.

NOTE 2 IEC TR 64000 provides a more comprehensive description of the overall structure of the IEC SC 45A standards series and of its relationship with other standards bodies and standards.

NUCLEAR FACILITIES – INSTRUMENTATION AND CONTROL, AND ELECTRICAL POWER SYSTEMS – ARTIFICIAL INTELLIGENCE APPLICATIONS

1 Scope

This document overviews the fundamentals of artificial intelligence (AI) as it could potentially be applied within nuclear facilities and identifies proven or potential applications, with the objective to foster better understanding and adoption of AI technologies within such facilities. With the objective of supporting future standard development work of IEC SC 45A in this technical area, this document takes the initiative to propose a structure for SC 45A standard series on nuclear AI applications, and recommends setting up a new dedicated working group to be responsible for and coordinate standard development efforts in this particular area, taking into account its cross-cutting nature.

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. However, this approach does not necessarily exclude the applicability of AI technologies in safety applications in nuclear facilities where the technology itself and the related regulatory framework support such potentials.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

IEC 61513, *Nuclear power plants – Instrumentation and control important to safety - General requirements for systems*

IEC 63046, *Nuclear power plants – Electrical power system – General requirements*

IEC TR 63400, *Nuclear facilities – Instrumentation, control and electrical power systems important to safety – Structure of the IEC SC45A standards series*

ISO/IEC 22989:2022, *Information technology – Artificial intelligence – Artificial intelligence concepts and terminology*

ISO/IEC 23053, *Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)*

ISO/IEC TR 29119-11, *Software and systems engineering – Software testing – Part 11: Guidelines on the testing of AI-based systems*

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

- IEC Electropedia: available at https://www.electropedia.org/
- ISO Online browsing platform: available at https://www.iso.org/obp

3.1
artificial intelligence
AI
research and development of mechanisms and applications of AI systems

Note 1 to entry: This definition is further expanded in Clause 5 of ISO/IEC 22989.

Note 2 to entry: For the purpose of this document, this definition can be supplemented by the definition given in Wikipedia, where Artificial Intelligence is defined as "intelligence—perceiving, synthesizing, and inferring information—demonstrated by machines, as opposed to intelligence displayed by animals and humans".

[SOURCE: ISO/IEC 22989:2022, 3.1.3]

3.2
artificial intelligence system
AI system
engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives

[SOURCE: ISO/IEC 22989:2022, 3.1.4]

3.3
autonomy
autonomous
characteristic of a system that is capable of modifying its operating domain or goal without external intervention, control or oversight

[SOURCE: ISO/IEC 22989:2022, 3.1.5]

3.4
automatic
automation
automated
pertaining to a process or system that, under specified conditions, functions without human intervention

[SOURCE: ISO/IEC 22989:2022, 3.1.7]

3.5
Bayesian network
probabilistic model that uses Bayesian inference for probability computations using a directed acyclic graph

[SOURCE: ISO/IEC 22989:2022, 3.3.1]

3.6
continuous learning
continual learning
lifelong learning
incremental training of an AI system that takes place on an ongoing basis during the operation phase of the AI system life cycle

[SOURCE: ISO/IEC 22989:2022, 3.1.9]

3.7
data mining
computational process that extracts patterns by analyzing quantitative data from different perspectives and dimensions, categorizing them, and summarizing potential relationships and impacts

[SOURCE: ISO 16439:2014, 3.13, modified – “identifies” has been replaced by “extracts”]

3.8
data sampling
process to select a subset of data samples intended to present patterns and trends similar to that of the larger dataset being analyzed

Note 1 to entry: Ideally, the subset of data samples will be representative of the larger dataset.

[SOURCE: ISO/IEC 22989:2022, 3.2.4]

3.9
dataset
collection of data with a shared format

Note 1 to entry: Datasets can be used for validating or testing an AI model. In a machine learning context, datasets can also be used to train a machine learning algorithm.

[SOURCE: ISO/IEC 22989:2022, 3.2.5]

3.10
decision tree
model for which inference is encoded as paths from the root to a leaf node in a tree structure

[SOURCE: ISO/IEC 22989:2022, 3.3.2]

3.11
declarative knowledge
knowledge represented by facts, rules and theorems

Note 1 to entry: Usually, declarative knowledge cannot be processed without first being translated into procedural knowledge.

[SOURCE: ISO/IEC 22989:2022, 3.1.12]

3.12
deep learning
deep neural network learning
<artificial intelligence> approach to creating rich hierarchical representations through the training of neural networks with many hidden layers

Note 1 to entry: Deep learning is a subset of ML.

[SOURCE: ISO/IEC 22989:2022, 3.4.4]

3.13
expert system
AI system that accumulates, combines and encapsulates knowledge provided by a human expert or experts in a specific domain to infer solutions to problems

[SOURCE: ISO/IEC 22989:2022, 3.1.13]

3.14
general AI
artificial general intelligence
AGI
type of AI system that addresses a broad range of tasks with a satisfactory level of performance

Note 1 to entry: Compared to narrow AI.

Note 2 to entry: AGI is often used in a stronger sense, meaning systems that not only can perform a wide variety of tasks, but all tasks that a human can perform.

[SOURCE: ISO/IEC 22989:2022, 3.1.14]

3.15
label
target variable assigned to a sample

[SOURCE: ISO/IEC 22989:2022, 3.2.10]

3.16
life cycle
evolution of a system, product, service, project or other human-made entity, from conception through retirement

[SOURCE: ISO/IEC/IEEE 15288:2015, 4.1.23]

3.17
long-short-term-memory network
LSTM
type of neural network that processes sequential data with a satisfactory performance for both long and short span dependencies

[SOURCE: ISO/IEC 22989:2022, 3.4.7, modified – "long-short-term" has been replaced by "long-short-term-memory"]

3.18
machine learning
ML
process of optimizing model parameters through computational techniques, such that the model's behavior reflects the data or experience

[SOURCE: ISO/IEC 22989:2022, 3.3.5]

3.19
machine learning model
mathematical construct that generates an inference, or prediction, based on input data or information

[SOURCE: ISO/IEC 22989:2022, 3.3.7]

3.20
model
physical, mathematical, or otherwise logical representation of a system, entity, phenomenon, process or data

[SOURCE: ISO/IEC 18023-1:2006, 3.1.11, modified –"or data" added]

3.21
narrow AI
type of AI system that is focused on defined tasks to address a specific problem

Note 1 to entry: Compared to general AI.

[SOURCE: ISO/IEC 22989:2022, 3.1.24]

3.22
neural network
neural net
artificial neural network
NN
network of one or more layers of neurons connected by weighted links with adjustable weights, which takes input data and produces an output

Note 1 to entry: Neural networks are a prominent example of the connectionist approach.

Note 2 to entry: Although the design of neural networks was initially inspired by the functioning of biological neurons, most works on neural networks do not follow that inspiration anymore.

[SOURCE: ISO/IEC 22989:2022, 3.4.8]

3.23
prediction
primary output of an AI system when provided with input data or information

Note 1 to entry: Predictions can be followed by additional outputs, such as recommendations, decisions and actions.

Note 2 to entry: Prediction does not necessarily refer to predicting something in the future.

Note 3 to entry: Predictions can refer to various kinds of data analysis or production applied to new data or historical data (including translating text, creating synthetic images or diagnosing a previous power failure).

3.24
reinforcement learning
RL
earning of an optimal sequence of actions to maximize a reward through interaction with an environment

[SOURCE: ISO/IEC 22989:2022, 3.1.27]

3.25
sample
atomic data element processed in quantities by a machine learning algorithm or an AI system

[SOURCE: ISO/IEC 22989:2022, 3.2.13]

3.26
semi-supervised machine learning
machine learning that makes use of both labelled and unlabelled data during training

[SOURCE: ISO/IEC 22989:2022, 3.3.11]

3.27
sub-symbolic AI
AI based on techniques and models that use an implicit encoding of information, that can be derived from experience or raw data.

Note 1 to entry: Compared to symbolic AI. Whereas symbolic AI produces declarative outputs, sub-symbolic AI is based on statistical approaches and produces outputs with a given probability of error.

[SOURCE: ISO/IEC 22989:2022, 3.1.34]

3.28
supervised machine learning
machine learning that makes use of labelled data during training

[SOURCE: ISO/IEC 22989:2022, 3.3.12]

3.29
symbolic AI
AI based on techniques and models that manipulate symbols and structures according to explicitly defined rules to obtain inferences

Note 1 to entry: Compared to sub-symbolic AI, symbolic AI produces declarative outputs, whereas sub-symbolic AI is based on statistical approaches and produces outputs with a given probability of error.

[SOURCE: ISO/IEC 22989:2022, 3.1.33]

3.30
test data
evaluation data
data used to assess the performance of a final model

Note 1 to entry: Test data is disjoint from training data and validation data.

[SOURCE: ISO/IEC 22989:2022, 3.2.14]

3.31
training
model training
process to establish or to improve the parameters of a machine learning model, based on a machine learning algorithm, by using training data

[SOURCE: ISO/IEC 22989:2022, 3.3.15]

3.32
unsupervised machine learning
machine learning that makes only use of unlabelled data during training

[SOURCE: ISO/IEC 22989:2022, 3.3.17]

3.33
validation data
development data
data used to compare the performance of different candidate models

Note 1 to entry: Validation data is disjoint from test data and generally also from training data. However, in cases where there is insufficient data for a three-way training, validation and test set split, the data is divided into only two sets – a test set and a training or validation set. Cross-validation or bootstrapping are common methods for then generating separate training and validation sets from the training or validation set.

