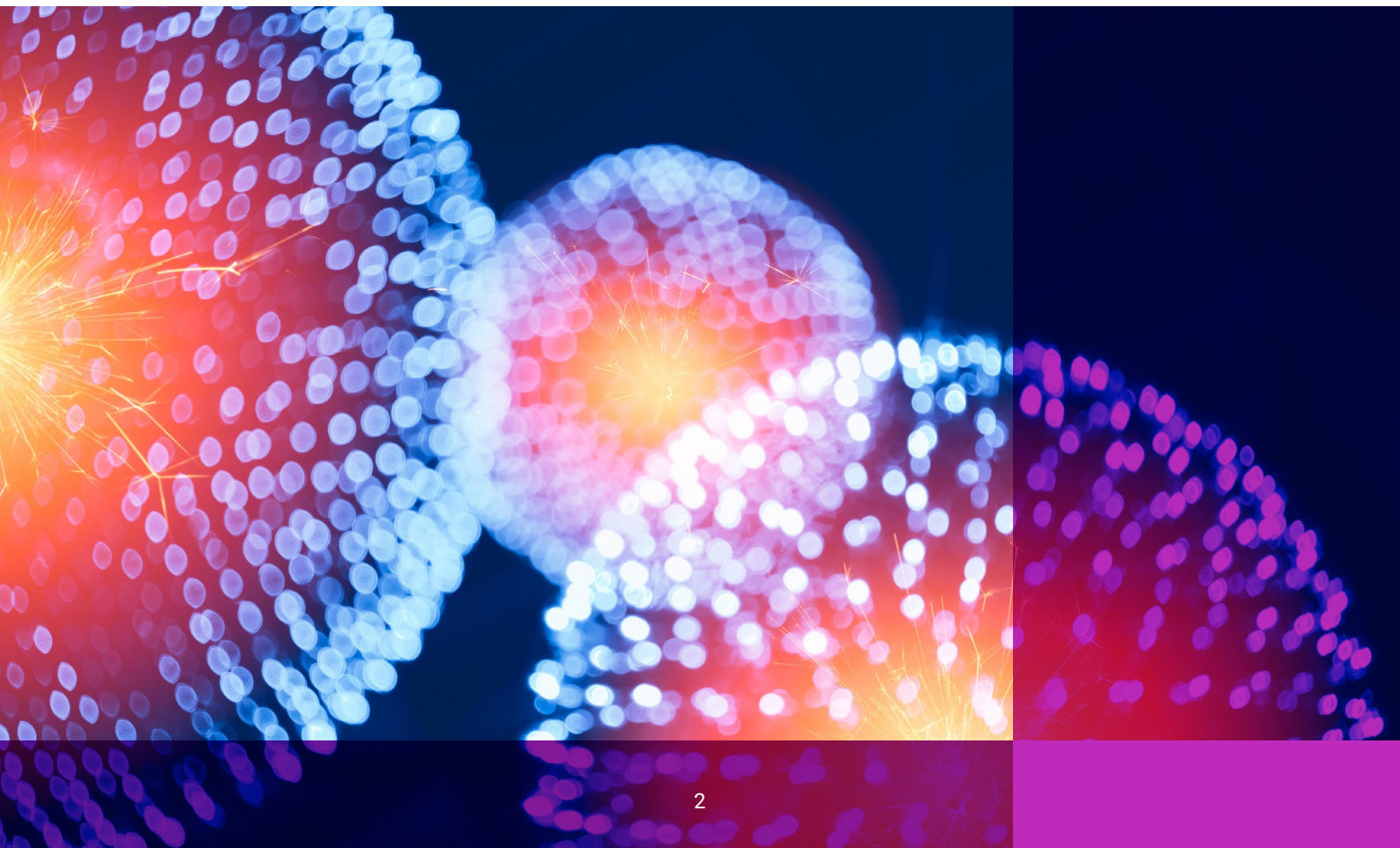
Innovate UK | Business Connect

Information and safety management in the nuclear industry

NLP Applications Edition

Contents



Introduction

This report from the AI for Nuclear (AI4N) group, will focus on the applications of Natural Language Processing (NLP) in the nuclear industry.

