The Institute for Ethical AI & Machine Learning

AI-RFX Procurement Framework v1.0

Machine Learning Maturity Model

Index

0 - Introduction

0.1 - Overview

This “**Machine Learning Maturity Model v1.0**” is part of the AI-RFX Procurement Framework, and it is the core of all the templates including the **”AI Request for Proposal Template”** & the **“AI Tender Competition Template”**. The web version is also available for reference.

The Machine Learning Maturity Model is an extension of The Principles for Responsible Machine Learning, which aims to convert the high level Responsible ML Principles into a practical checklist-style assessment criteria. This “checklist” goes beyond the machine learning algorithms themselves, and provides an assessment criteria to evaluate the maturity of the **infrastructure** and **processes** around the algorithms. The concept of **“Maturity”** is not just defined as a matter of technical excellence, scientific rigor, and robust products. It also essentially involves responsible innovation and development processes, with sensitivity to the relevant domains of expert knowledge and consideration of all relevant direct and indirect stakeholders.

The Machine Learning Maturity Model **should be a subset** of the overall assessment criteria required to evaluate a proposed solution, and it is specific to the machine learning piece. It should be complemented with a traditional assessment of other areas such as the specific features requested, services needed, and more domain-specific areas.

Each of the criteria was designed to be linked to each one of the Principles for Responsible Machine Learning, and consists of the following:

#	Assessment Criteria	Responsible ML Principle
#1	Practical benchmarks	Principle #6: Practical accuracy
#2	Explainability by justification	Principle #3: Explainability by justification
#3	Infrastructure for reproducible operations	Principle #4: Reproducible operations
#4	Data and model assessment processes	Principle #2: Bias Evaluation
#5	Privacy enforcing infrastructure	Principle #7: Trust by privacy
#6	Operational process design	Principle #1: Human Augmentation
#7	Change management capabilities	Principle #5: Displacement strategy
#8	Security risk processes	Principle #8: Security risks

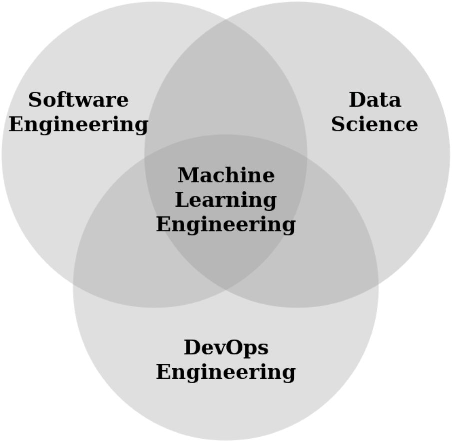
0.2 - About Us

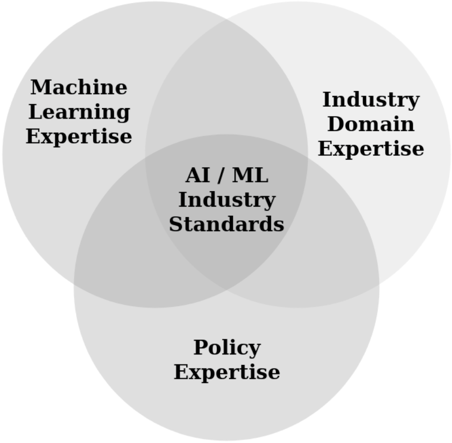
The Institute for Ethical AI & Machine Learning is a UK-based research centre that carries out world class research into responsible machine learning systems. We are formed by cross functional teams of applied STEM researchers, philosophers, industry experts, data scientists and software engineers.

Our vision is to mitigate risks of AI and unlock its full potential through frameworks that ensure ethical and conscientious development of intelligent systems across industrial sectors. We are building the Bell Labs of the 21st Century by delivering breakthrough contributions through applied AI research. You can find more information about us at http://ethical.institute.

0.3 - Motivation

There is currently a growing number of companies that are working towards introducing machine learning systems to automate critical processes at scale. This has required the "productisation" of machine learning models, which introduces new complexities. This complexity revolves around a new set of roles that fall under the umbrella of "Machine Learning Engineering". This new set of roles fall in the intersection between DevOps, data science and software engineering.