Note 2 to entry: Validation data can be used to tune hyper-parameters or to validate some algorithmic choices, up to the effect of including a given rule in an expert system.

[SOURCE: ISO/IEC 22989:2022, 3.2.15]

4 Abbreviated terms

AI	Artificial Intelligence
BWR	Boiling Water Reactor
CAP	Corrective Action Program
CFD	Computational Fluid Dynamics
CNN	Convolutional Neural Network
CRDM	Control Rod Drive Mechanism
CV	Computer Vision

DL	Deep Learning
ENIQ	European Network for Inspection and Qualification
FFD	Fitness For Duty
GOFAI	Good Old-Fashioned Artificial Intelligence
HST	Hot Spot Temperature
I&C	Instrumentation and Control
KG	Knowledge Graph
ML	Machine Learning
NDT	Non-Destructive Testing
NLP	Natural Language Processing
NN	Neural Network
NPP	Nuclear Power Plant
RUL	Remaining Useful Life
SAE	Society of Automotive Engineers
SSC	Systems, Structures, Components
V&V	Verification and Validation

5 AI overview from a nuclear perspective

5.1 Brief history of AI

The history of AI dates back to the 1940s when Warren McCulloch and Walter Pitts suggested connected neuron networks could learn. Alan Turing proposed the Turing test, machine learning, and reinforcement learning in his 1950's article "Computing Machinery and Intelligence."

The term "artificial intelligence" was coined at a Dartmouth workshop in 1956. The next two decades were golden years for AI and the field received extensive government funds for its promising potential for logic-based problem solving. However, by 1974, overly high expectations and limited capabilities led to the first "AI winter". The rise of knowledge-based expert systems brought new successes in the 1980s and following years. But the second "AI winter" started with the identification of expert systems limitations in 1987. AI returned to favor in 1993 with the help of increased computational power. From 2012, unprecedented availability of data and computational power enabled breakthroughs in machine learning, in particular deep machine learning, and ushered in greater success for AI. In the last decade, AI applications in such fields as image analysis, speech recognition and autonomous driving have greatly changed people's daily lives. With deep learning and other advanced techniques, AI now can outperform humans over a range of specific tasks such as image recognition and also beat human champions at games such as go. An intuitive development of different periods of AI is shown in Figure 1.

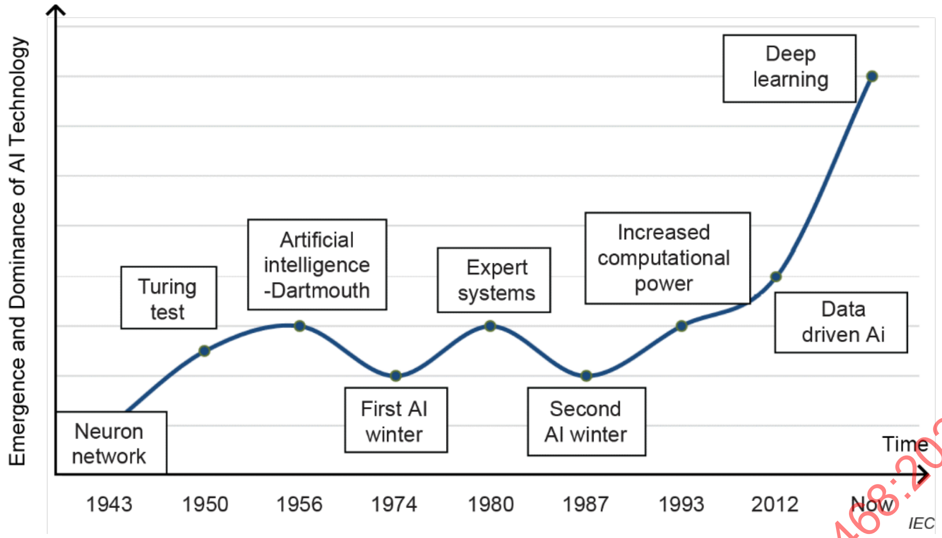

Figure 1 – Brief history of AI

The nuclear industry's effort in pursuing the application of AI techniques dates back to, at least, the early 1980s, when the AI itself was experiencing the second boom with the success of expert systems. As the industry shifted into improving the safety aspects of operating NPPs, following the Chernobyl accident, the interest in automation was overshadowed by the need to improve safety. Over decades of operations, the industry resorted to increasing staffing levels to meet increasing and emerging safety requirements, accompanied by increasing use of systematic processes and procedures in all aspects of industry operations. This increase of staffing requirements offset the economic advantage of the relatively low energy cost of nuclear fuel in comparison to other forms of baseload energy sources, i.e., mainly fossil energy.

However, the drop in oil and gas prices following the 2008 global financial crisis and the associated crash of fossil energy prices resulted in a substantial economical challenge to nuclear power, especially in non-subsidized nuclear energy markets. An NPP operating at 1 000 MWe could have more than 10 folds the number of a staff as a fossil plant operating at the same power level. This disadvantage of nuclear energy, coupled with the urge to fight climate change using nuclear energy as the main baseload clean source of energy, rejuvenated the interest in automation to offset manual activities performed by an NPP staff and enable the industry to become more economically sustainable.

In parallel to this realization by the industry, AI as a science have benefited from advancements in computational power and presented unique capabilities to enable the needed automation for the nuclear power industry. Given, all those factors, AI has found an exponentially growing number of applications in the nuclear industry. While NPPs 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 reactors design, fuel optimization, intelligent control, preventive maintenance, ageing management, non-destructive testing, physical protection, cybersecurity, and many other related fields.

5.2 Major concepts of AI

5.2.1 AI definition

Historically, AI has been approached from a number of 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.

This document will leverage the AI definition from ISO/IEC 22989, which views AI systems as engineered systems that generate outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives. It is worthwhile to note that this definition has been further expanded in Clause 5 of ISO/IEC 22989:2022 to make it more concrete. For the purpose of this document, the definition can be supplemented by the one given in Wikipedia, where AI is viewed as "intelligence—perceiving, synthesizing, and inferring information—demonstrated by machines, as opposed to intelligence displayed by animals and humans". This supplemental definition emphasized the machine intelligence aspects of ISO/IEC 22989 definition.

To put this definition into context, a simplified functional view of an AI system is presented in Figure 2, which is illustrated from a nuclear facility application perspective. The workflow for a deployed nuclear AI system is very straightforward, which is comprised of three major components, namely inputs, AI processing and outputs. Taking the AI system for identifying plant transients as an example, it takes inputs from plant instrumentation and control systems, performs intelligent analysis and inference, and then generates some outputs that predict the type of transients the NPP is experiencing.

Behind the abovementioned workflow, is an AI model that has been designed and developed by human developers. Essentially there are two major approaches to build the model, either encoding human knowledge (e.g., expert experience) into the AI model, or "educate" a model by enabling it to "learn" from data gathered for this specific purpose. Moreover, the model itself can autonomously evolve, in a process known as continuous learning. Technically, these two different ways for achieving intelligence are known as symbolic and sub-symbolic AI, and these terms will be explained in the coming sections. A combination of the two approaches is also possible by leveraging to various extent the data and human knowledge.

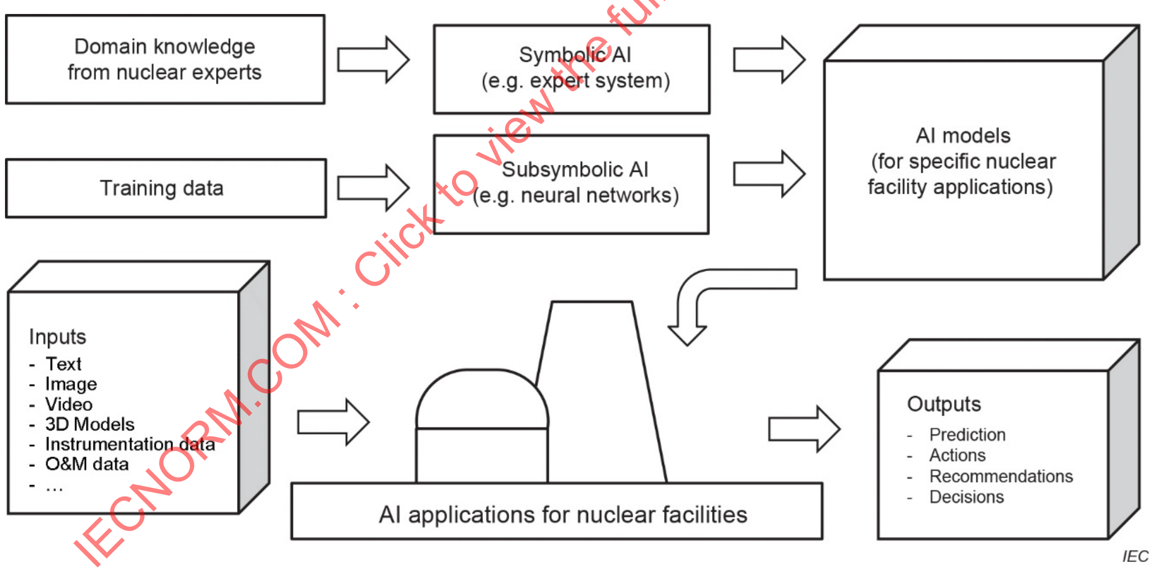

Figure 2 – Functional view of a nuclear AI system

5.2.2 Levels of intelligence and autonomy

It is worthwhile to note the difference between "strong AI" and "weak AI" which describes the levels of intelligence an AI system has. Simply put, 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 will possess an intelligence equal to humans, and it would have a self-aware consciousness that has 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.

