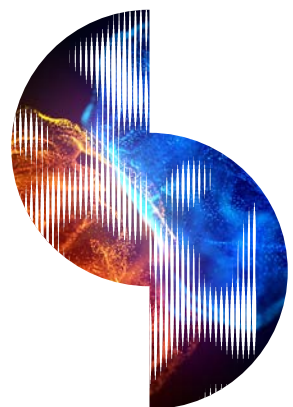


# Putting people at the centre of AI – impacted communities and missing voices

AI governance snapshot #2

**Shaping  
our Future**

Symposium 2024



## Putting people at the centre of AI – impacted communities and missing voices

AI governance snapshot #2

This document is the second in a series of Artificial Intelligence (AI) Governance Snapshots from the UTS Human Technology Institute (HTI). HTI has prepared these governance snapshots as part of the AI Corporate Governance Program to help build greater shared understanding and awareness of the practical steps that organisations seeking to transform AI governance practice can take.

These snapshots are living documents which evolve to incorporate the latest research and emerging trends in responsible AI governance transformation.

### Pursuing human-centred AI innovation

Engagement with people potentially impacted by an AI system is fundamental to any human-centred approach to AI innovation. Human-centred AI *“utilizes data to empower and enable its human users, while revealing its underlying values, biases, limitations, and the ethics of its data gathering and algorithms to foster ethical, interactive, and contestable use”*.

When meaningfully integrated into AI design, development and monitoring processes, the voice of impacted communities and their representatives play two crucial roles:

- **Strengthening accountability:** organisations must engage potentially impacted communities across the lifecycle to accurately identify and manage the harms of an AI system and reduce the risk of automated bias. This important accountability mechanism is most effective when coupled with responsive, meaningful changes to practice, along with transparency measures. Combined, these accountability measures can work to build trust and ensure technologies are beneficial for people and planet.
- **Supporting fit-for-purpose innovation:** engagement with diverse end users can ensure that the products, services and experiences driven by new technologies meet evolving needs and expectations of consumers and employees.

### AI Impact Assessments (AIAs) – a core component of AI governance frameworks

A growing number of AI governance policies and frameworks require consideration of the impact of AI systems on communities to identify, assess, and respond to potential harms and risk. These include:

- **Public sector AI governance frameworks:** including the Canadian Government Algorithmic Impact Assessment Tool<sup>ii</sup> and associated transparency requirements; UK Algorithmic Transparency Recording Standard<sup>iii</sup>; and NSW Government Artificial Intelligence Assurance Framework<sup>iv</sup>.
- **Voluntary AI risk management frameworks:** such as the NIST AI Risk Management Framework<sup>v</sup>, ISO Standard 42001:2023 on AI management systems.
- **Corporate policies:** for example, Microsoft Responsible AI Impact Assessment Template<sup>vi</sup>, and Atlassian’s Responsible Technology Review Template<sup>vii</sup>

In addition, ISO Standard 42005, currently under development, aims to provide guidance for organisations performing AI system impact assessments, including consideration of key documentation and appropriate stages of the AI system lifecycle<sup>viii</sup>.



# Putting people at the centre of AI – impacted communities and missing voices

## AI governance snapshot #2

Engaging potentially impacted communities when undertaking suitability and risk assessments is a fundamental way to build understanding of what harms may arise. Given the diversity of AI applications and input data, it is extremely unlikely that the requisite knowledge to make judgements about – or design controls for – potential AI harms will rest within a single organisation. Engagement with a diverse group of stakeholders is crucial to address blind spots and bias. Enacting the principle of ‘nothing about us without us’<sup>ix</sup> is crucial to both government and corporate AI governance systems.

### Existing stakeholder engagement and impact assessment mechanisms

Engagement with consumers, employees and citizens as an accountability practice has long-standing roots in several fields including the pursuit of consumer-centric market outcomes, workplace safety and deliberative democracy. Privacy Impact Assessments and Human Rights Impact Assessments are also undertaken by some organisations, with significant areas of overlap with the emerging field of AI Impact Assessments.

As engagement with impacted communities increasingly forms a central component of responsible AI governance frameworks, it is important to reflect on existing stakeholder engagement mechanisms and assessment processes and the lessons which may be relevant within an AI context.

### Organisational governance

Stakeholder engagement mechanisms at an organisational governance level are often designed to bring outside insight to the Board and Executive table. These structures tend to be advisory in nature without decision-making authority.