AI4N was launched in 2022 and immediately identified areas where Artificial Intelligence (AI) can be introduced in the nuclear sector to help overcome nuclear industry challenges. The details were published in the AI4N special edition.

This edition aims to deliver some of AI4N actions by sharing lessons learnt from case studies that demonstrate the use of NLP in the nuclear sector.

AI4N recognises the exponential growth of AI technology which has resulted in new opportunities and challenges in terms of improving the safety and economics of the nuclear industry. We also recognise that for nuclear energy to be more competitive and integral in modern energy systems, the nuclear industry must be not only economical and efficient but also safe, reliable, and sustainable. Natural Language Processing (NLP) has a big role to play in modernising the nuclear industry. For example, NLP can assist in nuclear decommissioning by extracting relevant information from historical data. The following are some of examples the applications of NLP in the nuclear industry.

Decommissioning & Operations

By using NLP, the nuclear industry can enhance safety, optimise operations, and improve decision-making processes. NLP has the potential to streamline information management, enable proactive maintenance, and support the overall efficiency and effectiveness of nuclear power plant operations and decommissioning.

- **Question Answering Systems**
  NLP models can be used to develop question answering systems that can understand and respond to queries related to nuclear decommissioning. These systems can provide quick access to specific information, reducing the time spent on searching through vast amounts of documentation. Workers can interrogate Life Time Quality Records (LTQRs) and ask questions about safety procedures, best practices, or technical specifications, and receive accurate and relevant answers.

- **Machine Translation**
  In an international context, where nuclear decommissioning projects often involve collaboration between teams from different countries, NLP can facilitate multilingual communication. Machine translation techniques can help translate training manuals, technical documents, safety guidelines, or communication materials between different languages, ensuring effective knowledge transfer and understanding.

- **Incident Analysis**
  NLP can be used to analyse incident reports, safety logs, and maintenance records to identify patterns and extract valuable insights. By processing textual data, NLP algorithms can help identify recurring issues, and root causes of incidents, and suggest preventive measures to enhance safety protocols.

- **Maintenance and Operations**
  NLP can assist in optimising maintenance schedules and predicting equipment failures. By analysing maintenance logs, sensor data, and historical records, NLP along with Machine Learning algorithms can identify patterns indicative of potential failures, allowing for proactive maintenance interventions and minimising unplanned downtime.

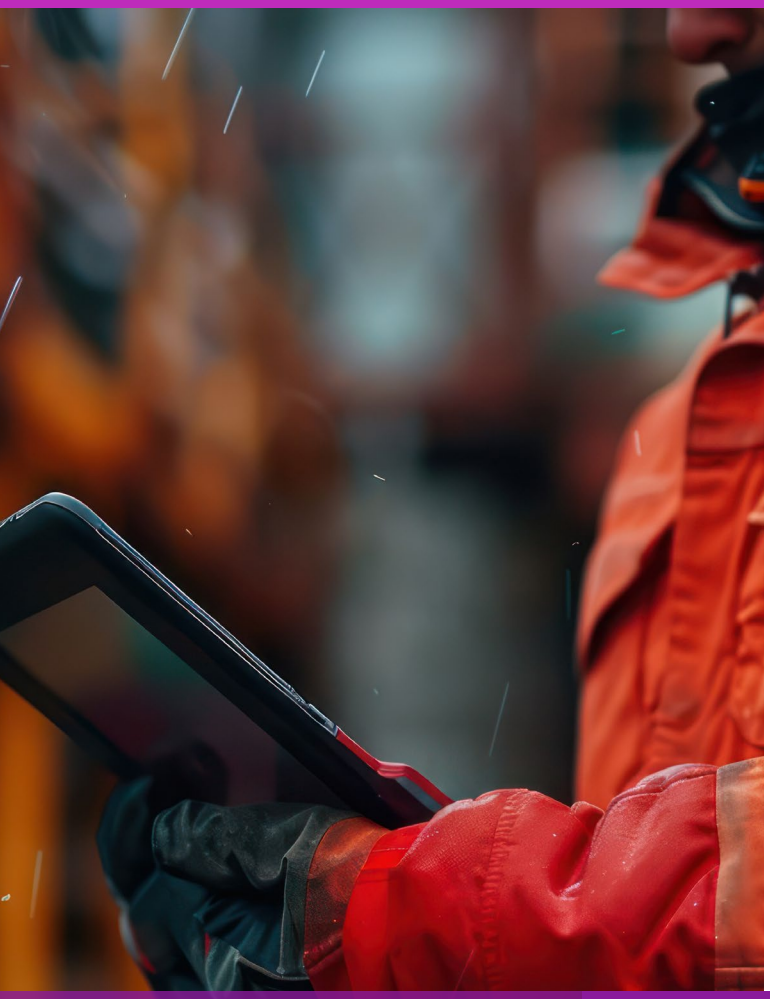
- **Safety Documentation Analysis**
  Nuclear power plants have extensive safety documentation, including manuals, procedures, and regulatory guidelines. NLP can help analyse and extract relevant information from these documents, making it easier for operators and engineers to access critical information quickly. This can aid in implementing safety protocols, ensuring compliance, and facilitating knowledge sharing.