Practically, this document will focus on weak AI, and there are different classifications of autonomy. In some scheme, autonomy can be described in three classes: algorithmically based, algorithmically driven and algorithmically-determined systems.

Considering the current level of AI development, the interpretability of AI prediction results is still an important content that scholars are studying, so it is suggested that the AI in this document will mainly play a recommendatory role in nuclear facilities, and currently it is not recommended to be directly used in important and safety-related scenarios.

Algorithmically based AI applications work as pure assistance systems without autonomous decision-making authority. An example in the nuclear industry is the use of AI to flag anomalies for a human to analyse and decide what needs to be done. Algorithm-driven AI applications take partial decisions from humans or shape human decisions through the results they calculate. As a result, the actual decision-making scope of humans and consequently their possibilities for self-determination shrink. AI-based predictive maintenance is an example of such approach, where AI estimates how long an equipment has before it needs to be maintained and the human makes the final decision on that prediction. Finally, algorithmically-determined AI applications make decisions independently and thus exhibit a high degree of autonomy. Due to the high degree of automation, there is no longer a human decision in individual cases, especially no human review of automated decisions. Autonomously operating micro reactors is a common futuristic use case of this type of AI.

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

- 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, which is called human-machine teaming. Overall, the presence of accountable supervision during operation can assist in ensuring that the AI system works as intended and avoids unwanted impacts on stakeholders;
- the system's degree of situated understanding, including the completeness and operationalizability 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;
- the ability to evaluate its own performance or fitness, including assessments against pre-set 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 Society of Automotive Engineers (SAE) which defines 6 levels of driving automation ranging from 0 (fully manual) to 5 (fully autonomous). It has been widely adopted by autonomous driving system companies worldwide, which is illustrated in Table 1 in the context of operating a valve for load following in an NPP.

For every application in nuclear facilities, AI systems deployed can have different degrees of decision autonomy. The deployment of various levels of autonomy has been gradually advancing from the low levels of autonomy in Table 1 to recently L4 type of applications, especially for futuristic uses of micro reactors that adopt a decentralized and autonomous modes of monitoring and control. However, enabling highly or fully autonomous AI use in nuclear facilities is still not practical presently, due to both technical and regulatory concerns.

Table 1 – Example of autonomy levels for nuclear facilities (Referring to the categorization scheme of autonomy levels from SAE)

Autonomy levels	Operating a valve in a nuclear load following example
L0: No automation	Manually operating a process valve based on operation procedures and grid information
L1: Assistance	AI recommends operating a valve based on grid condition but a human needs to act on it
L2: Partial automation	AI recommends operating a valve based on grid condition. A human approves grid demand and removes the valve manual operation override to allow AI to operate the valve
L3: Conditional automation	AI operates a valve based on grid condition under normal operating conditions. The human monitors its performance and can take control at any point of time
L4: High automation	AI operates a valve based on grid condition under normal operating conditions. The human does not exclusively monitor the valve performance but can take control at any point of time
L5: Full automation	AI operates a valve based on grid condition in all conditions. The human does not exclusively monitor the valve performance but can take control at any point of time

5.2.3 Data processing and specifications

Data are central to any data-driven AI system to learn the underlying models that represent input variables relationships to the desired output. Data can come in structured form (e.g., relational databases in NPP data historians) or unstructured form (e.g. event reports, specification documents, non-destructive testing images and files). Data are a key aspect of AI systems and they go through processes (as defined in ISO/IEC 22989) including:

– data acquisition, in which the data are obtained from one or more sources. The suitability of the data needs to be assessed, for example, whether it is biased in some ways or whether it is broad enough to be representative of the expected operational data input. The nuclear industry has been accumulating data over decades of operation, most of which is available in separate tools and databases, representing islands of data even in a single NPP. Despite the variation in data across the industry, the basic systems and tools used are consistent. Therefore, the data is considered broad enough to apply AI methods, provided it can be mapped to a standard set of data model across the industry;

– normalization, which is the adjustment of data values 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 datasets from the different plants. For example, one nuclear facility could prioritize events alphabetically, others might use numeric. Within a process, the process data units could also differ (e.g. gallons per minute vs m^3/s);

– data quality checking, in which the contents of the data are examined for completeness, for bias and other factors that affect its usefulness for the AI system. For example, sensors partial failure (e.g. sensors drifting) or complete loss could need to be rectified before the data is used;

– data labelling, in which datasets are labelled. To use supervised learning techniques, samples should be labelled manually or semi-manually. The nuclear industry data ranges from fully unlabelled data (e.g. process sensors missing labels on equipment condition) to full labelled data (e.g. an event assigned a priority by a human and logged into a database);

Depending on the use case and on the approach used, data in an AI system can be involved in several ways:

– training data is used specifically in the context of AI: it serves as the raw material from which the AI learning algorithm modifies its model to address the given task. This training data can be of any form such as time series, images, video, text, or acoustic. Because the current fleet of nuclear facilities are data abundant, this data usually exists. However, for newly developed nuclear reactors, this data is usually not available and a simulation of a model that represent the system is performed to generate synthetic data to use;
– validation data corresponds to data used by the developer to make or validate some algorithmic choices (hyper-parameter search, rule design, etc.). This set of data is usually extracted from the same data set from which the training data is extracted;
– test data is the data used to evaluate the performance of the AI system, before its deployment. It is expected to be similar to production data, and proper evaluation needs test data to be disjoint from any data used during development. The test data is usually used for supervised machine learning methods. Given that the unsupervised machine learning methods lack labelled data, dedicated methods exist to evaluate the performance of the unsupervised machine learning algorithm, which often involve human evaluations of the results. One example of the methods used to test unsupervised machine learning methods involve benchmarking the findings of one model or method against others;
– production data is the data processed by the AI system in the operation phase for nuclear facility applications.

5.2.4 Symbolic and sub-symbolic approaches

5.2.4.1 General

Since the foundation of AI as a discipline, two paradigms have developed: symbolic AI and sub-symbolic AI. Symbolic AI, also known as Good Old-Fashioned AI (GOFAI), focuses on attempting to express human knowledge clearly in symbols, that is, facts and rules. Sub-symbolic AI is characterized in particular by an inductive procedure, i.e. by the (algorithmic) derivation of general rules or relationships from individual cases, and the major technique is machine learning.

Modern AI systems typically contain elements of both symbolic AI and sub-symbolic AI. Such systems are called hybrid AI. Figure 3 shows the three approaches to AI and includes some typical examples algorithms/techniques under each category. These examples are not intended to be exhaustive.