Machine Learning Expertise
Industry Domain Expertise
AI / ML Industry Standards
Policy Expertise

To make things harder, the deployment of machine learning solutions in industry introduces an even bigger complexity. This involves the intersection of the new abstract "Machine Learning Engineering" roles, together with the industry domain experts and policy makers.

Because of this, there is a strong need to set the AI & ML standards, so practitioners are empowered to raise the bar for safety, quality and performance around AI solutions. **The AI-RFX Procurement Framework aims to achieve the first steps towards this.**

0.4 - How to use this document

0.4.1 - Using as reference

Many procurement managers may already own internally-approved assessment criteria. If that is the case, this document can be treated as a reference to obtain insights on key areas that should be taken into consideration when procuring and evaluation an AI / Machine Learning solution.

0.4.2 - Structure

Each subsection below consists of a **detailed explanation** of the criteria. It is followed by an **summary overview** of the requirements expected by the suppliers. Finally it contains a set of **detailed questions** that the supplier is expected to answer whether explicitly or implicitly in their proposal, together with **red flags** to look out for in each of the detailed questions.

0.4.3 - Example

The Machine Learning Maturity Model was used to build the **"AI Request for Proposal Template"** & the **"AI Tender Competition Template"**, which are part of the AI-RFX Procurement Framework.

0.4.4 - When to use

This template is relevant only for the procurement of machine learning systems, and hence it is only suitable when looking to automate a process that involves data analysis that is too complex to be tackled using simple RPA tools or rule-based systems.

0.5 - Template vs Reality

This document should serve as a guide, and doesn't require everything to be completed exactly as it's stated. Especially for smaller projects, the level of detail required may vary significantly, and some sections can be left out as required. This template attempts to to provide a high level overview on each chapter (and respective sections) so the procurement manager and suppliers can provide as much content as reasonable.

0.6 - Open Source License - Free as in freedom

0.6.1 - Open source License

This document is open source, which means that it can be updated by the community. The motivation to release this as open source is so that it is continuously improved by the community. This will ensure that the standards for safety, quality and performance of what is expected in machine learning systems will keep increasing, whilst being kept in check on a realistic level by both suppliers and companies.

0.6.2 - Contributing.md

The Institute for Ethical AI & Machine Learning's **AI-RFX committee** is in charge of the contributing community for all of the templates under the **AI-RFX Procurement Framework**. Anyone who would like to contribute, add suggestions, or provide example and practical uses of this template, please contact us through the website, or send us an email via a@ethical.institute.

0.6.3 - License

This document is registered under this MIT License (raw file), which means that anyone can re-use, modify or enhance this document as long as credit is given to The Institute for Ethical AI & Machine Learning. It also includes an "as is" disclaimer. **Please read the license before using this template**.

Machine Learning Maturity Model

1 - Practical benchmarks

This **Machine Learning Maturity Model** assessment criteria is directly aligned with the Responsible Machine Learning Principle #6 - Practical accuracy.

Explanation

- Having the right benchmark metrics is one of the most important points to consider during the evaluation of machine learning solutions. Relevant benchmarks that are considered in this section include **accuracy, time, time-to-accuracy, and computational resources.**
- The criteria of what makes good benchmarks can vary significantly depending on the task complexity, dataset size, etc. However the objective of this criteria is to assess that suppliers are able to follow best practices in data science, and make sure these are aligned with the use-case requirements.

Requirements

- Suppliers must be able to demonstrate best practices in software development, data science and industry-specific knowledge when presenting benchmarks. These benchmarks include:
  - **Time** - Supplier must provide estimated processing times
  - **Accuracy** - Supplier must provide metrics beyond accuracy as relevant
  - **Time-to-accuracy** - Supplier must provide information on the estimate time and resources it takes to train new models to a reasonable accuracy
  - **Computational resources** - Supplier must provide insight on computational resources required for efficient use of their system