- **Operator Training and Support**
  NLP can be used to develop interactive training modules and support tools for nuclear power plant operators. Virtual assistants powered by NLP can provide real-time assistance, answering questions and providing guidance on operational procedures, emergency response, and safety protocols. This can improve the efficiency and effectiveness of operator training and support systems.

- **Event Detection and Monitoring**
  NLP techniques can be employed to monitor and analyse various data sources, such as sensor readings, maintenance reports, and operational logs, to detect abnormal events or deviations from normal operating conditions. This can aid in the early detection of potential safety hazards or equipment malfunctions, enabling timely interventions.

Business Support for the Nuclear Industry

As NLP technology continues to advance, it holds the potential to enhance various aspects of the industry, ranging from research and development to safety and regulatory compliance. The following showcase some applications of NLP in the nuclear industry.



- **Information Extraction:** NLP techniques can be used to extract relevant information from large volumes of textual documents, such as technical reports, safety manuals, and regulatory guidelines. This can help decommissioning teams identify key insights, important requirements, and potential risks associated with a process.
- **Document Classification:** NLP algorithms can classify documents based on their content, enabling teams to organise and categorise large document repositories efficiently. For example, documents can be classified into categories such as waste management, radiation safety, regulatory compliance, or equipment dismantling. This facilitates document management during a project.
- **Horizon Scanning:** NLP can be used to analyse and extract information from vast amounts of text-based documents such as research papers, technical reports, regulations, and safety guidelines. This can help researchers and professionals in the nuclear industry to quickly access relevant information, identify trends, and make informed decisions.

- **Safety and Risk Assessment:** NLP techniques can assist in analysing safety reports, incident reports, and inspection records to identify patterns, trends, and potential risks. By automatically processing textual data, NLP can help identify safety-related issues, predict potential risks, and improve safety measures in nuclear facilities.
- **Regulatory Compliance:** NLP can aid in interpreting and understanding complex regulatory documents and compliance requirements. By analysing regulations and guidelines, NLP algorithms can help ensure that nuclear facilities adhere to the necessary safety standards, reducing the risk of regulatory non-compliance. NLP can also help in identifying changes in regulations that will help ensure that regulatory compliance remains up to date.
- **Knowledge Management:** NLP can be employed to build comprehensive knowledge bases by extracting and organising information from various sources. This can include gathering data from scientific publications, industry reports, and internal documentation. Such knowledge bases can facilitate efficient information retrieval, foster collaboration, and assist in decision-making processes. This will also ensure that any critical knowledge is not lost when professional staff leave an organisation.
- **Virtual Assistants and Chatbots:** NLP-powered virtual assistants or chatbots can provide quick and accurate responses to queries related to nuclear energy, safety protocols, and operational procedures. They can assist employees, researchers, and the general public by answering frequently asked questions, providing guidance, and disseminating relevant information.
- **Sentiment Analysis:** NLP can be used to analyse public sentiment and opinions related to nuclear energy through social media, online forums, and news articles. This can help gauge public perception, identify concerns, and enable better communication strategies regarding nuclear energy and its associated topics.

