

# New Zealand Government – Draft Algorithm Charter

## Response from Microsoft

*20 December 2019*

Microsoft welcomes the opportunity to contribute to this consultation on the Algorithm Charter.

We welcome and strongly support the Principles enumerated in the Charter. As a global company deeply engaged into Artificial Intelligence (AI) policy issues, we would like to acknowledge that with this Draft Algorithm Charter, New Zealand (NZ) Government places itself among the frontrunners in designing an internal operational framework for responsible AI adoption.

We believe that business leaders, policymakers, researchers, academics, representatives of nongovernmental groups, and communities must work together to ensure that AI-based technologies are designed and deployed in a manner that will earn trust. That means AI systems need to be designed and deployed with protections for fairness, reliability and safety, transparency and accountability, privacy, and be inclusive. Assessing and continuing to address potential risks are also important so that AI technologies are used responsibly.

Government agency commitment to use technology fairly, ethically and transparently – the Principles of the Charter - is essential to continued trust in the government and its use of emerging technologies.

Our response includes General Comments as well as thoughts on the specific Principles of the Charter.

## KEY RECOMMENDATIONS

**We encourage consideration of clarifying the scope of the Charter.** Specifically, we believe that the scope should be broader and not solely focused on *algorithms*. It would be perhaps appropriate to look at the algorithms as part of the *AI systems*. *Artificial Intelligence Charter* title would be able to capture broader aspects than solely algorithms and could take a more expansive approach, encompassing better i.e. humans' role in AI/algorithms adoption.

**We note that there has been progress in developing technical tools for managing some of the challenges** associated with responsible AI, including detecting bias, traceability through model management, AI security, and improving intelligibility (which is already raised in the Charter). We suggest leveraging further these existing efforts in the Charter and in subsequent operational work of the NZ Government and its agencies in line with the Charter.

**Consider proactive reporting on the use of AI systems by the Government.** Because the NZ Government would be one of the first in the world to adopt such set of Principles, it should consider publishing periodic public reports detailing the government's use of AI systems to further demonstrate Government's transparency and accountability in this regard.

**Consider developing practical implementation guidelines and sharing examples of government projects that have piloted the principles in the Charter.** We recommend to provide with the Government entities expected to apply the Charter with some real-life examples on how technical, organizational and policy safeguards could help them deliver on the Charter's objectives. Microsoft is ready to provide some of the best practices observed in other countries or entities.

## COMMENTS ON THE SPECIFIC PRINCIPLES OF THE DRAFT CHARTER

***Clearly explain how significant decisions are informed by algorithms and be clear where this isn't done for reasons of greater public good (for example, national security).***

This is an important Principle. When AI systems are used to help make decisions that impact people's lives, it is particularly important that people understand how those decisions were made. One approach that may engender trust with users and those affected by these systems is to provide explanations that include contextual information about how an AI system works and interacts with data. Such information will make it easier to identify and raise awareness of potential bias, errors, and unintended outcomes.<sup>1</sup>

In addition, where there is a significant decision informed by algorithms - a denial of consequential services, like access to transportation, housing, or healthcare; a risk of harm, like life or death decisions, law enforcement, healthcare, or vulnerable populations; or, a potential for infringement on human rights, like some uses by law enforcement – we suggest that there be a qualified human appropriately involved in the decision-making process by default, sometimes referred to as a “human in the loop.”

Also, the national security exception should be considered in a narrow sense and not used as a loophole to bypass transparency requirement. We believe that in majority of cases there is a degree of explanation that may be public without jeopardizing effectiveness of the use case or classified nature of some of the projects. Indeed, providing transparency that AI systems are being used – except where there is a genuine threat to national security if this is disclosed – is important for building public trust in AI.

***Embed a Te Ao Māori perspective in algorithm development or procurement.***

***Take into account the perspectives of communities, such as LGBTQI+, Pasifika and people with disabilities as appropriate.***

Inclusiveness is a fundamental Principle. All stakeholders have an important role to play in how to address AI issues and their implications. If we are to ensure that AI technologies benefit and empower everyone, they must incorporate and address a broad range of human needs and experiences.