#	Question	Red flags
1.1	Does the supplier have a process and/or infrastructure to make available statistical metrics beyond accuracy?	● Supplier doesn't have a process and/or infrastructure to provide statistical metrics beyond simple accuracy (e.g. true positive rate, false positive rate, precision, etc). ● Supplier doesn't provide reasonable insights (i.e. confusion matrix, learning curves, error bars, etc)
1.2	Does the supplier have a process to ensure their machine learning evaluation metrics (i.e.	● Supplier doesn't have a process to ensure that the cost functions they selected reflect the objectives of the use-case

	cost functions & benchmarks) are aligned to the objective of the use-case?	
1.3	Does the supplier have a process to validate the way they evaluate predictions as correct or incorrect?	● Supplier doesn't have a process to ensure the methods / function(s) they use to evaluate a prediction as correct or incorrect is aligned to the way the relevant domain expert would.
1.4	Does the supplier use reasonable statistical methods when comparing performance of different models?	● Supplier does not use standard comparison methods such as t-tests, ROC curves, or relevant metrics when comparing different solutions proposed.
1.5	Does the supplier provide comprehensible information on the time performance of their solution?	● Supplier doesn't provide reasonable time benchmarks for tasks, and how the time behaves as other variables change (data instance size, batch volumes, etc)
1.6	Does the supplier provide comprehensible estimates on time and resources required to develop a model from scratch to a reasonable accuracy?	● Supplier doesn't have reasonable estimates for time/resources required to build new models/capability that is of a reasonable or required accuracy
1.7	Does the supplier provide minimum and recommended system requirements?	● Supplier doesn't have reasonable guidance on minimum system requirements to operate platform efficiently in regards to number of cores, ram required based on load, storage, etc. unless not relevant (e.g. hosted in external cloud)
1.8	Does the supplier provide comprehensible documentation around their benchmarks?	● Supplier doesn't have a reasonable level of documentation provided with information about performance metrics
1.9	Does the supplier ensure staff in the benchmark processes have the right exp.?	● Supplier is not able to show the staff involved in the setting the benchmarks have a reasonable level of statistics, and that the domain experts are involved for decisions where reasonable

2 - Explainability by justification

This **Machine Learning Maturity Model** assessment criteria is directly aligned with the Responsible Machine Learning Principle #3 - Explainability by justification.

Explanation

- When domain experts are asked how they came to a specific conclusion, they don't answer by pointing to the neurons that fired in their brains. Instead domain experts provide a **"justifiable"** explanation of how they came to that conclusion.
- Similarly, with a machine learning model the objective is not to demand an explanation for every single weight in the algorithm. Instead, we look for a justifiable level of reasoning on the end-to-end process around and within the algorithm.
- The level of scrutiny for an explanation to be "justifiable" will most certainly vary depending on the critical nature of the use-case, as well as the level of feedback that can be analysed by humans.

Requirements

- This criteria is heavily dependent on Criteria 1 - Practical benchmarks, as suppliers have the right processes and capabilities around their accuracy metrics.
- Suppliers must make a reasonable case about how their solution (or solution + human) will be able to provide at least the same level (or higher) of justification when making a final decision on an instance of data analysis as a domain expert would.
- In order for suppliers to propose at least the same level of justification, they must also provide the current level of justification as a benchmark, from a quantitative perspective.

#	Question	Red flags
2.1	Does the supplier provide audit trails to assess the data that went through the models?	● Supplier doesn't have capabilities to provide human-readable audit trails where reasonable
2.2	Does the supplier have a process and/or infrastructure to explain input/feature importance?	● Supplier doesn't have a process and/or infrastructure in place to assess how inputs/features interact to result in specific predictions
2.3	Does the supplier provide capabilities to	● Supplier doesn't provide ways to explain how inputs/features result in the inference outcomes

	explain how input/features affect results?	where justification is required (e.g. when there’s a lack of human review, or critical nature of a use-case)
2.4	Does the supplier have the process and/or infrastructure to use model explainability techniques when developing deep learning / more complex models?	● Supplier doesn’t have processes and/or infrastructure to use explainability techniques (such as SHAP, LIME, aLIME, etc) to increase explainability of models where required
2.5	Does the supplier have process and/or infrastructure to work with domain experts to abstract their knowledge into models?	● Supplier doesn’t have a process and/or infrastructure to work with relevant domain experts and convert key knowledge into inputs/features that can introduce more levels of explainability to the machine learning process where reasonable
2.6	Does the supplier provide comprehensible information around their explainability processes?	● Supplier doesn’t have a reasonable level of documentation provided with information about the processes they involve around explainability
2.7	Does the supplier ensure the staff involved in the explainability processes have the right experience?	● Supplier is not able to show the staff involved in the analysis of machine learning models have a reasonable understanding of machine learning ● Supplier is not able to show the processes ensure they involve domain experts where reasonable

3 - Data and model assessment processes

This **Machine Learning Maturity Model** assessment criteria is directly aligned with the Responsible Machine Learning Principle #2: Bias evaluation.

