Case Study: Overview and Findings from Regulatory Sandbox of Two Nuclear AI Applications
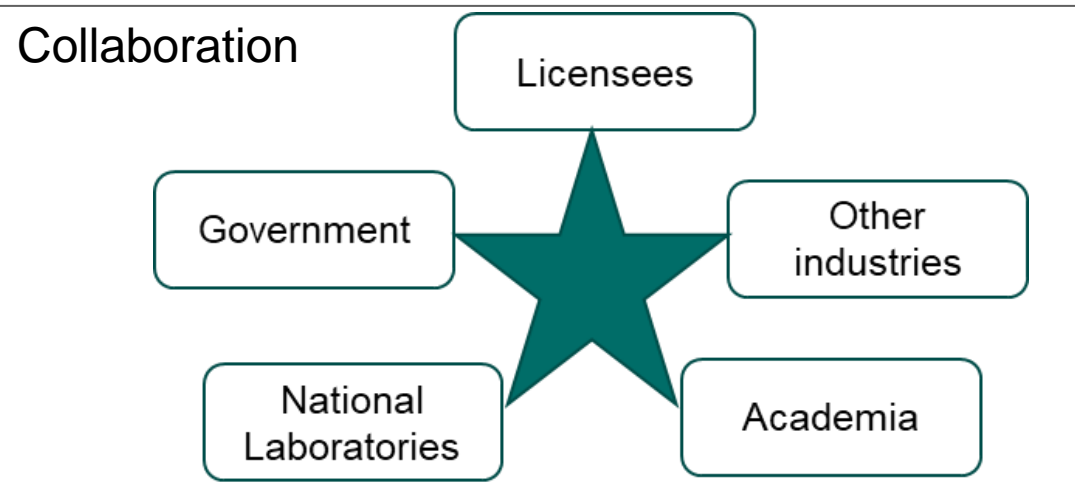
Andy White & Jake Surman

18 October 2024

Expert Panels & AI Research

Expert panels on the regulation of AI

Joint initiative with the Advanced Nuclear Skills and Innovation Campus (ANSIC) and with the Environment Agency

Aim - to establish a roadmap for effective and enabling regulation of AI in the nuclear sector.

Collaboration

Research – "Impact of AI / ML on nuclear regulation"

Research considerations
- Regulator capability and framework
- Evidence generation
- Data, architectural approaches, and cyber threats
- Standards development

www.onr.org.uk/documents/2021/onr-rrr-121.pdf

ONR Office for Nuclear Regulation

Environment Agency

Expert panel: Developing an approach to the regulation of AI

Three broad opportunities for the deployment of AI:

- Advisory

- Supervisory

- Control

Increasing challenge.

Development of AI systems
Good development lifecycle, configuration control, training data, cyber protection

Understand performance characteristics of the AI systems
How to understand performance and transfer info from other sectors

Confidence in performance of AI systems
Challenges with testing. Phased to build confidence / experience

Dealing with failure
Define / recognise failure. Use existing models (e.g. defence in depth)

Develop skills and experience including understanding the complexities of behaviours between humans and AI