Nuclear Power Plant

By using NLP, nuclear power plants can enhance safety, optimise operations, and improve decision-making processes. NLP has the potential to streamline information management, enable proactive maintenance, and support the overall efficiency and effectiveness of nuclear power plant operations.

- **Incident Analysis:** NLP can be used to analyse incident reports, safety logs, and maintenance records to identify patterns and extract valuable insights. By processing textual data, NLP algorithms can help identify recurring issues, and root causes of incidents, and suggest preventive measures to enhance safety protocols.
- **Maintenance and Operations:** NLP can assist in optimising maintenance schedules and predicting equipment failures. By analysing maintenance logs, sensor data, and historical records, NLP along with Machine Learning algorithms can identify patterns indicative of potential failures, allowing for proactive maintenance interventions and minimising unplanned downtime.
- **Safety Documentation Analysis:** Nuclear power plants have extensive safety documentation, including manuals, procedures, and regulatory guidelines. NLP can help analyse and extract relevant information from these documents, making it easier for operators and engineers to access critical information quickly. This can aid in implementing safety protocols, ensuring compliance, and facilitating knowledge sharing.
- **Operator Training and Support:** NLP can be used to develop interactive training modules and support tools for nuclear power plant operators. Virtual assistants powered by NLP can provide real-time assistance, answering questions and providing guidance on operational procedures, emergency response,

and safety protocols. This can improve the efficiency and effectiveness of operator training and support systems.

- **Regulatory Compliance:** NLP can assist in interpreting and understanding complex regulatory documents and requirements. By automatically processing regulatory guidelines, NLP algorithms can help identify gaps, assess compliance, and ensure that nuclear power plants adhere to the necessary safety standards.
- **Event Detection and Monitoring:** NLP techniques can be employed to monitor and analyse various data sources, such as sensor readings, maintenance reports, and operational logs, to detect abnormal events or deviations from normal operating conditions. This can aid in the early detection of potential safety hazards or equipment malfunctions, enabling timely interventions.
- **Public Perception Analysis:** NLP can be used to analyse public sentiment and opinions related to nuclear power plants through social media, news articles, and public forums. This can help plant operators and regulatory bodies gauge public perception, identify concerns, and design effective communication strategies.

In both the decommissioning and power plant activities NLP can be employed to analyse reports, surveys, research papers, and other text-based sources to extract valuable information related to the nuclear industry's energy consumption, water usage, carbon emissions, waste management, and sustainability initiatives. Techniques like named entity recognition, topic modelling, and sentiment analysis can help identify key entities, relevant topics, and the overall sentiment expressed in the text. By combining insights from the sparse measurement data and the NLP analysis, a carbon 'assistant' that can gain a more comprehensive and contextual understanding of the nuclear industry's carbon footprint can be created. For example, NLP can help identify specific areas in reports where certain environmental concerns are raised, which can be linked to relevant measurements and Earth Observation data to validate and prioritise actions.

NLP has capability to increase diversity and inclusion in nuclear industry. For example, NLP can be utilised to analyse job postings, descriptions, and candidate resumes to identify any biased language that may inadvertently discourage diverse candidates from applying. NLP can assist in identifying biases in training materials, ensuring that content is inclusive and accessible to all employees, including neurodiverse individuals. It can also help tailor learning programs to the unique needs of employees, fostering a supportive and inclusive learning environment.

Sentiment analysis using NLP can help gauge the sentiments and opinions of employees from diverse backgrounds. Understanding their experiences and feedback can enable organisations to address concerns and create a more inclusive work environment.