- **Advisory Boards** – often established to bring new insights, solve business problems; or explore new opportunities by stimulating high-quality strategic conversations with external expertise. They can also enable advocacy within approved guidelines for the organisation. Advisory boards do not have binding decision-making authority but can act as an important sounding board for directors, executives or shareholders<sup>x</sup>.
- **Customer Councils or Reference Groups** – common in essential markets and service delivery, Customer Councils generally comprise a diverse mix of customers, consumer or community representatives<sup>xixii</sup>, or a blend of the two<sup>xiii</sup>. These structures often include representation from executive or management teams and are consultative in nature, with no decision-making authority.
- **Ethics Boards and Committees** – these bodies are often formed with diverse leaders across an organisation and sometimes include external parties. They can be responsible for review, risk assessment and decision-making in relation to policies, practices, communications, research, products and services<sup>xiv</sup>. One of the common criticisms of such bodies is they often lack transparency<sup>xv</sup> or the appropriate decision-making authority or veto rights.



# Putting people at the centre of AI – impacted communities and missing voices

## AI governance snapshot #2

### System, project and product governance

A range of accountability mechanisms exist in the fields of privacy, human rights, consumer protection and workplace safety which require organisations to engage with and/or make assessments about the impact of operations or products and services on consumers, employees and citizens. They range from voluntary to mandatory obligations to ‘assess impact’, ‘engage’ or ‘consult’ with impacted communities. Examples include:

- **Privacy Impact Assessments (PIAs):** A PIA is ‘a systematic assessment of a project<sup>xvi</sup> that identifies the impact that the project might have on the privacy of individuals, and sets out recommendations for managing, minimising, or eliminating that impact’.<sup>xvii</sup> Australian Government agencies subject to the Privacy Act 1988 (Cth) must conduct a PIA for all high privacy risk projects and list PIAs conducted on a register. Identifying and consulting with stakeholders is a key step in the recommended PIA process and crucial to determining whether the impact may be ‘significant’.<sup>xix</sup> An evaluation of the quality and effectiveness of such engagement processes is yet to be undertaken. There is no transparency requirement for PIAs that are produced.
- **Human Rights Impact Assessments (HRIAs):** Businesses have a responsibility to respect human rights. To meet that responsibility, they should have in place due diligence processes to identify, prevent, mitigate; and account for their impact on human rights.<sup>xx</sup> These processes should involve meaningful consultation with affected stakeholders.<sup>xxi</sup> Whilst HRIAs are not mandatory, they provide businesses with a process for identifying, understanding, assessing and addressing the human rights impacts of a project or activity on affected stakeholders.<sup>xxii</sup> As a voluntary framework, there is no requirement for transparency.

- **Financial product Design and Distribution Obligations (DDOs):** Financial Service Licence (FSL) holders have specific product governance obligations to design financial products that are consistent with the objectives and needs of the consumers for whom they are intended. DDOs include requirements for product design, distribution, information sharing and outcomes monitoring. Issuers must assess suitability of the product for the likely objectives, financial situation and needs of consumers in the defined target market<sup>xxiii</sup>. Australian Security and Investments Commission (ASIC) guidance however does not yet stipulate how consumer needs are to be assessed, nor ‘customer classes’ defined by issuers via consumer consultation or engagement mechanisms, particularly with vulnerable or at-risk communities. Target market determinations which describe who a product is for must be publicly available and free of charge.
- **Safe Work Method Statements (SWMS):** A SWMS is required by all persons conducting a business or undertaking engaging in high-risk construction work activities. A SWMS must be completed before work begins and stipulates the high-risk work activities to be carried out in the workplace; the hazards arising; and measures to be put in place to control the risk<sup>xxiv</sup>. SWMSs must be prepared in consultation with workers. This duty to consult recognises that worker participation improves WHS decision-making and assists in reducing work-related injuries and disease<sup>xxv</sup>. SWMSs must be transparent and easily understood by workers.

As these diverse examples demonstrate, very little consistency or guidance is provided as to how organisations are to engage with potentially impacted parties, and what transparency measures are appropriate.



# Putting people at the centre of AI – impacted communities and missing voices

## AI governance snapshot #2

### Key considerations for the engagement of impacted communities as an AI accountability measure

Additional guidance is required to ensure that the engagement of potentially impacted communities forms a meaningful and practical AI guardrail, as well as an accelerator of human-centered innovation.