Explanation

- Any any non-trivial decisions (defined as having more than 1 option) always carry an inherent bias without exception
- Hence the objective is not to remove bias from a machine learning completely. Instead, the objective is to ensure that the "desired bias" is aligned with our accuracy/objectives, and "undesired bias" is identified and mitigated.
- To be more specific, bias in machine learning boils down to the error between development and production. As a result of this, all machine learning models start to “degrade” as soon as they are put in production. The reasons for this include:
    - Unseen data is not representative to the data used in development
    - Temporal data changes as time goes on (e.g. inflation affects price)
    - Human-generated data changes as people and projects change
- Bias in machine learning is a challenge that can be tackled by ensuring there are processes in place to identify, document and mitigate bias

Requirements

- This assessment criteria is heavily dependent on Criteria 1 - Explainability by justification, and Criteria 2 - Practical benchmarks being in place.
- Suppliers must be able to demonstrate processes and infrastructure they have to identify undesired bias through best practices in data science as well as awareness of domain-specific considerations

The Institute for Ethical AI & Machine Learning is working with the IEEE p7003 working group to develop the **p7003 Algorithmic Bias Considerations standard** that will facilitate this assessment criteria once it is released as suppliers that obtain this certification will verify that they have the relevant process towards data and model assessment.

#	Question	Red flags
3.1	Does the supplier have a process to assess representability of datasets?	● Supplier doesn’t have a process in place to assess representability of training data
3.2	Does the supplier have a process to identify and document undesired	● No process in place to analyse input/feature importance during the development of a model ● No process in place to obtain a breakdown of

	biases during the development of their models?	accuracy metrics on an input/feature level to identify undesired bias where reasonable ● No process in place to identify wanted/unwanted correlations within the input/features where reasonable
3.3	Does the supplier have capabilities to track performance metrics in production to identify and mitigate new bias?	● No process and/or infrastructure in place to identify metrics that should be tracked in production to alert when a model drops under certain thresholds where reasonable ● If metrics are tracked, there is no explicit awareness of why they need to be tracked where required or where not obvious
3.4	Does the supplier provide comprehensible information around their data and model evaluation processes?	● Supplier doesn’t have a reasonable level of documentation provided with information about the processes they involve around explainability
3.5	Does the supplier demonstrate the team they have allocated has the right expertise to perform the data and model assessment efficiently?	● Supplier is not able to show the staff involved in the analysis of machine learning models have strong background on statistics and/or machine learning ● Supplier is not able to show the processes ensure they involve domain experts where reasonable

4 - Infrastructure for reproducible operations

This **Machine Learning Maturity Model** assessment criteria is directly aligned with the Responsible Machine Learning Principle #4 - Reproducible operations.

Explanation

- Similar to production software, machine learning requires infrastructure to ensure reliable and robust service offerings
- Different to traditional software however, machine learning introduces complexities beyond the code, such as versioning and orchestration of models
- This requirements demand the suppliers to be conscious of this, and ensure their infrastructure is able to cope with these challenges

Requirements

- Suppliers must also be able to demonstrate their capabilities to version, roll-back, diagnose and/or deploy models to production
- Suppliers must have the processes and/or infrastructure to be able to separate the development of new models (i.e. new capabilities) from the serving in production of the models
- Suppliers must demonstrate the ability to scale their services as required by the use-case.