To realise the role of NLP in the Nuclear Industry, the AI4N steering committee organised a workshop which was held in 2023 to demonstrate the role of NLP in the nuclear industry.

The Workshop

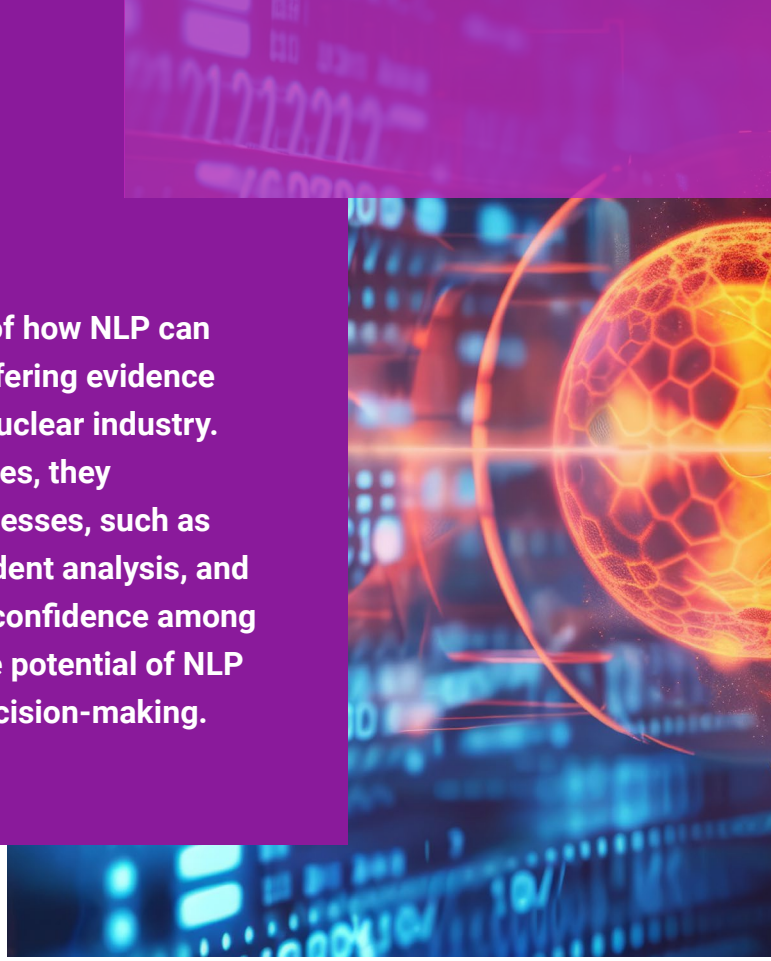
The workshop had 54 attendees with an introductory session followed by the case studies which demonstrated the use of NLP in the nuclear industry.

The opening session for the workshop discussed current news items related to NLP; firstly, the Large Language Models such as ChatGPT, their potential as well as the risks of applying them in the nuclear industry. Among the risks includes, inaccurate or misleading information, lack of domain-specific expertise, safety and security risks, ethical concerns, lack of accountability and transparency, dependency and human complacency (Over-reliance on large language models can lead to reduced human involvement, potentially fostering complacency among human operators and experts).

Secondly, the use of AI in the EU will now be regulated by the AI Act to ensure that AI developed and used in Europe complies with EU rights. The European Union is pressing ahead with its flagship AI Act governing the safe and transparent application of artificial intelligence. As a result, the EU is proposing a full ban on AI for biometric surveillance, emotion recognition and predictive policing, while generative AI systems like ChatGPT must disclose that resulting content is AI-generated. EU is also classifying AI systems used to influence voters in elections as high-risk, and they may be banned.