We applaud the recognition of a Te Ao Māori vital perspective that must be considered in AI as it is developed and deployed – it will be critically important to the Government's Social License in an AI area that will impact everybody's life.

---

<sup>1</sup> Microsoft, *The Future Computed*, page 74 (2018).

Considering the perspectives of potentially impacted communities and populations should also mean that it enables the technology to empower those communities. Hence, we believe that on the inclusive design practices should not only take into account *perspectives* but also *needs* of the communities and ultimately will help system developers to build products that will serve people of all abilities and backgrounds.<sup>2</sup>

***Identify and consult with groups or stakeholders with an interest in algorithm development.***

Fostering dialogue and sharing best practices will help to deliver broad-based benefits of AI, mitigate risks, and minimize unintended consequences. Open discussions with civil society, academic researchers, other individuals and organizations can create a culture of cooperation, identify concerns, and assist in the sharing of best practices.

We, therefore, encourage NZ Government agencies implementing this provision to setup regular working groups in order to provide with appropriate fora for ongoing dialogue.

***Publish information about how data are collected, secured and stored.***

Transparency about the collection of data and utilization is important. Likewise, information about how data is stored and used should be disclosed. We suggest to consider adding a security angle to this principle.

We believe that data practices are foundational to trust. In that spirit we believe that this Principle should consider additional elements.

- Information should also be provided, where appropriate, about other important aspects related to data, such as a person's ability to access and correct the data, or how their data might be shared. People should also be informed when they can consent to the data collection for specific purposes.
- Risk assessments should be considered with respect to data used in AI systems and policies and procedures should include practices related to data used in such applications.
- AI systems should follow applicable privacy laws – developers as well as agencies implementing such systems need to consider *privacy-by-design* principle.
- Data that is collected and stored should be secured so that it is protected from misuse or theft – *security-by-design*. In particular, developers and implementing agencies should understand and mitigate threats to AI systems that are different to traditional IT systems – for instance, intentional input data pollution in order to manipulate algorithmic training and ultimately alternate AI systems' decisions.

---

<sup>2</sup> Microsoft, *The Future Computed*, page 71 (2018).

***Upon request, offer technical information about algorithms and the data they use.***

We support transparency about technical information and data. We also believe that such information should be sufficient to allow for testing and include documentation. We suggest including more specificity in this Principle (or in the potential accompanying guidelines), especially where AI systems are used to make consequential or significant decisions.

- Application programming interfaces (APIs) or other technical capabilities should be made available to enable reasonable testing of AI systems for accuracy and unfair bias.
- Documentation should be provided that explains the capabilities and limitations of the technology in terms that the deploying agency can understand.
- Documentation to be provided should therefore include both technical and non-technical information depending on the audience. A key finding from Microsoft’s work on Responsible AI is that the context in which an AI system is used needs to be considered when implementing the transparency principle. Deployment of the same system may warrant different approaches to transparency depending on the context and audience. For example, in sharing information on the functioning of Microsoft’s Face API (a building block for a facial recognition system), we have issued both technical documentation for developers, as well as a less technical Transparency Note that explains for a less technical audience the intent, capabilities, and limitations of Face API.
- The Government will also need to balance transparency in AI functions and data with any potential tradeoffs. That is, it should make information transparent keeping in mind to protect confidentiality, privacy, cybersecurity, and other concerns.
- As the Government seeks to implement best practices on transparency of AI systems, we would like to draw the Government’s attention to a significant project being led by the Partnership on AI (the “ABOUT-ML” project) to define best practices for implementing transparency through the ML lifecycle<sup>3</sup>. The learnings from this project will be valuable as New Zealand seeks to improve transparency – and we would also encourage the government to share its insights with the Partnership on AI.

***Use tools and processes to ensure that privacy, ethics, and human rights considerations are integrated as a part of algorithm development and procurement.***

We strongly support this Principle. It is vital to ensure privacy, ethics, and human rights are considered early in the process - and not only considered at the implementation phase, when significant investments have been already made into a development of the AI system and there may be organizational pressure to deploy it. Steps that can be taken to meet this Principle include:

- Conducting a thorough risk assessment. The assessment should describe the risks associated with uses of the technology, including risks to the privacy of individuals and the risk of disparate impacts for certain groups or communities, and the specific steps the

---

<sup>3</sup> See <https://www.partnershiponai.org/about-ml/>

agency plans to take to minimize those potential risks and to prevent unauthorized use of the technology.