#	Question	Red flags
4.1	Does the supplier have process and/or infrastructure to version models?	● No infrastructure and/or processes to version different machine learning models where reasonable
4.2	Does the supplier have process/infrastructure to re-train previous version of models?	● No infrastructure and/or processes to re-train previous versions of models where reasonable
4.3	Does the supplier have a clear separation for their workflows around development and production of models?	● No clear separation for the workflows of development and production/orchestration of models
4.4	Does the supplier have a	● No capabilities to separate the production and

	QA/BETA process within their machine learning lifecycle?	QA/BETA process for testing where reasonable
4.5	Does the supplier have a reasonable process to deploy and revert back models in production?	● No reasonable process and/or infrastructure to move models from development to production ● No reasonable process and/or infrastructure to revert models in production without an unreasonable level of disruption
4.6	Does the supplier have the infrastructure to reproduce and diagnose errors observed in production efficiently?	● No process and/or infrastructure to diagnose errors in production by reproducing executions where reasonable
4.7	Does the supplier have the capabilities to scale their computation horizontally?	● No reasonable process and/or infrastructure to increase the number of horizontal workers to process larger loads where required
4.8	Does the supplier have the capabilities to scale their computation vertically?	● No reasonable process and/or infrastructure to handle computational loads that need to be broken up across multiple nodes for processing (e.g. due to not fitting in memory, or the task requiring too much load for one instance)
4.9	Does the supplier have the capabilities to provision the relevant resources required by for the computation of the models across their infrastructure?	● No reasonable process and/or infrastructure to manage the resources across their server infrastructure efficiently as required
4.10	Does the supplier have a stable release cycle and method to provide updates to machine learning infrastructure?	● No process and/or infrastructure to develop updates
4.11	Does the supplier provide the required functionality to extend the features provided?	● Supplier doesn't provide processes or infrastructure to extend functionality of the solution through APIs, SDKs or relevant interfaces as required.

4.12	Does the supplier provide comprehensible documentation on the infrastructure around their machine learning?	● Supplier doesn’t have a reasonable level of documentation provided with information about the infrastructure they provide, which may include: ○ Deployment instructions ○ Minimum requirements overview ○ User manual ○ Technical documentation
4.13	Does the supplier demonstrate the team they have allocated for the design, development and delivery of the machine learning infrastructure have the right expertise?	● Supplier is not able to show the staff involved in the development and/or delivery of the machine learning solution have strong background in software development, machine learning and sysadmin/devops engineering. ● Supplier outsources large part of their core software offerings without a sensible reason.

5 - Privacy enforcing infrastructure

This **Machine Learning Maturity Model** assessment criteria is directly aligned with the Responsible Machine Learning Principle #7 - Trust beyond the user.

Explanation

- Both suppliers and companies have a responsibility to protect the user's privacy when their data is collected and stored into the solution.
- Different to the B2C world where there might be a clear user, in B2B solutions there may be a large number of different user types that interact with the solution (directly and indirectly), increasing the complexity of data protection.
- Whilst section "8 - Security risks processes" focuses on addressing data security risks from **"external"** threats, this section focuses on mitigating privacy violations that can arise from **"internal"** parties (whether intentional or unintentional)

Requirements

- Suppliers must demonstrate awareness and capability to protect privacy of users in the platform across multiple levels.
- The right communication channels should be established to ensure that all the stakeholders from the supplier are capable of separating and protecting personally identifiable information where reasonable.
- When anonymization techniques are used, suppliers must demonstrate the use of reasonable techniques that can mitigate re-identification attacks.

#	Question	Red flags
5.1	Does the supplier have the process and/or infrastructure to restrict access to user data?	● Supplier doesn't have the processes and/or infrastructure that ensures only users with explicitly granted permissions have access to user data
5.2	Does the supplier have processes and/or infrastructure in place to ensure data is anonymised where reasonable?	● Supplier doesn't have the processes and/or infrastructure to ensure user data is anonymized where reasonable (e.g. through automated differential privacy/anonimisation or manual/batch anonimisation) ● Supplier uses methods of anonymization that are not reasonable for the confidentiality level of the data presented
5.3	Does the supplier have	● No capability and/or infrastructure to show a