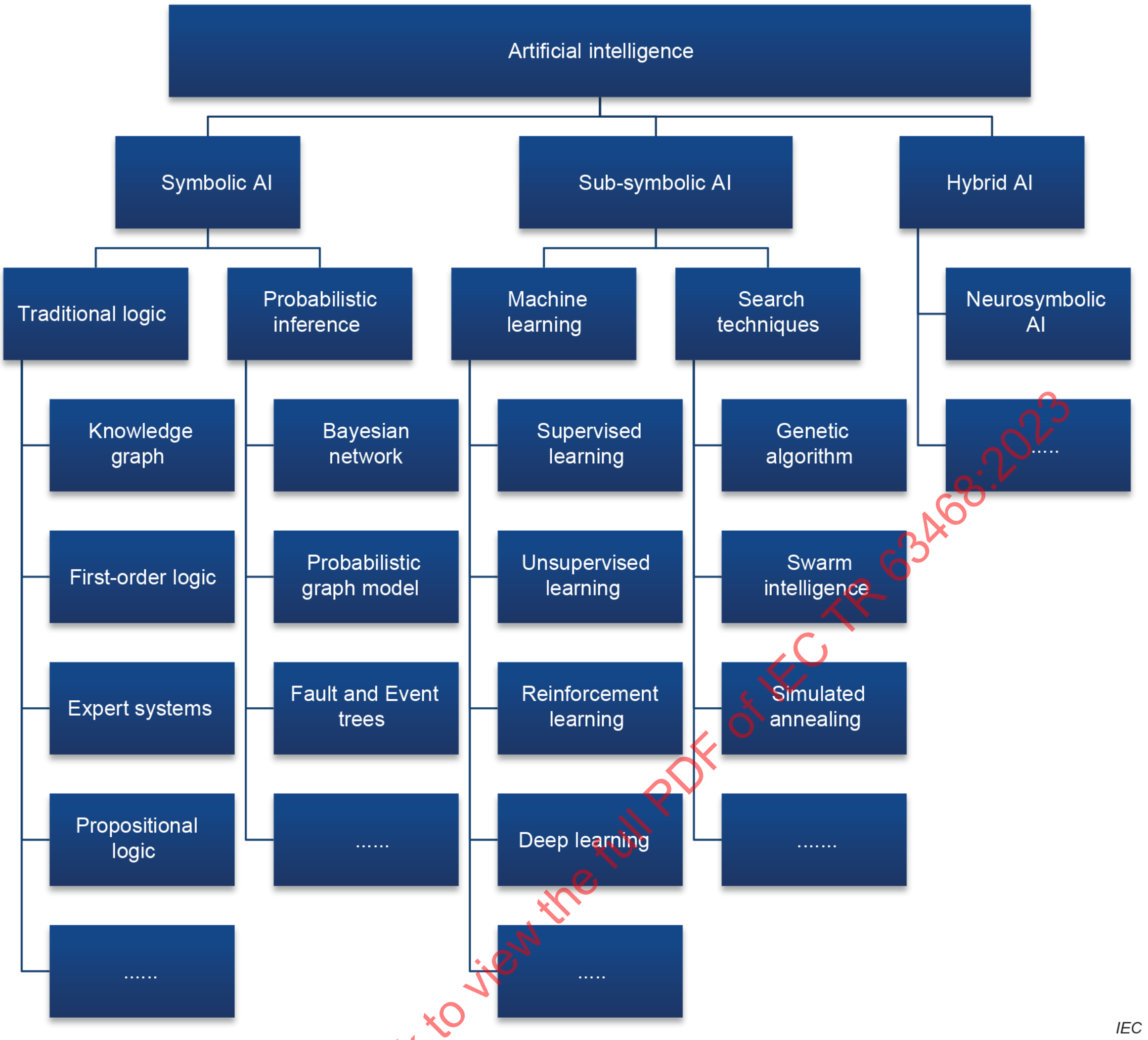

Figure 3 – Major approaches to AI

5.2.4.2 Symbolic AI

5.2.4.2.1 General

Symbolic AI involves encoding knowledge with symbols and structures and it mostly uses logics or mathematical formulation to model reasoning processes. Information is represented using a formal syntax that can be processed by a machine and is explainable to a human. In the nuclear industry, explainability of any automated decision is a key requirement to enable independent evaluation of the underlying logic and models. Therefore, symbolic AI is important for safety-critical applications, as its reasoning process can be tracked, verified and explained.

5.2.4.2.2 Traditional logic

5.2.4.2.2.1 General

Traditional logic is a category of AI that converts reasoning statements of associations between events into forms of logical relationship between the input variables and the output.

5.2.4.2.2.2 Propositional logic

Propositional logic, also known as sentential logic, statement logic and Boolean logic, provides methods to create an incremental set of structures to design a complicated structure. This structure builds a logical relationship between a predefined set of inputs and outputs.

Propositional logic is represented by a language structure that is based on simple undividable states joined together with logical connectors. For example,

P: The NPP with an seismic trip system experienced a beyond design basis earthquake.

Q: The control rods were automatically released.

P $\rightarrow$ Q.

Due to this unique simplicity and powerful outcome for logical reasoning, propositional logic is widely adopted in AI in solving complex problems that can be disassembled into simple logic relationships, such as sequential control routines in NPPs. While propositional logic is a useful tool for reasoning, it has some limitations such as:

– It cannot address relations like ‘some’.
– It has limited expressive ability.
– It is prone to human errors.

5.2.4.2.2.3 First-order logic

First-order logic is another way of knowledge representation in AI, and is sufficiently expressive to represent the natural language statements in a concise way. First-order logic does assume that the world contains facts like propositional logic but also contains the following contents with syntax and semantics:

– Objects: valves, pumps, nuclear fuel, work orders.
– Relations: unary relation (approved version, within calibration, verified) or n-any relation (a component of, supervisor of, in between).
– Functions: accident precursor, power module of, end of.

Because first-order logic is close to the semantics of natural language, it can be used for knowledge engineering (the process of constructing a knowledge-base) in a special-purpose knowledge base (in contrast to general purpose knowledge base). However, first-order logic still has its limitations on its expressiveness and expressive logics like second-order/higher-order were developed to overcome those limitations.

5.2.4.2.2.4 Expert systems

Expert systems, also called “knowledge-based systems”, could match or exceed the performance of human experts on narrowly defined tasks, if given the appropriate domain knowledge.

An expert system consists of an inference engine and the knowledge base. The knowledge base is materialized using rules, which are then applied by the inference engine on known facts to deduce new facts.

The expert systems played an important role in the nuclear industry. According to IAEA-TECDOC-542, expert systems for plant data management, training, condition/safety status monitoring, alarm analysis and diagnosis, accident management, emergency planning have been applied to NPPs worldwide since 1980.

5.2.4.2.2.5 Knowledge graph

The term “knowledge graph (KG)” was introduced by Google in 2012 to refer to its general-purpose knowledge base. Today, KGs are used extensively in anything from search engines and product recommenders to autonomous systems. Technically, a KG is a directed labelled graph in which the labels have well-defined meanings. A directed labelled graph consists of nodes, edges, and labels. Anything can act as a node, for example, NPP process systems, equipment such as valves or pumps, etc. An edge connects a pair of nodes and captures the relationship of interest between them, for example, is_a_component_of relationship between the reactor coolant system and the reactor coolant pump. The labels capture the meaning of the relationship, for example, is_a_component_of in the previous example can be the label.

5.2.4.2.3 Probabilistic inference

5.2.4.2.3.1 General

Probabilistic inference converts reasoning statements of associations between events into forms of probabilistic cause and effect relationship between the input variables and the output.

5.2.4.2.3.2 Bayesian networks

A Bayesian network is a formulation of a belief that one event is likely caused by, or associated with another. This belief is quantified through a probabilistic and statistical relationship that is constructed using historical occurrence of the correlated events.

In nuclear facilities, Bayesian networks are usually constrained by a set of symbolic and predefined structures that are used as the basis to develop the probabilistic relationship. They can be used to predict the likelihood of an event occurring, such as an equipment failure. They are also used to aid in the diagnosis of the reactor's state, as well as plant operation risks modelling.

5.2.4.2.3.3 Decision trees

Decision trees are used to cascade a set of events into a holistic event of interest. Fault or event trees are a common example in nuclear. A fault or event tree is a logical decision tree consisting of structure nodes that connect two or more leaf nodes and/or other structure nodes to provide an estimate of the likelihood of a fault or event occurrence and the consequence of a series of occurrences. The structure of the tree consists of a set of logical rules that are represented by the nodes with each branch of the node.

Fault trees and event trees have been historically used in safety evaluation and licensing of NPPs as they are capable of aggregating a series of events likelihood of a severe accident occurrence. Dynamic forms of the trees have recently been explored to enable the trees to adapt to plant conditions and adaptively risk-inform its operations.

5.2.4.2.3.4 Probabilistic graph model

Probabilistic graph models are a branch of machine learning whose purpose is to describe and reproduce things in the world using the overall probability distribution. The main goal of the probability graph algorithm is to achieve more accurate probability inference under limited computing power. Given some observed variables, based on the relationship between the observed variable and the inferred variable, the probability of the inferred variable we want is deduced.

The concept of probability graph model is mainly divided into three parts:

a) Representation: how to express it with a model from a real-world problem.
b) Inference: How to infer the probability of the variable you want to ask from the model generated above.
c) Learning: How to further fit our model with real-world data.

5.2.4.3 Sub-symbolic AI

5.2.4.3.1 General

The other AI approach is sub-symbolic AI. This approach is not based on symbolic reasoning; rather, it relies on the implicit and data-based encoding of knowledge. This implicit knowledge representation is fundamentally based on statistical modeling of experience or raw data. Examples of this type of AI system are various machine learning systems, including the different forms of deep neural networks.

In nuclear facilities, this AI approach has been rapidly growing from an application perspective, mainly because of the complexity of the targeted application and the difficulty to develop representative and explainable relationships and structures for symbolic AI.

5.2.4.3.2 Machine learning

5.2.4.3.2.1 General

Machine learning is a branch of AI technologies that enables a machine to learn data patterns and use them to make predictions. Machine learning continues to learn and improve its performance during and after making observations. ISO/IEC 22989 defines machine learning as the process of optimizing model parameters through computational techniques, such that the model's behavior reflects the data or experience.

Figure 4 summarizes the major process for developing and deploying a machine learning application, as described in ISO/IEC 23053. It includes the phases of data acquisition, data pre-processing, modelling, model development, and operation. NPPs accumulate huge volumes of data during operations, which offer great opportunities to explore AI for many applications.

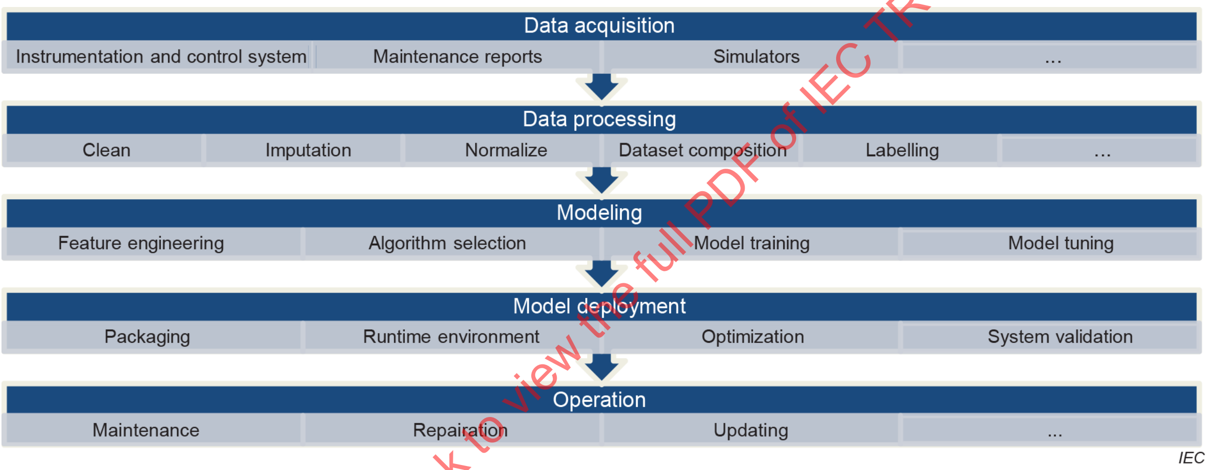

Figure 4 – Machine learning pipeline (workflow)

Machine learning methods can be classified into three approaches: supervised machine learning, unsupervised machine learning, and reinforcement machine learning. The major difference between supervised machine learning and unsupervised machine learning lies in that supervised machine learning expects the training data to be labelled and the labels are used to train the model to predict similar labels if similar data is provided.

