**Algorithm Impact Assessment (AIA) Questionnaire**

**Instructions:**

* Please complete this AIA Questionnaire when directed by the decision maker who reviewed the associated [**Algorithm Threshold Assessment**](https://data.govt.nz/assets/data-ethics/algorithm/Algorithm-threshold-assessment-questionnaire.docx)**.**
* Please read alongside the [**AIA User Guide**](https://data.govt.nz/docs/algorithm-impact-assessment-user-guide), whichprovides guidance and support to help you include the information required to enable **informed decision-making** on the impacts of using the proposed algorithm. See the [***Glossary***](https://data.govt.nz/docs/algorithm-impact-assessment-user-guide/#glossary) in the **User Guide** for the defined terms used in the AIA documentation.
* AIAs should be performed with **diverse** and **multi-disciplinary input**, initiated at an **early stage of the project** and **updated** as required across the Project. While the questions are future focused, please also provide the requested information for algorithms already in use.
* Please use **clear, plain language** to providesuccinct but comprehensive answers that avoid the use of jargon and technical terms. Although your Project may involve more than one algorithm, this AIA refers to an algorithm in the **singular** for simplicity and consistency.
* Please attach/link to relevant **supplementary** **information** about the algorithm *(for example, business case, specifications, general project documentation)*
* Use the answers in the AIA Questionnaire to produce an [**AIA Report**](https://data.govt.nz/assets/data-ethics/algorithm/Algorithm-impact-assessment-report-template.docx) and send both documents to the reviewer/decision maker identified below.
* Please answer the following questions and then send the document to the Reviewer/decision maker listed below.

|  |  |  |  |
| --- | --- | --- | --- |
| 1. AIA details | | | |
| **1.1** | **Algorithm project name** | (the “**Project**”). | |
| **1.2** | **Personnel details:** | ***Specify name, role, email address and date completed/reviewed:*** | |
| * **Author** *(person completing this form)* |  | |
| * **Business Owner** |  | |
| * **Executive Sponsor** *(if any; for example, DCE)* |  | |
| * **Reviewer/decision maker** |  | |
| 1. Project information | | | |
| **2.1** Describe the **issue** the Project is aiming to solve or the **goal** it’s trying to achieve. | | |  |
| **2.2** How does the algorithm help **achieve the objectives** of the Project (the “**Purpose**”)? | | |  |
| **2.3** Describe the **source** of the **proposed algorithm**, how it **works** andwill be **used**, the nature of its **outputs** and **why** it was selected. How will using the algorithm be different from the **status quo.** How is it a better **alternative** andis it the **best solution** for the issue? | | |  |
| **2.4** Describe the **key opportunities and benefits** the algorithm is intended to provide and its **core value proposition**. **Who** will benefit and why? What are the **public benefits**? | | |  |
| **2.5 Who** will be **impacted** by the algorithm, **directly** and **indirectly** (“**Impacted People**”)? | | |  |
| **2.6** Describe the **users** of the algorithm, including **how** and **when** they will use it. | | |  |
| **2.7** If the algorithm will be used within an **AI model** or **system** (**tool**) - including Generative AI (*see* [*User Guide glossary*](https://data.govt.nz/docs/algorithm-impact-assessment-user-guide/#glossary)*) -* please describe the nature of that tool, where it comes from and how it will be used. Will any new data or content be **created** and, if so, how will it be used? Is it the best AI tool for the Purpose? | | |  |
| **2.8** Describe any **hardware/software systems** needed for the algorithm to function. | | |  |
| **2.9** Will the algorithmbe **developed internally** or **obtained** from an **external supplier**? | | |  |
| **2.10** Indicate the relevant **dates** for each stage of the **algorithm’s lifecycle**, including for each of planning, design, development, testing, deployment, review and retirement. | | | ***Expected dates:*** |
| **2.11** Describe the **team** handling data and developing the algorithmand the extent of **diversity** in that group. What **perspectives** might be **missing** and how will this be addressed? How is a **Te Tiriti** perspective incorporated? | | |  |