	capabilities to ensure privacy protection on the data that is used on models?	reasonable level of confidentiality for personal data that is used to train models
5.4	Does the supplier have the internal capabilities to ensure compliance with the relevant regulations?	● No process and/or infrastructure to deal with relevant privacy regulations that require handling of personal data (such as GDPR)
5.5	Does the supplier have the capability to ensure privacy protection on the metadata contained within their models?	● No infrastructure and/or process to ensure a reasonable level of privacy and protection to personally identifiable information both as raw data, and within a trained machine learning model ● When using simpler / more explainable models, no process or capability to ensure protection of data in machine learning models
5.6	Does the supplier have processes in place that ensure user privacy measures are in place based on the consent that has been given?	● Supplier doesn't have processes to ensure a project is evaluated from user privacy implications perspective
5.7	Does the supplier provide comprehensible documentation on the infrastructure around their machine learning?	● Supplier doesn't provide comprehensible documentation and information about personal storage methods, how personal data will be used, and how it will be protected where reasonable
5.8	Does the supplier demonstrate the team they have allocated has the right expertise to ensure the required level of privacy is provided to the users in the system?	● Supplier is not able to show the staff involved in the development and/or delivery of the machine learning solution have strong background in software development, machine learning and relevant domain expertise to make relevant decisions on personal data (e.g. legal, ethical, industry expertise)

6 - Operational process design

This **Machine Learning Maturity Model** assessment criteria is directly aligned with the Responsible Machine Learning Principle #1: Human augmentation.

Explanation

- This criteria focuses on the processes required for the responsible operation of the solution. It assesses whether the proposed process takes into consideration the fail-safe steps in place to mitigate the impact of errors / incorrect predictions.
- The operational process involves the full end-to-end steps that are performed **around** the machine learning system, including human intervention or analysis at any relevant step of the process

Requirements

- This criteria is heavily dependent on Criteria 3 - Data and model assessment process, as suppliers have the plan in place in order for the operational/business/digital transformation to take place.
- The main objective of this criteria is to ensure suppliers show explicitly they have considered the impact of incorrect predictions or errors, and are able to mitigate some of these through the operational steps introduced beyond the technology.

#	Question	Red flags
6.1	Does the supplier have a process to assess the need for a human-in-the-loop design process based on the impact of incorrect predictions?	● Supplier doesn't have a process to assess the need of a human review process based on the impact of incorrect predictions or errors in proposed process
6.2	Does the supplier have a process to assess whether a human-in-the-loop review process is or isn't necessary?	● Supplier doesn't have a reasonable process of why a human-in-the-loop review process is or isn't necessary ● Supplier's process doesn't use key points from Criteria 1, Criteria 2, and Criteria 3 in their process
6.3	Does the supplier have a process to assess whether a temporary human-in-the-loop	● Supplier doesn't have a reasonable process to assess whether a temporary human-in-the-loop review process is or isn't necessary to monitor newly deployed models

	process is necessary?	● Supplier's process doesn't use key points from Criteria 1, Criteria 2, and Criteria 3 in their process
6.4	Does the supplier have a process to assess whether a scheduled human review is or isn't necessary after the deployment of a model?	● Supplier doesn't have a reasonable process to assess whether a scheduled (e.g. weekly, monthly, annual, etc) human review is or isn't necessary after the deployment of a model ● Supplier's process doesn't use key points from Criteria 1, Criteria 2, and Criteria 3 in their process
6.5	Is the supplier able to ensure the right domain experts are involved in the human in the loop review process?	● Supplier doesn't have a process to ensure that stakeholders with the right skill-set or capabilities are in place as required
6.6	Does the supplier demonstrate the team they have allocated has the right expertise to ensure safety, quality and performance in the design of the operational process?	● Supplier is not able to show the staff involved in the development and/or delivery of the machine learning solution have strong background in software development, machine learning and relevant domain expertise to make relevant decisions on the impact of incorrect predictions and add human review processes where reasonable (e.g. legal, ethical, industry expertise)

7 - Change management capabilities

This **Machine Learning Maturity Model** assessment criteria is directly aligned with the Responsible Machine Learning Principle #5 - Displacement strategy.

Explanation

- When performing any large-scale IT projects it is important to have the right change management plans in place. With AI & machine learning systems it is no different.
- Large-scale machine learning systems are often introduced to create a new data analysis capability and/or automate an existing process. This results in people that will need to be re-trained to use the systems in the right way, as well as ensuring that the gains in efficiency are distributed by allocating the now available time accordingly.