- Prior to deployment there should be a robust testing of the system, including thorough pilot implementations intended to observe system behavior in operational conditions. Such testing should include assessments of relative performance across different populations (including groups determined by similar characteristics as well as communities). Those intending to use such technologies should also consider engaging local community members to participate in such pilots and provide feedback and input about the system's intended uses and performance.
- There should be adequate training for all persons tasked with operating and using the technology or any data obtained from the technology. Such training should cover the capabilities and limitations of the technology, as well as how to interpret and act on the output of the technology. The procedures should also require appointing qualified staff to oversee the deployment of the AI system.
- In line with the recommendation above to share implementation guidance on the Charter, we recommend that information is shared across government agencies on existing and emerging technical methods that have a role in addressing issues such as fairness, bias, traceability and transparency relating to AI systems. For example, a Microsoft Research project has resulted in an open source set of tools that may support the detection and mitigation of bias<sup>4</sup>. Another example is an open source toolkit released by Microsoft Research that addresses transparency and accountability concerns by making "black box" Machine Learning models interpretable<sup>5</sup>.

***Regularly collect and review data relating to the implementation and operation of algorithms, and periodically assess this for unintended consequences, for example bias.***

Ensuring AI systems are operating as intended is important. Establishing a process for periodic assessment, including for unintended consequences, like bias, is essential for transparency and accountability.

- Consideration should be given to publishing the results of the assessment.
- Procedures for a periodic review should also include a process or mechanism for addressing the findings, including mitigating any unintended consequences, like bias.

---

<sup>4</sup> See <https://www.microsoft.com/en-us/research/publication/a-reductions-approach-to-fair-classification/>

<sup>5</sup> See: <https://github.com/interpretml/interpret>

***Have a robust approach for peer-reviewing these findings.***

As pointed out above, we believe that establishing relevant review processes is important.

- In addition to peer-reviews, the review process should include ways to receive feedback from communities that are affected using the AI system.
- The approach should also create procedures for responding to the peer and community reviews.

***Clearly explain who is responsible for automated decisions and what methods exist for challenge or appeal via a human.***

This is a critical Principle – to ensure consequential decisions are subject to meaningful human review. The purpose of meaningful human review is to help detect and resolve cases of misidentification or other failures.

- While we recognize the importance of educating users that challenging AI system's decision to a human is possible and makes part of the regular course of the system's operational practice, consideration should be given to providing for meaningful human review not just when a decision is challenged, but more generally, when decisions are made. The mere fact that the decision was not challenged does not mean it was an optimal one (providing best outcome to citizen[s]) or that no discrimination has occurred.

## CONCLUSIONS

Adopting the Charter and implementing its Principles will provide increased confidence in how the government uses AI. Government agencies will also be leading by example – and therefore help move the responsible use of AI forward. As mentioned before, we also believe this will provide an opportunity for the NZ Government to be a frontrunner in practical implementation of AI policy.

We further believe that considering the scope extension of the Charter to cover broader Artificial Intelligence would make it more future-ready. We believe that the attempt, already featured in the charter, to balance transparency, innovation, community considerations and accountability provides great foundation for this enhanced scope.

Microsoft is currently in the process of implementing its own internal policies around Responsible AI<sup>6</sup>, as we develop and deploy AI across many governments, enterprises and non-governmental organisations. We would welcome continued opportunities to exchange with NZ Government each other's learnings from the AI implementation processes, including best practices or governance mechanisms.

We believe NZ Government's Charter will be a foundational step to help instill confidence that the government's commitment to using AI in a responsible way is real. How the Charter is implemented and the resulting practices, will help ensure ongoing trust in the government's use of AI.

*Kind regards,*

**Maciej Surowiec**

Government Affairs Manager

Microsoft

[Maciej.Surowiec@Microsoft.com](mailto:Maciej.Surowiec@Microsoft.com)

---

<sup>6</sup> <https://www.microsoft.com/en-us/AI/our-approach-to-ai>