ONR Office for Nuclear Regulation

Environment Agency

Regulatory sandboxing

A regulatory sandbox is a controlled real-life or <u>realistic</u> environment (digital, physical or table-top exercise) whereby regulators and industry <u>collaborate</u> to test innovations, e.g. new products, services, business models or technologies. The intent of a regulatory sandbox is to <u>explore</u> how the sandboxed innovation can be <u>regulated and safely deployed</u>
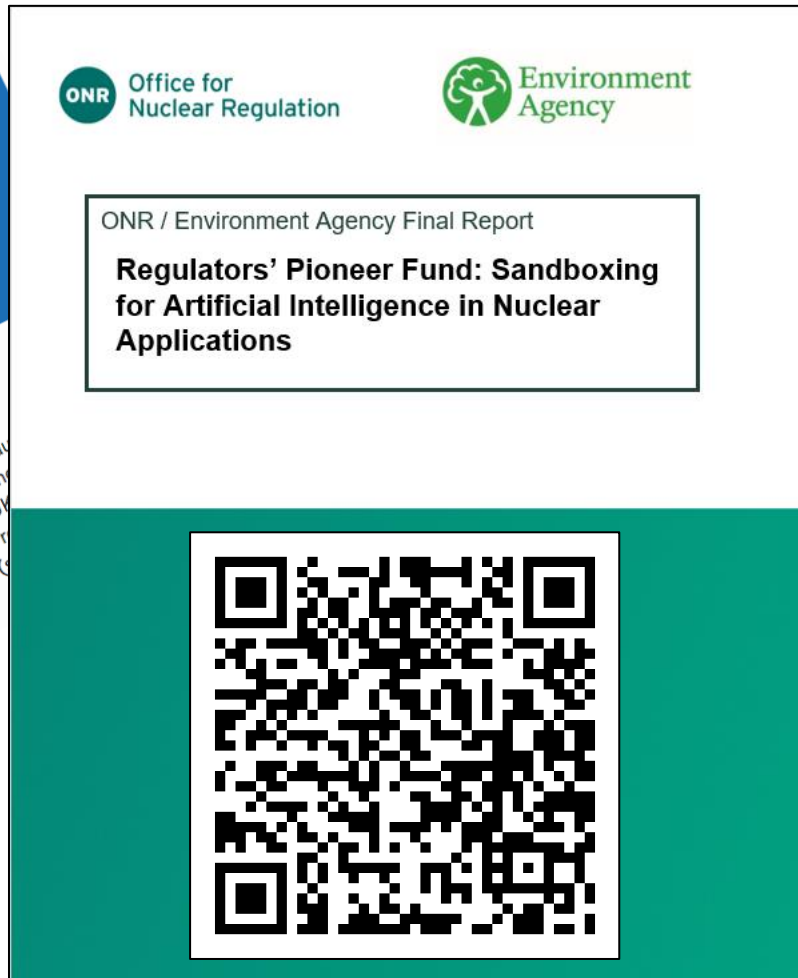
Regulatory Sandbox: Two Nuclear Applications of AI

Regulatory Sandboxing: AI applications

Structural integrity:

Use AI to derive information from plant to inform structural integrity claims in a safety case to help demonstrate reliability.
It is thought that this will assist in the development of digital twins and probabilistic assessment to demonstrate asset in-service operational life.

Robotic arm in a glovebox:

Use AI for real-time application to inform operations and understanding stresses and potential environmental constraints to, for example, optimise robotic movements.