Requirements

- This criteria is heavily dependent on Criteria 6 - Operational Process Design, as suppliers have the plan in place in order for the operational/business/digital transformation to take place.
- For larger system rollouts, suppliers must be able to demonstrate the right processes and infrastructure to deal with the impact of their proposed automation, in regards to the stakeholders that are being partially or fully automated.
- Suppliers must also be able to show they have the processes and infrastructure to perform training and handover for relevant stakeholders that will be operating the solution

#	Question	Red flags
7.1	Does the supplier provide a comprehensible plan to roll out the solution internally?	● Supplier doesn't provide a comprehensible plan to roll out solution internally when large number of stakeholders will be affected
7.2	Does the supplier have capability to carry out the business change plan proposed?	● Supplier doesn't demonstrate it has the relevant expertise or man-power to carry out required operational transformation piece as required for the project
7.3	Does the supplier have the capability to provide	● Supplier doesn't have a reasonable capacity, process or infrastructure for training key

	training for stakeholders that will operate the solution?	stakeholders and build the relevant internal capabilities required.
7.4	Does the supplier provide a comprehensible plan for handover?	● Supplier doesn’t provide a plan for handover of the technology, including the ability for the company to have the right skills in-house where relevant.
7.5	Does the supplier have a process and/or capability to support the transition of stakeholders that are automated?	● Supplier doesn’t show awareness or provide a plan for the impact the solution will have on the automation for the business, or a proposed plan for transitioning the time freed once the solution is in place
5.7	Does the supplier provide comprehensible documentation and material on change management plans and training material?	● Supplier doesn’t provide comprehensible documentation and material around the change management plans, as well as the training and handover material for education and reference
5.8	Does the supplier demonstrate the team they have allocated has the right expertise to perform the change management delivery, trainings and handover?	● Supplier is not able to show the staff involved in the delivery of the change management piece of the machine learning solution have strong background in the relevant areas required for the delivery, training and handover.

8 - Security risk processes

This **Machine Learning Maturity Model** assessment criteria is directly aligned with the Responsible Machine Learning Principle #8 - Data risk awareness.

Explanation

- This section focuses on the “external” threats & risks around the data contained in the systems proposed by the solution.
- This section doesn’t only require suppliers to have the right technical safeguards, but also to have the right education for relevant users that will interact with potentially complex systems
- This section should be complementary to the standard security questions that are often provided through questionnaires in tender processes.

Requirements

- Suppliers must demonstrate awareness around data and system security, including around their machine learning models, as well as the infrastructure around it.
- Suppliers should provide an overview of how their systems as well as processes are secured.
- Suppliers should demonstrate they have the right processes in place around stakeholders that operate the solution
- Suppliers should demonstrate they have the right processes in place to educate the stakeholders that operate the solution

#	Question	Red flags
8.1	Does the supplier have processes to ensure only privileged users have access	● No internal capabilities or infrastructure in place for supplier to identify security risks in machine learning infrastructure
8.2	Does the supplier ensure all machine learning model data is encrypted at transport?	● No process and/or infrastructure to ensure machine learning data encrypted on transport
8.3	Does the supplier ensure all machine learning model data is encrypted at rest?	● No process and/or infrastructure to ensure machine learning data is encrypted at rest
8.4	Does the supplier have	● Supplier doesn’t provide comprehensible process

	processes and/or infrastructure in place to ensure that all the encryption keys and relevant passwords are not shared or repeated across deployments?	and/or infrastructure to ensure that keys and passwords are not repeated
8.5	Does the supplier have processes and/or infrastructure to assess the level of protection require based on exposure?	● Supplier doesn't have a process to analyse the level of security measures for machine learning models that are exposed, such as protection from adversarial attacks, etc.
8.6	Does the supplier have a process to ensure confidential information is not exposed through logs or other mediums of metrics?	● Supplier doesn't provide a comprehensible process that ensures a new update of their system exposes confidential information through logs or other mediums of metrics
8.7	Does the supplier have processes and/or infrastructure to ensure system access is restricted to privileged users as required?	● Supplier doesn't provide comprehensible processes and/or infrastructure to ensure system access is restricted
8.8	Does the supplier demonstrate the team they have allocated has the right expertise to fortify their machine learning infrastructure against external threats?	● Supplier is not able to show the staff involved in the development of the core product or delivery of key solution has strong skills on the security requirements based on the critical nature and scale or the project