In unsupervised machine learning, data is not labelled, and the machine attempts to establish its own set of labels without any form of assurance that the labels are correctly assigned.

Semi-supervised machine learning combines the two approaches and is suitable when the data is partially labelled.

In nuclear facilities, anomaly detection algorithms applied to process sensors data usually rely on unsupervised machine learning, because of decades of operations process data that are not labelled as failures occurred. Recently, the industry starts to use some of the condition reports that log any issues in the plant to label parts of the process data, which is why semi-supervised machine learning became of interest. For specific equipment that has a full log of events, it is possible to fully label the historical data, enabling the use of supervised machine learning methods.

Given that the labels represent a form of knowledge that assist the machine in its learning, supervised machine learning methods are generally more capable than unsupervised or semi-supervised machine learning methods.

In reinforcement learning, the machine learns by means of rewards and punishment functions. In order to learn, the machine is allowed to iterate, fail, and learn from its failure by means of a reward or punishment function. As the system learns, the number of failures drops, and successes increase to the point where it achieves the desired success performance.

There are many techniques that have been developed under this umbrella, and Figure 5 provides an overview of these techniques.

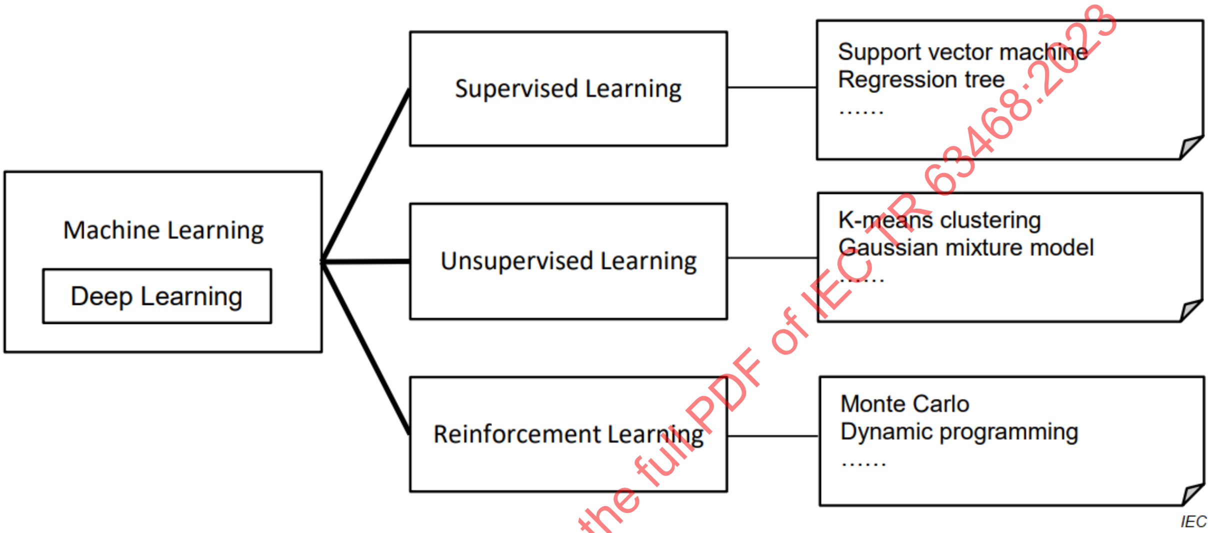

Figure 5 – Major approaches to machine learning

5.2.4.3.2.2 Supervised learning

Supervised machine learning uses labelled data during the training process. Labelled data consists of any form of data samples (e.g., sensors, text, images) with inputs mapped to outputs of known states. Thus, the training data are organized as pairs of input variables and “known” outputs. Known outputs are also referred to as labels, target variables, and ground truth, depending on the context. The labels can be of any types including categorical, binary, or numeric values. Supervised learning can be used for classification and regression, as well as more complex tasks pertaining to prediction.

Supervised learning, as an example, can be used to train a machine to detect a fire in a video stream to automate a nuclear facility fire watch. The dataset consisting of images containing fire and images with no fire are labelled by a human and used by a machine to determine the common features of fire. The resulting machine learning model would be able to determine if a new video contains fire.

Another example is the use of supervised machine learning to automatically review and classify condition reports in an NPP. Those reports have been classified by plant staff for decades. This human-based classification represents a label that a machine can consider as the ground truth to learn patterns in text that contributed to this classification decision and predict future classifications of condition reports.

5.2.4.3.2.3 Unsupervised learning

In contrast to supervised machine learning, unsupervised machine learning uses unlabelled data. It is used to identify similarities, patterns, or clusters of data and can be used to reduce dimensionality irrespective of any label. Unsupervised machine learning can be used in the discovery of dominant factors (i.e., capture the “essence” of the data).

NPPs are complex systems with different combinations of process alignments and equipment working modes. The data collected during plant operation contain mostly data representing normal condition of the plant equipment but also faulty conditions that are unlabelled. If the equipment failure results in a shift in the operational data and the unsupervised machine learning is used to cluster the equipment into two clusters, it is possible that the resulting clusters would represent the two states, normal and abnormal. Therefore, despite not having clear labels of failures, unsupervised machine learning can separate the two states by the data only.

5.2.4.3.2.4 Reinforcement learning

Reinforcement machine learning is the process in which a machine learning agent learns through an iterative process by trial and error. The agent’s goal is to find a model that results in the best rewards from its action’s impact on the environment. For each trial (successful or failure), feedback is provided by the environment to the model. The mode is then adjusted based on this feedback.

Reinforcement learning has been used to optimize nuclear fuel assembly designs. In a typical reactor, fuel rods are lined up on a grid, or assembly, by their levels of burnup, like chess pieces on a board. In an ideal layout, these competing impulses balance out to drive efficient reactions. As the fuel is shuffled, the reward function represented by metrics such as a uniform or flat power profile, is evaluated and the model is tuned until the desired power profile is achieved. Reinforcement learning can find optimal solutions faster than a human and quickly modify designs in a safe, simulated environment.

5.2.4.3.2.5 Deep learning

Deep learning is a machine learning technique that constructs artificial neural networks to mimic the structure and function of the human brain. In practice, deep learning uses a large number of hidden layers - typically more than 6 but often much higher – of nonlinear processing to extract features from data and transform the data into different levels of abstraction (representations).

Deep learning has been the main innovation that has renewed interest in AI in the past years, helping solve many critical problems in computer vision, natural language processing, and speech recognition. Models are trained by using a large set of data and neural network architectures that contain many layers. It can extract features that a human used to develop and therefore achieve results that were not possible before, exceeding human-level performance. For example, instead of developing features for a neural network to evaluate fire video in a model to detect fires in a video stream, such as color or shape of fire, a deep neural network has several layers that automatically extract those features and associates those features with the output.

5.2.4.3.3 Search techniques

5.2.4.3.3.1 General

There are many complex problems that make it impossible or extremely challenging to find the exact solution in a strictly analytical manner. To address these issues, many novel algorithms have been developed to search for a feasible solution, rather than the absolute optimal solution, and many of them try to emulate biological or natural phenomena. This subfield of AI has been extensively explored in engineering systems, and nuclear industry also benefits from its progress.

5.2.4.3.3.2 Genetic algorithms

A genetic algorithm represents the variables in the problem domain in the form of chromosomes of fixed length. The parameters like the initial size of the chromosome population need to be defined, probability of crossover, and mutation, also a fitness function is decided to select the chromosomes that will get the opportunity to mate during the phase of reproduction. The individuals which result from the mating of the parent chromosomes form the next generation of solutions and the process is repeated. Genetic algorithms are often used in optimization problems in nuclear fields, such as the optimization of fuel assembly shuffling during refueling to maximize a certain fitness function, representing a flat power profile.

5.2.4.3.3.3 Swarm intelligence

Swarm intelligence algorithms are a form of nature-based optimization algorithms. Their main inspiration is the cooperative behavior of animals within specific communities. This can be described as simple behaviors of individuals along with the mechanisms for sharing knowledge between them, resulting in the complex behavior of the entire community. Examples of such behavior can be found in ant colonies, bee swarms, schools of fish or bird flocks.

Swarm intelligence algorithms are used to solve difficult optimization problems similar to the fuel management and shuffling example discussed earlier.