Thirdly, the UK AI government recognises that AI is already delivering real social and economic benefits for people. The government does not want organisations to be held back from using AI to its full potential resulting from the patchwork of legal regimes that could cause confusion for businesses trying to comply with rules. With that, the government will not give responsibility for AI governance to a new single regulator, instead, it wants the existing regulators - such as the Health and Safety Executive, Equality and Human Rights Commission and Competition and Markets Authority - to come up with their own approaches that suit the way AI is actually being used in their sectors.
A pro-innovation approach to AI regulation which is addressing the regulation issues was published in March this year.

Case studies



These case studies provide concrete examples of how NLP can be effectively applied in the nuclear industry. Offering evidence of the effectiveness of NLP applications in the nuclear industry. By showcasing real-world scenarios and outcomes, they demonstrate how NLP can enhance various processes, such as document analysis, compliance monitoring, incident analysis, and knowledge management. This validation builds confidence among stakeholders and decision-makers regarding the potential of NLP to improve operational efficiency, safety, and decision-making.

How AI reduced the administrative burden by 90%

Sellafield, the birthplace of the UK's nuclear industry, continues to play an important role as the storage centre for the country's nuclear waste. To operate safely, the teams on site rely on scores of processes set out in complex documents. Every time legislation or regulations change, these documents and processes may have to change to stay compliant and pass inspections.

Until Sellafield's collaboration with PA Consulting, the process of checking for these changes and working out what they mean for Sellafield remained highly manual and could take weeks per document. Sellafield saw an opportunity to be industry leaders by adopting AI to radically accelerate and enhance their existing specialised processes.

We built an AI solution (combining Machine Learning and Natural Language Processing) that identifies 1) what's new, 2) what's relevant to you, and 3) what's the impact – for example, which Sellafield documents would need to be updated as a result of a specific legislative or regulatory change.

This solution has:

- Cut through the existing complexity associated with text data
- Accelerated document review from weeks to minutes
- Reduced overall administrative burden by 90%
- Most importantly, given engineers time back to focus on the most important activities – evaluating the "so what" of incoming change

Derreck van Gelderen
PA Consulting – Head of AI Strategy and AI Energy lead

Using NLP to extract information from Deep Storage archives

The Energy, Security & Technology (ES&T) business unit within Jacobs currently has approximately 25,000 boxes held in a secured offsite storage location. These boxes contain information relating to nuclear projects, dating back over 40+ years. The information contained within these boxes could be vital to support future research and projects. In a pilot project, 200 boxes were identified and retrieved from the secure offsite storage. 2943 documents from these 200 boxes were meticulously scanned and converted into PDF files.

In total nearly half a million pages of information was converted into a digital 'Knowledge Base' by a propriety software. A number of complex technical questions were asked related with reactor design and nuclear site investigations. The search results were provided in order of relevancy along with other related topics which may be of interest. Each search provided accurate information and took less than 20 seconds to complete. A manual search would have taken days.

Prof. Nawal K Prinja
Jacobs – Technology Director

IFE Document AI: Auto text annotation using NLP for Licensee Event Report (LER)

The nuclear power industry has a long history of safety initiatives, with event reporting being a consolidated practice. Analysing the recorded information, usually in the form of technical reports, is a difficult and time-consuming task. To address this challenge, Natural Language Processing offers tools to help decision-makers gain better insights in a shorter time. A case study leveraging these advanced techniques is performed, which extracts information from Licensee Event Reports (LER). These reports provide detailed accounts of incidents during events, encompassing component or system failures, along with corrective actions taken or planned to mitigate recurrence.

The research explores two distinct approaches: transfer learning and active learning. Transfer learning involves an algorithmic method that iteratively queries a user to label new data points with desired outputs, potentially achieving higher accuracy with fewer training labels. On the other hand, active learning constitutes a subfield of machine learning concerned with storing knowledge acquired while addressing one problem and applying it to a different yet related problem.

Active Learning, represented by the INCEpTION application, offers an intuitive interface that enhances the efficiency and speed of annotating text. Despite its

user-friendly interface, INCEpTION operates as a standalone application, posing a challenge when integrating it into other systems. Transfer Learning, exemplified by implementing the spaCy framework, provides a convenient package suitable for production and seamless integration. As a conclusion, in our application, we utilise INCEpTION for generating the training set due to its ease of use, while integrating spaCy into our system to achieve improved results in terms of precision and other performance metrics.