|  |  |
| --- | --- |
| 1. Overall risk profile (potential best and worst-case scenarios) | |
| **3.1** Describe the **best-case scenario(s)** that could realistically arise from using the algorithm. Who benefits, how and why? How **likely** is this? |  |
| **3.2** Describe the **worst-case scenario(s)** that could realistically arise from using this algorithm. How **likely** is this? What is the possible **media headline** and who is likely to be held **accountable**? How would the problems be **addressed** and **by whom**? |  |
| **3.3** What are the **potential harms** arising from the worst-case scenario and the related risks to the agency? How **likely** and what is the severity of those harms and risks? **Who** is most at risk of being harmed and **why**? |  |

|  |  |
| --- | --- |
| 1. Governance and human oversight | |
| **4.1 Governance:** Describe the **governance mechanisms** in place and what performance and risk management **reporting** will be providedover thealgorithm lifecycle. |  |
| **4.2 Accountability:** Who has **ultimate accountability** if things go wrong? Who is **responsible** for each stage of the algorithm lifecycle, including risk management and **addressing** errors, failures or unexpected outputs? |  |
| **4.3 Policies:** Describe what **systems and policies** apply to govern algorithms, any **restrictions** on use of the algorithm and how such restrictions are **monitored**. |  |
| **4.4 Audit:** Describe what **recording** and **audit** mechanisms will be in place and how the algorithm’s **decision-making processes** will be **documented**. How often will this AIA be **reviewed** to ensure the information and assessments remain current? |  |
| **4.5 Human review:** Explain whether the algorithm will **replace** or **assist** human decision-making and if people can **opt out** of being subject to the algorithm’s decisions. What **human review processes** will be in place? |  |
| **4.6 Automation bias:** Describe the extent to which algorithm users might rely on recommendations instead of their own judgement. How will this risk be addressed? |  |
| **4.7 Law:** Please share/summarise any **legal** **advice** on this Project, including compliance with **applicable** **laws** and any **material legal risks** that have been identified to date. |  |
| **4.8 Appeals & recourse:** Describe how algorithm-based decisions can be **challenged, appealed** or **reviewed** and what **processes for actionable recourse** will be in place for people negatively impacted by algorithm outputs and decisions. |  |
| **4.9 Guidance & training:** Describe what support will be provided to help users understand how to **responsibly use the algorithm**, including in relation to checking the **accuracy** of outputs and **minimising risks**. |  |

|  |  |
| --- | --- |
| 1. Partnership with Māori | |
| **5.1** Describe what **engagement with Māori** has occurred/is planned, how **Te Ao Māori** perspectives will be incorporatedinto the design and use of the algorithm and how the Project **upholds the principles and intention of Te Tiriti o Waitangi***.* |  |
| **5.2** What **Māori voices** been involved in the **co-design** of the algorithm, including **Māori individuals, whānau, hapū, iwi, organisations and communities**? If not, why not? What issues were raised and how will they be **addressed**? |  |
| **5.3** Describe what **Māori data** will be used in the algorithm, how it will be treated as **taonga** and how **Māori data sovereignty** will be maintained. |  |
| **5.4** Describe the **impacts** of using the algorithm to make decisions affecting Māori andwhat **actions** will be taken to **mitigate** any **potential negative impacts.** |  |

|  |  |  |
| --- | --- | --- |
| 1. Data | | |
| **6.1 Nature of data:** Describe the **type**, **sensitivity** and **source** of the **data** to be used in relation to the algorithm. |  | |
| **6.2 Personal information:** Will the algorithm use any **personal** or **sensitive** **information** orthedata or images of **children and/or young adults**? Was it collected withthe **consent** or **awareness** of the relevant people? Will people be able to access and correct their personal information used in the algorithm? |  | |
| **6.3 Storage & security:** Describe **where** and **how** training and production data will be **stored** and **secured,** including **security** **arrangements**, **access controls** and data **retention** periods. |  | |
| **Data quality** | ***TRAINING DATA*** | ***PRODUCTION DATA*** |
| **6.4 Data source:** For each data type, describe the data **source** (including country), **who** originally collected it, **how** and **why.** Was it obtained **legally**? |  |  |
| **6.5 Relevance, sufficiency & representation**:   * **When** was the data collected and for how long will it be **relevant** and **accurate**? Is there **enough** of it to **achieve the** **Purpose?** * **What** and **who** is **represented** in the data? Do the data sets accurately **represent** the populations to which the algorithm will be applied? How will risks associated with **over** or **under representation** be addressed? |  |  |
| **6.6 Accuracy -** Describe:   * Whether the data is sufficiently **accurate** and **reliable** to support the **Purpose and** whetherthe production data is of **comparable quality** to the training data. * How **inaccuracies** in the data will be **identified** and **addressed** and data **quality** will be maintained over the Project’s lifecycle. |  |  |
| **6.7 Organisation:** Has the data been appropriately structured, classified and labelled? |  |  |