5.2.4.3.3.4 Simulated annealing

Simulated annealing is a technique used to find solutions to optimization problems. It is based on the idea of annealing in metallurgy, where a metal is heated and then cooled slowly in order to reduce its brittleness. In the same way, simulated annealing can be used to find solutions to optimization problems by slowly changing the values of the variables in the problem until a solution is found. The advantage of simulated annealing over other optimization methods is that it is less likely to get stuck in a local minimum, where the solution is not the best possible but is good enough.

5.2.4.4 Hybrid AI

5.2.4.4.1 General

Hybrid AI is a special type of AI that uses symbolic and sub-symbolic AI to various extent as each of the two is not capable of satisfactorily making a decision on its own. Using hybrid AI can improve the accuracy of the prediction, reduce the data requirement, and require less time to make the decisions than sub-symbolic AI.

5.2.4.4.2 Neurosymbolic AI

Neurosymbolic AI is a hybrid AI approach that couples symbolic AI to neural networks. One use case is when AI is used to count a certain number of objects in a picture that has specific attributes. For example, if AI is used to count large red balls in a picture, then a neural network (i.e. sub-symbolic AI) is used to extract the red balls and a set of rules (i.e. symbolic AI) is used to qualify the object as large based on its size in the image and then count all the large red balls.

Though neurosymbolic AI is relatively a new field of research, it provides a key advantage to the nuclear industry, which is the added level of explainability. The symbolic parts of the AI could be used to make a critical part of the decision-making process, because of its explainability, while the less critical part leverages sub-symbolic AI decision making. This could be especially beneficial in designing AI-based safety systems for autonomous control of nuclear reactors.

5.3 Specific fields of AI applications

5.3.1 Data mining

Data mining refers to the application of algorithms for the discovery of valid, novel and useful information from data. Data Mining is a subset of Machine Learning that centres around exploratory data analysis which came into prominence in the late 1990s.

The end goal is to extract relevant information (and not the “extraction” of raw data itself) from datasets and transform the same into operational insights for further use. The important aspect of data mining is that it helps to answer questions we did not know to ask by proactively identifying non-intuitive data patterns through algorithms.

Data mining methods can be directly applied to raw data to perform prediction, usually through regression. Regression analysis is about understanding which factors within a data set are most important to a certain output, which can be ignored, and how these factors interact to predict the output. A typical example in nuclear facilities is the use of prediction to perform anomaly detection in process sensors data. Anomaly detection is the process of finding data that does not conform to the pattern. In anomaly detection patterns extracted from the data are used to uncover unusual data by predicting what the process value should be then comparing it to what is observed. This process can help to alert plant operator to situations deviating from normal operations before it reaches alarming or control system responses.

Data mining can also be used in combination with other fields of AI, such as natural language processing or computer vision to perform classification. Data points are assigned to groups, or classes, based on a specific question or decision to make. Classification can be performed using labelled data (i.e., data is assigned to previously defined groups) in supervised machine learning methods. It can also be performed by unsupervised machine learning methods through clustering. Clustering looks for similarities within a data set, separating data points that share common traits into subsets. Given that process anomalies detection mostly leverages unsupervised machine learning, clustering separate normal equipment from ones that are demonstrating some level of anomaly.

Data mining can be also used to perform sensitivity analysis through input variables correlation development. This function seeks to uncover the relationships between data points; it is used to determine whether a specific action or variable has any traits that can be linked to other actions. For example, it can be used to identify if there are notable correlations between the occurrence of a certain event in a nuclear facility condition report and equipment failure.

On a broader sense, the data correlation can be used to perform sequential correlation of events (referred to as planning in ISO/IEC 22989). This entails using a machine to automatically find a procedural sequence of actions, for reaching certain goals while optimizing certain performance measures. From the perspective of planning, a system occupies a certain state. The execution of an action can change the system state and the sequence of actions proposed by planning can move the system from the initial state closer to the goal state. This is beneficial to in nuclear projects throughout its life cycle, from planning the schedule of construction, and planning the outage activities, throughout to planning the decommissioning work. For example, during an NPP outage, AI can help operators optimize their work schedule to optimize the resources and time schedule.

5.3.2 Natural language processing

Natural language processing is information processing based upon natural language context understanding and natural language generation. This encompasses natural language analysis and generation, with text or speech. By using NLP capabilities, computers can analyse text that is written in human language and identify concepts, entities, keywords, relations, emotions, sentiments and other characteristics, allowing users to draw insights from content. With those capabilities, computers can also generate text or speech to communicate with users. Any system that takes natural language as input or output, in text or speech form, and is capable of processing it is using natural language processing components. An example of such a system is an automated review of condition reports or work order in an NPP. Natural language processing can be used to understand free text entries and classify the condition or work order text and suggest decisions and actions, thereby improving the efficiency of the process.

5.3.3 Computer vision

Computer vision uses AI to acquire, process, and interpret data representing images or video. Computer vision encompasses image recognition, e.g., the processing of digital images. Digital images exist as a matrix of numbers that represent the grey scales or colors in the captured image or in other cases a collection of vectors. Digital images can include metadata that describes characteristics and attributes associated with the image.

Visual data ordinarily originates from a digital image sensor, a digitally scanned analogue image or some other graphical input device. For the purposes of this document, digital images include both fixed and moving variants (i.e., video). Fundamental tasks for computer vision include image acquisition, re-sampling, scaling, noise reduction, contrast enhancement, feature extraction, segmentation, object detection and classification.

In the nuclear industry, AI applications based on computer vision could identify specific images out of a set of images (e.g., detect fire in a video stream monitoring an area when the fire protection system is not functional). It can be used in quality control (e.g., spotting defective or corroding parts on a steam generator or non-destructive test images to detect cracks or containment condition), or security applications such as facial recognition for NPP security check.

5.4 Challenges of AI applications in nuclear facilities

5.4.1 General

As shown in Clause 6, AI has been successfully applied to, or has the potential to be applied to many aspects of nuclear facilities. However, because AI systems are typically complex (e.g. deep neural networks), can be poorly specified and can be non-deterministic, it creates some new challenges.

Take NPP regulations as an example. With AI being used to replace or augment a safety-related or risk-significant system, AI needs to adapt to current standards that are adopted by the regulatory framework. 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. Due to 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.

5.4.2 Trustworthiness

AI applications have raised some challenges due to its high level of intelligence and low level of transparency. This concern is addressed by demonstrating its trustworthiness which refers to characteristics as follows which help relevant stakeholders understand whether the AI system meets their expectations. This includes demonstrating:

– AI explainability: the important factors influencing a decision can be expressed in a way that humans can understand.
– AI transparency: human centred objectives for the system.
– AI robustness: describe the ability to maintain their level of performance.
– AI resilience: describe the ability of a system or an entity within that system to perform its required functions under unexpected conditions for a specific period of time.
– AI controllability: an external agent can intervene in its functioning.
– AI predictability: reliable assumptions by stakeholders about the output.

– AI bias and fairness: different cases call for different treatment.
– AI reproducibility: obtain the same results by using the same methodology.
– AI security: protected from cybersecurity risks that might lead to physical and/or digital harm.
– AI safety: freedom from unacceptable risk.

5.4.3 AI verification and validation

As in the application of digital I&C systems for the nuclear industry, verification and validation play a central role in enabling the use of AI systems for nuclear sectors. However, due to the black-box nature of some AI systems, there are special considerations in their verification and validation processes, including:

– Some systems are partially verifiable and partially validatable (e.g. at least one system component can individually be verified, and the remaining system components can be validated, as can the complete system).
– Some systems are unverifiable but validatable (e.g. no system component can be verified, but all system components can be validated, as can the complete system).
– Some systems are unverifiable and partially validatable (e.g. no system component can be verified, but at least one system component can individually be validated).
– Some systems are unverifiable and unvalidatable (e.g. no system component can either be verified or validated).

This challenge has been addressed in ISO/IEC TR 29119-11, which offers guidelines on the testing of AI-based systems. It covers testing of AI systems across the life cycle and gives guidelines on how AI-based systems in general can be tested using black-box approaches and introduces white-box testing specifically for neural networks.

6 Some potential nuclear AI applications

6.1 General

AI techniques have been developed and used in many fields of the nuclear industry, including, but not limited to, reactor system research and development, plant design, equipment testing, as well as plant operation and maintenance. The efforts started at least as early as 1980s and have been continuing to evolve in the past decades. In particular, recent success of deep learning has inspired more exploration in the global industry. Several influential technical meetings and conferences have been organized.

This section presents some of the scenarios of AI applications in nuclear industry. These scenarios are not exhaustive, and it is intended to help capture the general landscape of this subject and facilitate the discussion on future standardization work in this direction. In addition, it also serves to foster more investigation and exploration of AI techniques in nuclear industry.

AI application potentials cover the full life cycle of a nuclear facility. The coming sections present applications that pertain to the scope of IEC SC 45A, and more examples beyond this scope are available in Annex A.

6.2 Virtual sensors

There are some process variables that cannot be directly measured but of significant value to plant operation and maintenance. In some cases, these variables can be inferred from available measurements by solving the first principle equations representing the physical mechanisms of performance. The inferred variable is usually called a virtual sensor.

In other cases, it is not easy to solve the analytical equations to find the correlation between the virtual sensor and other existing measurements. To address this issue, a high fidelity multi-physics model can be developed, such that values of virtual sensors and other measureable variables are generated. Then, a supervised learning algorithm can be explored to find the hidden correlation. The trained model is then deployed to operate as the virtual sensor.

6.3 Intelligent control

Though this document is intended to focus on non-safety related applications in nuclear industry, there are some ongoing research on intelligent control systems.