Regulatory Sandbox Pilot
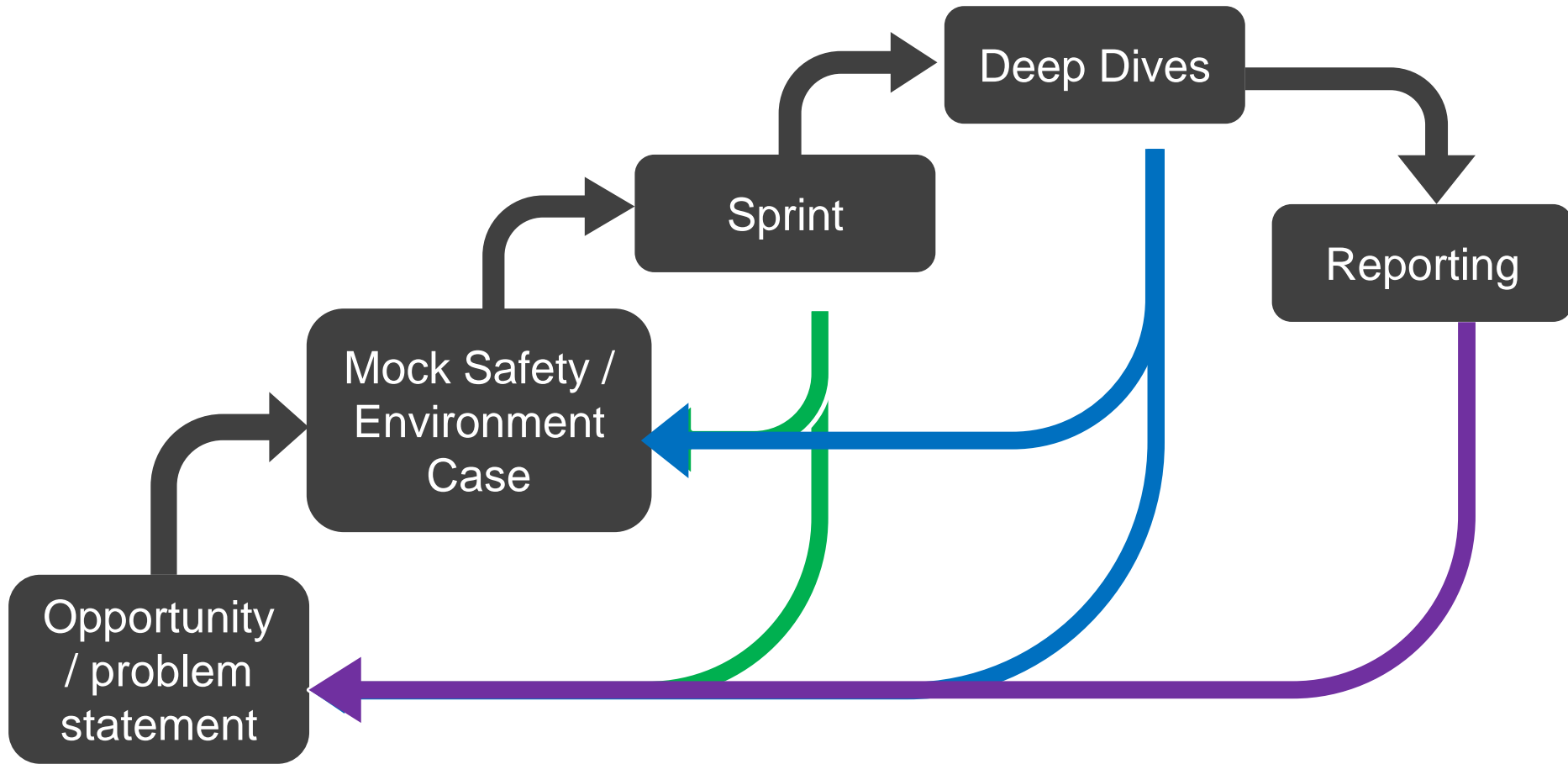
Regulatory sandbox

Evaluated as successful from a regulatory and industry perspective, it:

- Provided a safe, collaborative space to allow industry experts & regulators to express views in an open-minded and inclusive manner. Developing a collective / shared understanding.

- Focused on the practical aspects and the crux of the issues related to deployment, via considering AI components as part of a whole system.

- Enabled identification of existing good practice relevant to the innovation which could help enable it's deployment & where further work is needed.

Overall, allowed industry access to regulatory view & regulators to gain an understanding of an area of innovation & how it could be deployed in a safe & secure manner, maintaining protection of the environment.

ONR Office for Nuclear Regulation

Environment Agency

AI Sandboxing: Findings (1/3)

- Benefits of AI should be clearly articulated and compared with alternative (traditional) technologies.

- Risks associated with AI need to be understood and managed through robust arrangements e.g.

  - The level of authority associated with the AI system
  - The level of safety, security and environmental significance
  - The level of continuous learning

- Deployment of AI systems should be phased in order to build confidence and experience

ONR Office for Nuclear Regulation

Environment Agency

AI Sandboxing: Findings (2/3)

- Key considerations for the use of AI differed for the different applications which may suggest a principles based approach to regulation is preferred.

- Substantiating AI is difficult, so we should consider the consequences of its failure and manage these instead. The performance of the AI could still be monitored through understanding the systems functional requirements.

- When transferring the AI to a new operation or phase, it should be assumed the existing training data is no longer adequate.

- Benchmarking the AI system should be considered, to indicate performance and operational envelope.

AI Sandboxing: Findings (3/3)

- A hazard analysis of the AI system is needed for each potential deployment mode.

- Three key aspects of development of skills and guidance have been identified –

  - o Need to access AI expertise
  - o Operational experience of the AI application
  - o Behaviour and culture, to retain a challenging safety / environmental culture.

- It is important to understand the complexity of the human/system interaction.

- Guidance / Relevant Good Practice (RGP) in this area should initially at least be in the form principles and case studies.

ONR Office for Nuclear Regulation

Environment Agency

Any questions?

Andy White:

andy.white@onr.gov.uk

Jake Surman:

jake.surman@environment-agency.gov.uk