|  |  |
| --- | --- |
| 1. Privacy | |
| *Where* ***personal information*** *is used in the Project, attach/link to a copy of your* ***Privacy Impact Assessment*** *and answer the following in addition to the privacy-related questions already asked throughout the questionnaire.* | |
| **7.1** Will users’ input data be used to **train the algorithm** and to what extent will such data be **retained** in the algorithm or elsewhere? |  |
| **7.2** Is there any potential for the algorithm to be used for any form of **surveillance**, whether **actual or perceived**? If so, who is likely to be surveilled and by whom? What justifications, if any, are there for this type of surveillance? |  |
| **7.3** Where any **biometric information** and/or **templates** are in use, describe:   * **How, when, where** and **why** such information is collected and **by whom**. * **Who** creates the biometric templates, where are they **stored,** how arethey **protected** and for how long will they be **retained**? Who can **access** them? |  |
| **7.4** If the Project involves **image matching** (for example, facial recognition), describe how the algorithm assesses different **skin tones** and the relevant accuracy levels, whether a **watchlist** will beused and by whom and any **confidence scores** used. |  |

|  |  |
| --- | --- |
| 1. Bias and other unfair outcomes | |
| **8.1** Describe how **“fairness”** will be assessed given the **business context** and the **Purpose.** How will **criteria for optimising** the algorithm’s **performance** be **balanced** against the risks of potential **harm.** Describe any potential **trade-offs.** |  |
| **8.2** Describe how potential **bias, discrimination** and/or other **unfair, unintended** or **unexpected outcomes**(together “**unfair outcomes”**) could be present in the data, the algorithm and the wider Project. For each unfair outcome, describe the **nature** of the harm, **who** might be harmed, **how**, **when** and **why.** |  |
| **8.3 Proxy variables:** Describe what **proxy variables** may be used, the **rationale** for using them, their potential **limitations** and how they might affect the algorithm’s **outputs**. What **steps** have been taken to **mitigate** potential bias risks associated with them? |  |
| **8.4** Are there any **alternative** tools or approaches that could achieve similar results with a **lower risk of inaccurate or unfair outcomes**? |  |