A nuclear reactor is a complex system, and comprehensive control of it is not trivial. Besides well-known control of thermal power and coolant temperature, reactor controllers take care of plenty of other aspects such as operational safety permitting operation only within given limits, homogenization of burnup, burnup compensation, compensation of the poisoning, shaping of the power density distribution, downstream conditions of the plant, support of flexible electricity production, operation economy, etc.

Machine learning based techniques can be used to improve the traditional core-control methods and to allow improved predictive control plant, especially for future reactors. AI allows consideration of an arbitrary large number of goals even if quite different in nature. In the case of reactor governing, safety, ergonomics, operation economy and grid services can be processed simultaneously in accord with each other.

6.4 Ageing management

Harsh environment in nuclear facilities causes significant ageing and degradation of materials, which might cause degraded plant performance or an unplanned shutdown. Ageing management program is developed to cope with this challenge. Degradation modeling is essential to address component aging problems and provide an accurate prediction of their failure points or remaining useful life (RUL). There are various models for estimating degradation of material or component in I&C and electrical systems. However, these models usually rely heavily on physics and expertise, which may have limited capacity to learn from massive measured or simulated data. In other words, when new data is available, it is not easy to improve the model.

Deep neural networks have found applications in this area. To predict the RUL for electric valves, a convolutional auto-encoder is used to extract features and a recurrent neural network is used for time-series data processing. Though the learning capabilities of deep neural networks are excellent, it is also advisable to be guided by the physics underlying the ageing process.

The RUL of insulation paper used in an NPP transformer is largely determined by the winding hot spot temperature (HST), which is not directly monitored and inferred from other measurements. The estimation of aging parameters is complex and non-deterministic. Accordingly, uncertainty modelling is critical for well-informed predictions. Integrating machine learning and experimental models in the Bayesian framework has proved to improve transformer RUL prediction.

6.5 Preventive 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, labour, and unnecessary planned, forced, or extended outages.

AI has been extensively explored and deployed to optimize maintenance decisions in many industries, including nuclear. It 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 a raft of other areas. As a matter of fact, preventive maintenance is one of the major fields where AI has found extensive and in-depth applications in nuclear facilities.

An example of a preventive maintenance task that can be improved through application of AI techniques is the periodic inspections of control rod drive mechanisms (CRDMs). Problems with the coils can cause the CRDM to inadvertently drop rods which can result in significant plant downtime. However, detecting these types of issues is difficult and often involves evaluation of thousands of coil current measurements by multiple human experts that can take hundreds of man-hours to perform. Supervised ML has been used to automate the detection of CRDM coil current issues in near real-time and predict the maintenance needs.

In addition, a significant part of NPP operations relate to staff walking around the plant to inspect the plant condition and environment and collect needed information. Computer vision and machine learning enables drones to recognize features of their surrounding environment, eliminating or reducing the need for the human rounds.

6.6 Anomaly detection

Currently, numerous NPPs rely on alarming systems to detect abnormal situations, which is less effective, as it is often too late for operators to take actions when the alarms appear. Lead time is desired to offer the operator more opportunity to elegantly plan and take necessary actions. Capabilities of detecting subtle anomaly, or small deviation from norm operating conditions, is essential to achieve that objective.

Anomaly detection is a well-established research field whose primary objective is to identify events or samples deviating from what it could be considered "ordinary" or "normal" behavior within an application domain. Machine learning community has developed a repertoire of tools to detect anomalies. Supervised, unsupervised and semi-supervised machine learning algorithms have been studied to identify anomaly in NPPs. For example, auto-encoder, an unsupervised deep learning algorithm, has been investigated to learn the hidden patterns in the huge volume of operation data, and then deployed to predict anomaly occurring in plant processes.

6.7 Operational decision support

In an operating NPP, a typical work process comprises of 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. Natural language processing (NLP) methods promise to mimic the human decision-making processes and improve the work process efficiency.

As an example, NLP has been used for automating and streamlining the corrective action program (CAP) process. Incident reports from NPPs are used to train the NLP models, which then can be used to predict whether the incident report has significant impact and requires reviewing.

NLP is also explored to improve preventative maintenance strategies by analysing maintenance records. In one research, 18 million maintenance work order records from 10 utilities over a few years are gathered. NLP is then used to compare the work order history of similar components across utilities and plants. This enables examining the impact of different preventive maintenance strategies on the overall maintenance.

6.8 Cyber security

With the increased use of digital systems in nuclear industry, maintaining effective cyber security 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 cyber attack in nuclear applications.

The workflow of applying AI techniques to cyber attack detection consists of four steps, namely, data acquisition, feature extraction, cyber anomaly detection, attack identification. In addition to network traffic data, process-specific features such as measurements, control commands, and set-points are also considered.

For cyber anomaly detection, various AI methods for anomaly detection can be applied based on the types of features extracted. Since the abnormal data related with cyber incidents are not sufficient in nuclear domains, the unsupervised learning methods which do not require two classes of tagged data, normal and abnormal, are suitable to address the cyber anomaly detection problem.

6.9 Human factor engineering

The Three Mile Island (TMI) and Chernobyl accidents revealed that plant operators may not always adequately handle challenging situations, and human factor engineering have played an essential role 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.

Operators' fitness for duty (FFD) has been highlighted as one the primary reasons for human error in nuclear accidents. Current FFD systems have been reported to be far from effective in practice, because many elements of the program involve tests that are subjective and infrequent. A continuous operator monitoring system can hence overcome the many of the challenges of existing FFD systems.

Recent studies show the potential of machine learning methods for monitoring a worker's FFD status using biosignals. Important biosignal markers of FFD are selected and fed to a multi-class support vector machine (SVM) classifier to achieve fast identification of a potential at-risk worker. In addition, NLP and computer vision (CV) techniques are also investigated to monitor operator cognitive factors. NLP empowers the automatic assessment of an operator's mental workload through oral assessments. CV based human tracking can be used to detect anomalous behavior of field workers.

7 Proposed structure for SC 45A AI standards

7.1 General

From Clause 6, it can be concluded that AI applications for nuclear facilities are very broad and extensive. Standards are very important to promote and facilitate its wide adoption in a fast and low cost way. It is envisioned that many standards are needed to accommodate the practical needs.

IEC SC 45A has a systematic approach to standard development work, best summarized in IEC TR 63400 which discusses the structure of SC 45A standard series. This clause outlines a nuclear AI standard structure that complies with current nuclear systems standards and fundamental state of practice and art.

7.2 Key criteria for structure design

7.2.1 Technical coverage

The extent of coverage that may be provided by the IEC SC 45A AI standard series is as follows:

- All types of nuclear facilities, including but not limited to nuclear power plants, research reactors and adjoining radioisotope production facilities; facilities for the enrichment of uranium; nuclear fuel fabrication facilities; facilities for the reprocessing of spent fuel; storage and disposal facilities for spent fuel and radioactive waste;
- All I&C systems and equipment and electrical power systems and equipment;
- All stages of the life cycle of I&C and electrical power systems and equipment in a nuclear facility;
- All the aspects or activities needed for developing and deploying an AI system in a nuclear facility.

7.2.2 Hierarchical levels

The structure shall follow the hierarchy of 4 levels used in IEC SC 45A standards, which is included here for reference as defined in IEC TR 63400.

L1: General requirements

L2: General requirements for specific topics

L3: Specific requirements for specific topics

L4: Technical Reports (which are not normative)

The reader is reminded to distinguish these levels from those used in Table 1 where intended to indicate AI autonomy levels, which should be self-evident in the context.

7.2.3 Entry point documents

According to IEC TR 63400, an entry point document is the term applied to the small number of IEC SC 45A standards that enable a new user of the standard series to most effectively find their way into the series and explore the hierarchy.

In this respect, it is desired there will be a single entry point document for the SC 45A AI standard series. This document will cover the full scope (or a major part of the scope) at a general level and other documents will go into further detail for parts of that scope.

However, it is worthwhile to note that this entry point document will not be an L1 standard. As defined in IEC TR 63400, only IEC 61513 and IEC 63046 belong to the category of L1 standards.

7.2.4 Reference to other standards

Many IEC SC 45A standards make reference to other SC 45A standards and, where applicable, also make reference to various non-SC 45A IEC and ISO standards rather than repeating their requirements.

As a general criterion, the developed structure will refer to generic international standards (e.g. IEC, ISO) that have been developed for general purpose AI applications.

When there is a generic standard for a topic which cannot be used directly or which needs adaptation for the nuclear sector, guidelines are developed for its use or "nuclear equivalent" standards are developed. In such cases, the IEC SC 45A "application" guidelines and "nuclear equivalent" standards make reference to and adhere as closely as possible to the generic international standard.

7.3 Structure of AI standard series

The 4-level hierarchy adopted in IEC SC 45A standards, as described in IEC TR 63400, provides a good starting points for the structure design on nuclear AI standards.

Internationally, there are a number of AI standardization roadmaps that have been published. It is common practice in these roadmaps to organize the structure in horizontal and vertical standards.

Taking both of the abovementioned factors into consideration, this document proposes to arrange the SC 45A AI application standard structure in three levels, which are standards on horizontal requirements, standards on vertical requirements, and technical reports (which are not normative) respectively. These levels essentially correspond to L2, L3, and L4 as defined in IEC TR 63400. In this sense, the structure naturally follows the general structure adopted by SC 45A.

