Governance

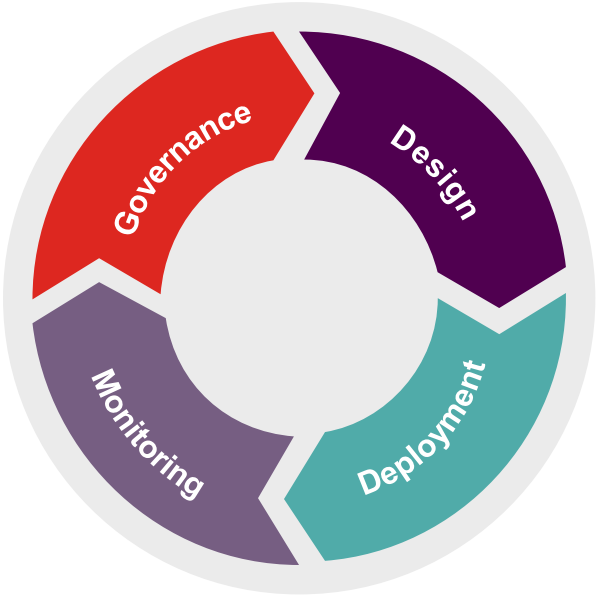
Existing governance, technology management, procurement and data privacy frameworks should be updated to reflect the following considerations relevant to the design, deployment and monitoring of AI technologies:

- ***Documentation*** – nature, technical specification, monitoring regime and system issues to be documented and maintained
- ***Accountability*** – allocation of human responsibility for system performance and compliance across all parts of the AI lifecycle (from both a technical and legal and regulatory perspective)
- ***Model selection*** – framework for exercising judgement in relation to model selection, including trade-offs between model complexity and explainability in the context of the relevant use case
- ***Supply chain*** – appropriateness of and control over outsourcing arrangements and/or use of any off the shelf tools (particularly where not specifically designed for the business purpose)

System Monitoring

System behaviour should be monitored to assess system performance and compliance throughout operation:

- ***System logic information*** – undertaking direct interpretation (where possible) and indirect analysis using explainability methods that are, in each case, appropriate to address the nature of the relevant matter being monitored, model type and/or use case
- ***User interaction*** – to be analysed to monitor integration with existing processes and avoid occurrences of over reliance or undue distrust in the system
- ***Audit and validation*** – system performance against design and deployment objectives should be audited on a regular basis and validated through self-certification or independent audit (as appropriate in the context)



System Design

The model selection, technical specification and build should be designed to reflect requirements for the specific AI technologies in the context of the relevant use case relating to:

- ***Business purpose and technical scope*** – use case and scope of the AI technologies should be clear and understood by internal and external stakeholders
- ***Data inputs*** – model training data should be accurate, up-to-date, complete, conceptually valid and representative
- ***Model opacity*** – to facilitate adequate interrogation of system logic information, particularly input-output relationships and counterfactual explanations (i.e. the conditions under which certain outputs are yielded)
- ***Communication of system logic*** – to be accurate, intuitive and intelligible to internal and external stakeholders (in each relevant context)
- ***Responsiveness*** – to be capable of responding to user requests for information, assistance and rectification (including human-in-the-loop or human-on-the-loop mechanisms where appropriate)
- ***System compliance*** – to align with applicable laws, including those relating to financial regulation, competition, data protection, equality and any specific frameworks relevant to AI technologies
- ***Robustness / safety*** – to facilitate identification of and protection against adversarial attacks through data poisoning
- ***Monitoring*** – to facilitate system monitoring, particularly explainability and auditability

System Performance

The system should be deployed in a manner that reflects the following requirements for the specific AI technologies in the context of the relevant use case:

- ***Operation*** - system should be used in a manner consistent with that for which it was designed
- ***Training*** - user training bespoke to each relevant user capacity to be rolled out on an ongoing basis
- ***System updates and retirement*** – modifications identified as part of monitoring processes should be implemented as and when needed to improve performance and/or prevent performance deterioration (specific consideration should be given as to retraining requirements in relation supervised and/or static machine learning models)

NRF