|  |  |
| --- | --- |
| Name: **Data Readiness Questionnaire** Use Case:  Date: | |
| What data will be provided by your agency to support the planned capability? | |
| How will the data be used?   * Training data * Ground truth data * Validation or test data * Target data * I don’t know * Other   *[if data will be used for training or ground truth or target data be sure the vendors can keep the data private and secure, and that the data is not re-used in the AI system.]*  *[if the provided data is used for training data, is the data representative? Is it balanced? Does it contain all target classes/topic/demographics the system will classify or predict? What are the known biases in the data? - you may need to do a data audit and balance the data]* | What is the status of ground truth?   * Ground truth data exists * Could transform data into ground truth * Using pre-trained model * None exists * I don’t know   *[if you are not providing ground truth or training data, you will be using a pre-trained model or will have to create ground truth to validate performance or data to train or tune models. In many cases, pre trained models will need additional data to perform well. Both scenarios involve additional costs and risks.*  *For some high-value tasks the method of creating training data may need to be audited. If synthetic data will be used for training or ground truth data, document the method the data was created and how data quality and representation is assured?]* |
| Who is depicted in the data? What privacy or disclosure responsibilities do you have related to the people depicted in the data? | Do you want to or need to restrict some future use of data input or generated from this process?  Y/ N / IDK  If so, how? |
| Who owns the data? Is the data protected by privacy, IP, or other rights?  *[Review with your General Counsel data rights and permissions before proceeding.]* | How will the resulting/output data be managed over time? |
| How will you use the resulting/output data? | How will the resulting/output data be monitored over time? |
| Data Quality  **Will you have documented processes in place to test datasets against biases and other unexpected outcomes? This could include experience in applying frameworks, methods, guidelines or other assessment tools.**  Yes  No  Will you be developing a process to document how data quality issues were resolved during the design process?  Yes  No  Will you be making this information publicly available?  Yes  No  Will you undertake a Gender Based Analysis Plus of the data?  Yes  No  Will you be making this information publicly available?  Yes  No  Have you assigned accountability in your institution for the design, development, maintenance, and improvement of the system?  Yes  No  Will you have a documented process to manage the risk that outdated or unreliable data is used to make an automated decision?  Yes  No  Will you be making this information publicly available?  Yes  No  Will the data used for this system be posted on the Open Government Portal?  Yes  No  **Procedural Fairness**  Will the audit trail identify the authority or delegated authority identified in legislation?  Yes  No  Will the system provide an audit trail that records all the recommendations or decisions made by the system?  Yes  No  Will all key decision points be identifiable in the audit trail?  Yes  No  Will all key decision points within the automated system's logic be linked to the relevant legislation, policy or procedure?  Yes  No  Will you maintain a log detailing all of the changes made to the model and the system?  Yes  No  Will the audit trail clearly set out all decision points made by the system?  Yes  No  Could the audit trail generated by the system be used to help generate a notification of the decision (including a statement of reasons or other notification) where required?  Yes  No  Will the audit trail identify precisely which version of the system was used for each decision it supports?  Yes  No  Will the audit trail show who the authorized decision-maker is?  Yes  No  Will the system be able to produce reasons for its decisions or recommendations when required?  Yes  No  Will there be a process in place to grant, monitor, and revoke access permission to the system?  Yes  No  Will there be a mechanism to capture feedback by users of the system?  Yes  No  Will there be a recourse process planned or established for clients that wish to challenge the decision?  Yes  No  Will the system enable human override of system decisions?  Yes  No  Will there be a process in place to log the instances when overrides were performed?  Yes  No  Will the audit trail include change control processes to record modifications to the system's operation or performance?  Yes  No  Will you be preparing a concept case to the Government of Canada Enterprise Architecture Review Board?  Yes  No  **PRIVACY**  Privacy  If your system uses or creates personal information, will you undertake or have you undertaken a Privacy Impact Assessment, or updated an existing one?  Yes  No  Will you design and build security and privacy into your systems from the concept stage of the project?  Yes  No  Will information be used within a closed system (i.e. no connections to the Internet, Intranet or any other system)?  Yes  No  If the sharing of personal information is involved, has an agreement or arrangement with appropriate safeguards been established?  Yes  No  Will you de-identify any personal information used or created by the system at any point in the lifecycle?  Yes  No | How do we do this? Against what do we measure?  Unpack this more-  Look this up  Good question, relates to implementation  Data processing plan question  Or any shared repository?  Good question, in the model card documentation?  Same as above  Maybe these questions are only applicable to decision-making enhancements? |