Chau Thi Thuy Tran
Researcher – Applied Data Science Dept, Institution for Energy (IFE)

While the recent advancements are promising, the AI4Nuclear group acknowledges the increasing utilisation of NLP-based generative AI tools, such as OpenAI's ChatGPT and Microsoft's Copilot, over the past year. Understanding both their capabilities and limitations is crucial. Despite their impressive ability to generate human-like text, these tools present specific challenges, some of which are discussed on **Maddyness.com** and outlined below.

1. **Contextual Understanding:** These tools may falter with complex technical subjects, potentially leading to imprecise responses or misinterpretations of specific queries.
2. **Bias and Training Data:** As these models are trained on extensive existing text datasets, they are susceptible to inheriting biases which can be reflected in their outputs, sometimes perpetuating stereotypes or disseminating misinformation.
3. **Information Accuracy ("Hallucination"):** Generative AI tools may produce entirely fictitious content with undue confidence, termed "hallucination." Users must be vigilant and critical of the information provided.
4. **Source Attribution:** Unlike human authors, these AI models do not fully cite sources, complicating the verification of information accuracy.
5. **Safety-Critical Applications:** It is prudent to avoid relying solely on AI for decisions in safety-critical areas, such as medical diagnoses or engineering tasks. Human supervision is indispensable to ensure validity and reliability.

Discussions are underway to validate strategies aimed at mitigating the aforementioned constraints. One proposed approach involves avoiding the use of generic datasets for training generative AI models in favour of data that is specific to the business context. While this adjustment can yield improvements, it is crucial to acknowledge that the proficiency of Generative AI primarily lies in its ability to select appropriate vocabulary and construct sentences that are grammatically sound and persuasive. However, this does not inherently guarantee the accuracy of the content produced, regardless of whether the training data is generalised or context-specific.

For this reason, the AI4Nuclear group believes that Generative AI tools are best used to enhance human judgment rather than replace it. Humans must be in the loop, especially in contexts where safety is crucial.

Conclusion and next steps

This edition of the workshop has effectively showcased the considerable potential of Natural Language Processing (NLP) in enhancing information management and safety within the nuclear industry. Furthermore, the workshop has successfully realised one of the plans and objectives outlined by the AI4N group.

AI4N will further investigate best practices for large language models (LLMs) and identify methods to develop a framework that will support the sector in benefiting from the technology while minimising risks.

The presented case studies during the workshop have demonstrated that NLP is a pivotal technology capable of bolstering safety measures in the nuclear sector. It offers valuable support in managing a broad spectrum of information-related tasks, including regulatory document handling, public communication, training initiatives, and event detection and monitoring.

The remarkable outcomes of this workshop, which substantiate the power of NLP, present a compelling challenge: the widespread implementation of NLP within the nuclear sector in the United Kingdom must be advocated for and embraced. However, the adoption of NLP tools in this sector should proceed with caution. Tools such as those based on generative AI still have significant limitations and require human oversight when being used.

AI for Nuclear Steering Group

Dr Caroline Chibelushi
AI Lead, Innovate UK Business Connect

Dr Melissa Willis
RATIRT Lead, Sellafield Ltd

Prof. Nawal Prinja
Technology Director, Jacobs

Ray Chegwin
Nuclear Lead, Innovate UK Business Connect

Dr Mark Bankhead
Technology Manager, NNL

Andrew Cooney
Innovation Lead, Enterprise R&D, Sellafield Ltd

Carl Dalby
Group Head of AI & Digital, Nuclear Decommissioning Authority

Innovate UK

Jacobs

NATIONAL NUCLEAR LABORATORY

the nuclear institute

Sellafield Ltd

ukri.org/councils/innovate-uk

03003 214357 • support@iuk.ukri.org • @innovateuk

Innovate UK, Polaris House, North Star Avenue, Swindon, SN2 1FL