|  |  |
| --- | --- |
| 1. Algorithm development, procurement and monitoring | |
| **Internal algorithm development** | |
| **9.1** Describe the **methodology** for developing the algorithm, including **who** isdoing this work, **where**, **when** and with whose **oversight** and **accountability**. Are any **open source solutions** being used? Detail how the algorithm will be published/made available and its performance metrics to support the Purpose (what does success look like?). Please share documentation of the algorithm’s technical features. |  |
| **External procurement** | |
| **9.2 Nature of procurement:** Describe what is being procured (for example, **off-the-shelf** solution; **open source** solutions, **bespoke development**), **why** that supplier was selected and how the supplier and algorithm support and align with the **Purpose**. |  |
| **9.3 Evaluation:** Describe the **evaluation criteria** used to assess potential suppliers and sources and **performance** against those criteria. What **concerns or risks** were identified and how will they be **addressed**? |  |
| **9.4 Contracts:** Describe the supplier’s **key** **contractual obligations**. |  |
| **9.5 Supplier training data:** How has the supplier provided sufficient **visibility** of the relevance of its **training data** to the proposed usage in New Zealand? Describe any potential for **unfair outcomes** resulting from that training data and how they will be **identified** and **addressed**, especially where the training data is not disclosed. |  |
| **9.6 Supplier access & use:** Describe whether the **supplier** will have **access** to the production data, why that is required and how such access will be **managed.** Will the supplier use or keep production and/or algorithm user data to continue training its own algorithms/models and, if so, who will receive any commercial benefits? |  |
| **Algorithm performance, testing and monitoring** | |
| **9.7 Performance and testing -** Provide details of:   * The **results** of any **testing** conducted **to date**, including whether the algorithm performed **accurately** andin alignment with the Purpose, whether it produced any **unfair outcomes** and its performance compared to the status quo. Please include breakdowns of the performance metrics. * **How** and **when** the algorithm will be **tested in future,** what will that entail, and how **performance, accuracy** and **unfair outcomes** will be measured. * If the algorithm uses a **confidence score**, describe how you determined the appropriate decision **thresholds**, the percentage of **false positives** and **negatives,** the resulting **risks** and any **trade-offs** that have been determined to be **reasonable** in the context of the Purpose. |  |
| **9.8 Ongoing monitoring -** Describe:   * **when, why and how** the algorithm will be **monitored** and reviewed across its lifecycle andhow errors and unfair outcomes will be **identified** * the processes and responsibility for **addressing** errors and unfair outcomes * what will **trigger** changes to the algorithm. |  |
| **9.9 Retire:** Describewhen the algorithm is likely to become **obsolete** and be **retired**? |  |

|  |  |
| --- | --- |
| 1. Safety, security and reliability | |
| *Attach or link to a copy of your* ***Security Risk Assessment*** *and answer the additional questions below* | |
| **10.1** Describe possible algorithm **errors**, their **likely impact** and how will they be **addressed.** How will they be **communicated** to those impacted and more widely? |  |
| **10.2** Describe the extent to which the algorithm could be exposed to external **malicious attack,** including cyber-attacks, adversarial attacks, data poisoning. |  |
| **10.3** Describe the extent to which the algorithm could be used to generate or perpetuate **misinformation** or **disinformation** *(see the* [*User Guide*](https://data.govt.nz/docs/algorithm-impact-assessment-user-guide/#Safety-security-and-reliability) *for more detail).* |  |
| **10.4** If the algorithm **degrades** or its performance is **impaired**, describe the potential **impacts**, **contingency plans,** how it can be **shut down** and **who** should be informed. |  |
| **10.5** Describe measures to ensure data and algorithm **reliability, integrity and security**, including **access controls**, monitoring and auditing. |  |

|  |  |
| --- | --- |
| 1. Community engagement | |
| **11.1** Describe the **consultation** or **engagement** with those most likely to be **impacted** by the algorithm, detailing **who** will be consulted, **how, when** and **why**. Describe what **issues** were raised and how they are being **addressed.** |  |
| **11.2** Describe what **engagement** has already occurred/is planned with people, communities and groups with an **interest in the algorithm generally**. Describe what **issues** were raised and how are they being **addressed.** |  |
| **11.3** Describe how **diverse perspectives and expertise** from relevant individuals or groups have been collected and factored into the algorithm. For example, community representatives, disabled people, children and young people, refugees, migrants, service providers, frontline personnel and cultural experts. |  |

|  |  |  |
| --- | --- | --- |
| 1. Transparency and explainability | | |
| *Answer the questions below in relation to the following groups:* | **Those impacted** | **The wider public** |
| **12.1 Transparency: How** and **when** will the following be communicated:   * **How** and **why** the algorithm is being **used, what** **decisions** are being made by the algorithm, who the algorithm will **impact** and **how** * What **personal information** is involved and how **training and production data** has been collected, secured and stored * Where a user is **interacting with an algorithm,** not a person.   Where applicable, explain **why** (including for **legal compliance** reasons) information about algorithm data cannot or should **not be shared**? |  |  |
| **12.2 Explainability:** Explain **how** and **why** the algorithm produces its **outputs**, including the logic behind its design, how it makes any decisions and whether it is doing so in accordance with the relevant **legal authority**. |  | |