Data ethics requirements for RM6200 Artificial Intelligence Suppliers

It is important that Suppliers follow the data ethics framework, mitigate bias and ensure diversity in the team that developed/ is developing a solution; as well as transparency/ interpretability and explainability of the results, including audits.

The supplier will need to be open around how the AI service was built.

Below is a list of screening questions that could be added to an invitation to tender and asked of a Supplier where there is an ethical dimension to the tender.

Purpose:

Describe the area of the problem space that is addressed by your solution

Is your approach built on an existing AI system (COTS) or will it be custom made?

Describe what algorithms or techniques you anticipate this service to implement?

Describe the approach to ensuring that use of AI/ML is both necessary and proportionate in the solution

Describe how you have ensured that the data used to power the AI solution is sufficient in quantity, accuracy and relevance to the data available, and what measures have been taken to mitigate bias in the model

Explain how end users have been considered throughout the design and implementation process

Explain how you will demonstrate accountability for the goals and outcomes of the solution

Consent and control:

Provide evidence that you have adopted legally sound and ethical consent for processing and capturing the data throughout the full lifecycle of the solution

Describe the level of human decision-making at critical control points

Privacy and cybersecurity

Describe your privacy and cybersecurity approach for the proposed solution, in particular how the data will be protected

Describe the potential threats and vulnerabilities to the system or solution from external or internal adversaries

Explain your test processes, including the specialist expertise used to assess the solution

Provide, if applicable, evidence of previous case studies of where the solution has been implemented and how the output has been interpreted, highlighting best practice

Describe how your organisation draws on specialist knowledge and expertise to develop and maintain the solution

Explainability

Describe the capabilities in the business to ensure that the outputs of AI technology are explainable and that this explanation is widely available and understandable to a non-expert audience

Would you allow independent, third party audit(s) of the AI solution? If your answer is no, please explain

Ethical considerations relation to data limitations, fairness and bias

What data limitations have you identified and what strategies will you implement to address these data limitations?

(Applicable where solution is COTS and/or government has shared data as part of the procurement process)

How will you ensure that the AI system fits the requirements of the Data Ethics Framework (or can you ensure that you meet the requirements of the Data Ethics Framework during the tool development?)

Describe the approach to eliminate (or minimise) bias, ethical issues, or other safety risks as a result of using the service

Describe the process for ensuring that the development team adopts an ethical mindset

Explain how the solution will be checked to detect bias and the remediation steps if it is introduced

What training data was used, which variables have contributed most to a result, and the types of audit and assurance the model went through, with respect to intrinsic attributes such as considerations of fairness and mitigation of bias. This should be included in documentation supplied by your provider.

Concept drift

Explain how you will ensure the solution or service does not drift from its intended purpose or outcome

Interoperability

Explain how your system or service conforms to specific international or local open interoperability standards or other relevant standards relating to cyber security, coding quality, safety or testing for example

Due diligence on existing algorithms or COTS solutions:

Describe the architecture of the solution, including use of external COTS or open source components and the function they provide in the solution. This should consider the data used by each component of the solution and how the output of that component was validated

Documentation on toolkit and auditability

Please provide information about:

1. the available toolkit, including the list of software tools the provider proposes to use
2. the origin and nature of any data the provider plans on bringing to the project
3. data used to train algorithms the provider will bring to the project
4. and algorithms used
5. documentation that provides information about the algorithms used (e.g. data used for training algorithms, whether the model is based on supervised, unsupervised, or reinforcement learning, and any limitations).

provide information on their model building methodology, including how they select variables, build samples (where applicable), and validate the model.

Documentation that provides information about the algorithms used (e.g. data used for training algorithms, whether the model is based on supervised, unsupervised, or reinforcement learning, and any limitations).

provide information on their model building methodology, including how they select variables, build samples (where applicable), and validate the model.

What training data was used, which variables have contributed most to a result, and the types of audit and assurance the model went through, with respect to intrinsic attributes such as considerations of fairness and mitigation of bias. This should be included in documentation supplied by your provider.

Enable end-to-end auditability by implementing process logs that gather the data across the modelling, training, testing, verifying, and implementation phases of the project lifecycle. Such a log should allow for the variable accessibility and presentation of information with different users in mind to achieve interpretable and justifiable AI.

Lifecycle management

How do you envisage training and knowledge transfer of the AI system development and deployment to the public sector delivery team?

Explain how you will ensure usability for non-trained staff

Explain how the AI system will be maintained, how it’s accuracy and integrity will be sustained over time, and whether third party providers could be engaged for these activities

Monitoring and feedback loops: For AI-powered solutions in the public sector, implementation plans, sustainable and ongoing evaluation methods, and mechanisms to feed back into the data model are crucial to ensure ethical use.

Make clear in your invitation to tender that such considerations by the provider count and will be discussed during procurement

Testing: establish with the provider how the efficacy of the model will be monitored once deployed

Skills

Can you demonstrate how you will assess the skills, qualifications and diversity of the team that will develop and deploy the AI system?