The proposed structure is shown in Figure 6.

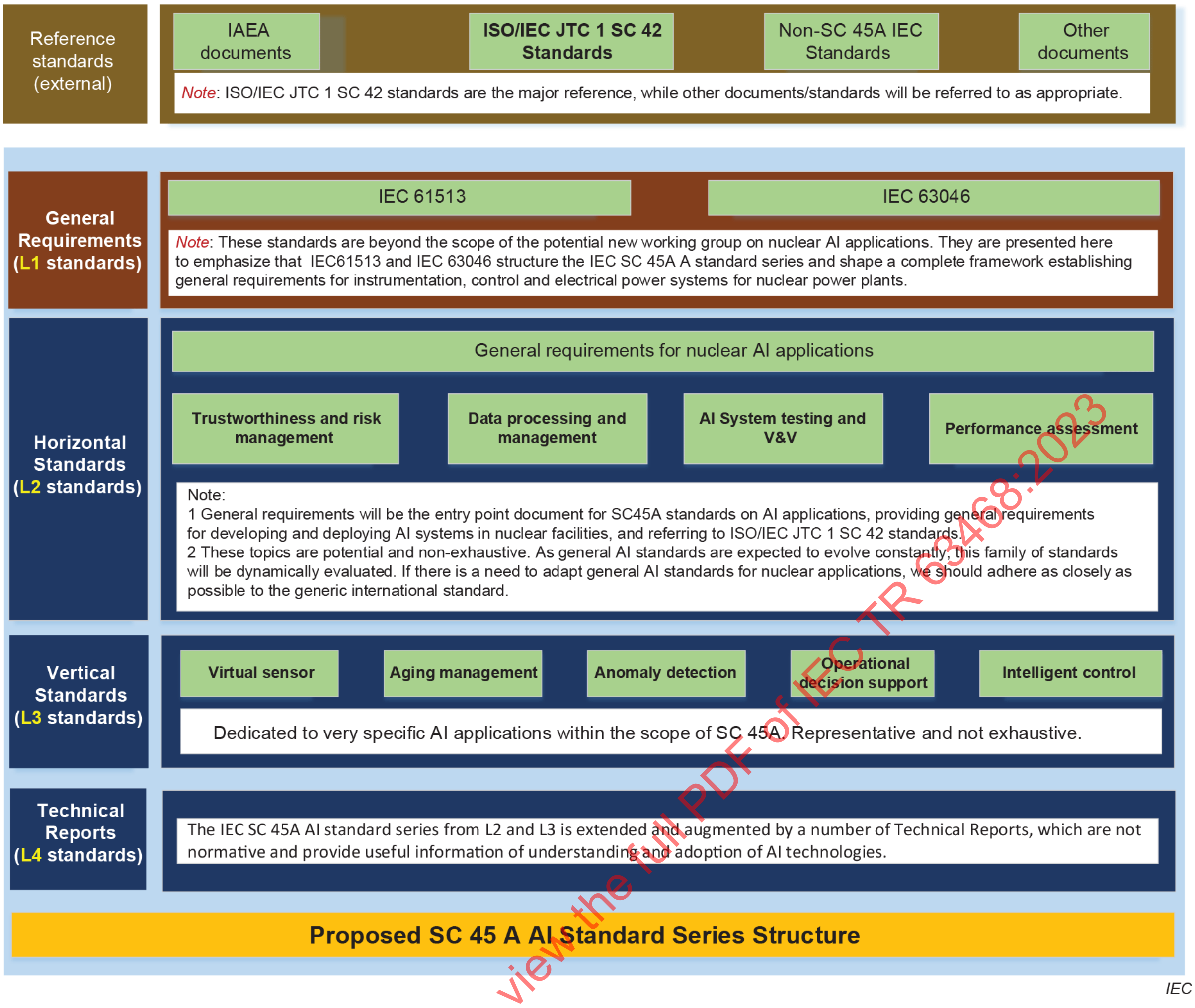

Figure 6 – Proposed structure for SC 45A AI standards

8 Near-term development priorities

The structure for the SC 45A standard series on AI applications, as proposed in the preceding section, will be the fundamental reference for discussions on priority new work items.

A systematic approach is needed to better coordinate work on AI standard development within SC 45A. The priority items are evident if the intrinsic logic is respected in the abovementioned structure.

The first priority is to develop the entry point document in L2, which will provide general requirements for developing and deploying AI systems in a nuclear facility. This will enable the community to have a common language when talking about AI techniques.

It follows that other L2 standards should be in place. This will enable the SC 45A community to have a common position when talking about specific characteristic or issues of AI techniques.

Vast work is foreseen to hinge on vertical applications, i.e. L3 standards. These standards will be diverse in nature, and more experts who are traditionally outside the SC 45A community will join and contribute. They will produce the “practical value” for end users, but the success of these vertical standards rely on the existence of L2 standards. There is no specific priority in this family of vertical standards, initiatives from experts and maturity of applications will play a central role in which standards are developed earlier than others.

L4 documents may be prepared as necessary during the process of expanding the hierarchy of SC 45A AI standards, capturing information valuable for understanding and adoption of AI technologies.

9 Organizational challenges and recommendation

9.1 Cross-cutting characteristics of nuclear AI standards

AI is an enabling technology, just like that combustion engines or electricity can be used across many industries. This remains true when we are considering AI applications in nuclear facilities. As can be seen from the non-exhaustive overview of applications in Clause 6, AI is enabling many different areas of nuclear facilities.

From the perspective of AI standard development, the diverse applications areas mean that this topic is cross-cutting with many of the working groups of SC 45A. There are 8 working groups of SC 45A, as listed in Table 2.

Table 2 – Working groups of IEC SC 45A

Working group	Title of the working group
WGA2	Sensors and measurement techniques
WGA3	Instrumentation and control systems: architecture and system specific aspects
WGA5	Special process measurements and radiation monitoring
WGA7	Functional and safety fundamentals of instrumentation, control and electrical power systems
WGA8	Control rooms
WGA9	System performance and robustness toward external stress stems
WGA10	Ageing management of instrumentation, control and electrical power systems in NPP
WGA11	Electrical power systems: architecture and system specific aspects

Based on samples of AI applications in nuclear facilities from Clause 6, and the responsibilities of the SC 45A working groups, Table 3 is prepared to capture the cross-cutting nature of this topic. As more applications are identified, the working groups of relevance to a particular area will expand accordingly.

Table 3 – Cross-cutting between nuclear AI applications and SC 45A working groups

Selected application examples	Typical AI methods	Relevance to SC 45A working groups
Virtual sensors	supervised learning	WGA2, WGA5
Intelligent control	machine learning	WGA7, WGA8, WGA9, WGA11
Ageing management	recurrent neural network	WGA10
Preventive maintenance	supervised learning	WGA2, WGA9, WGA10, WGA11
Anomaly detection	clustering	WGA2,WGA5,WGA9,WGA10, WGA11
Operational decision support	NLP	WGA8
Cyber security	machine learning	WGA9
Human factor engineering	NLP, CV	WGA8

In addition to the specific applications and their relevance to the working groups, it is also important to look at the general characteristics of AI technologies and their relevance to the working groups (see Table 4). These general topics, whether or not identified here, are expected to be covered by the relevant working groups respectively. To put it another way, current working groups will take these general topics into account when develop or revise their documents, and these new requirements will be followed by other specific AI standards in SC 45A.

Table 4 – Cross-cutting between general AI topics and SC 45A working groups

	General topics of AI techniques	Relevance to SC 45A working groups
1	V&V of AI systems	WGA3
2	Function categorization of AI systems	WGA7
3	Life cycle of AI systems	WGA3
4	Security of AI systems	WGA9
5	AI integration with human operators	WGA8
6	Risk analysis of AI systems	WGA7

9.2 Organizational challenges

The complex nature of AI techniques and their applications pose special challenges to SC 45A with regard to the question who will be organizing the standard development activities in this emerging area.

It is self-evident from the preceding subclauses that AI applications in nuclear facilities cover topical areas that are diversely dispersed and beyond the scope of any specific working group in SC 45A. In other words, no single working group can handle the majority of AI topics. It does not take much effort to notice that AI has cross-cutting with almost every working group in SC 45A.

Subsequently, this raises the organizational challenges for SC 45A. How to effectively and efficiently coordinate the AI standard development efforts within the subcommittee so that it can better serve the industry needs in a consistent manner?

These challenges are not easy to resolve under current working group settings.

9.3 Recommendations

9.3.1 General

These organizational challenges are by no means unique for SC 45A. Similar challenges are present for ISO/IEC JTC1 which is the technical committee for Information Technology. With the advent of AI boom in recent years, it is found that current sub-committees are not appropriate to accommodate the diverse needs of standard developments on AI. That is why a new sub-committee, ISO/IEC JTC1 SC 42, was created to coordinate the efforts in this field.

The same approach is recommended to resolve the organizational challenges of SC 45A. A dedicated new working group can be created, to focus on the AI aspects of nuclear facilities. The details are described as follows.

In addition, it is worthwhile to note that this recommendation is also a fulfillment of the requirement in 45A/1363A/DC and 45A/1374/INF. Taking into account crosscutting nature of this topic, in-depth cooperation with all of the existing WGs in SC 45A will be essential in order to draft new and revised standards, so as to sustain consistency and harmony of SC 45A standard portfolios. When close collaboration is needed, project teams may be appropriate for such work.