Not all participation by impacted communities is equal and there is no ‘one size fits all’ approach. It is important that organisations seeking to embed participatory processes identify the appropriate method depending on the context and objectives.

For example, Ada Lovelace Institute has proposed that participatory data stewardship occurs along a spectrum<sup>xxvi</sup> (see Figure 1), and that public participation throughout the AI lifecycle is often for three distinct purposes:

- *work* - recognising the intensive labour that goes into AI system production or refinement
- *consultation* - gathering feedback from experts or potential users in design processes
- *justice* - longer-term partnerships which provide impacted communities agency over design and infrastructure.



Figure 1: Ada’s spectrum of participatory data stewardship

HTI will be producing further research in this field and will update this snapshot as more findings on leading practice become available.

Organisations seeking to establish engagement mechanisms as part of any AI governance framework should consider:

1. Clearly defining the purpose of stakeholder participation and engagement and the extent of stakeholder agency as part of that process.
2. Any AI impact assessment process should have regard to existing individual rights-based obligations (human rights, consumer rights and workplace rights), including any overlaps and gaps with other suitability or assessment frameworks.
3. Complementary measures such as transparency requirements about assessment processes and their outcomes are key to increasing accountability and building trust.
4. Impacted communities and their representatives must be appropriately resourced to participate in engagement processes. Engagement processes themselves must be designed inclusively to ensure broad and diverse participation.
5. Significant challenges exist for public participation as an accountability mechanism for general purpose AI / foundation models.



## Putting people at the centre of AI – impacted communities and missing voices

### AI governance snapshot #2

- i. Tara Capel and Margot Brereton, 'What is Human-Centered about Human-Centered AI? A Map of the Research Landscape' (Conference Paper, CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing, April 23 2023) Systems  
<[https://eprints.qut.edu.au/240755/1/What\\_is\\_human\\_centered\\_about\\_human\\_centered\\_AI\\_a\\_map\\_of\\_the\\_research\\_landscape.pdf](https://eprints.qut.edu.au/240755/1/What_is_human_centered_about_human_centered_AI_a_map_of_the_research_landscape.pdf)>.
- ii. Government of Canada, Algorithmic Impact Assessment (Web Page, 12 April 2023)  
<<https://open.canada.ca/aia-eia-js/?lang=en>>.
- iii. United Kingdom Central Digital and Data Office and Centre for Data Ethics and Innovation, Algorithmic Transparency Recording Standard (Web Page, 6 December 2021)  
<<https://www.gov.uk/government/publications/algorithmic-transparency-template>>.
- iv. NSW Government, NSW Artificial Intelligence Assurance Framework (Web Page)  
<<https://www.digital.nsw.gov.au/policy/artificial-intelligence/nsw-artificial-intelligence-assurance-framework>>.
- v. National Institute of Standards and Technology (NIST), Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1, January 2023)  
<<https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf>>.
- vi. Microsoft, Responsible AI Impact Assessment Template (Template, June 2022)  
<<https://blogs.microsoft.com/wp-content/uploads/prod/sites/5/2022/06/Microsoft-RAI-Impact-Assessment-Template.pdf>>.
- vii. Atlassian, Responsible Technology Review Template (Template, 2023)  
<<https://www.atlassian.com/trust/responsible-tech-principles>>.
- viii. International Organisation for Standardization, ISO/IEC DIS 42005 Information Technology Artificial Intelligence AI system impact assessment (Web page)  
<<https://www.iso.org/standard/44545.html>>.
- ix. ARC Centre of Excellence for Automated Decision-Making and Society, Safe and Responsible AI in Australia Discussion Paper ADM+S Submission (Submission, 4 August 2023)  
<<https://apo.org.au/sites/default/files/resource-files/2023-08/apo-nid323896.pdf>>.
- x. Advisory Board Centre, What are advisory boards roles & responsibilities? (Article, 14 January 2023)  
<<https://www.advisoryboardcentre.com/insight/what-are-the-advisory-board-roles-and-responsibilities/>>.
- xi. AGL, Our Customer Council (Web Page)  
<<https://www.agl.com.au/about-agl/who-we-are/our-commitments/our-customer-council?rmxt3r=vwt06z>>.
- xii. Sydney Water, Customer and Community Reference Group (Web Page, 2024)  
<<https://www.sydneywater.com.au/about-us/our-people/who-we-are/customer-forums.html>>.
- xiii. SA Power Networks, Customer and stakeholder engagement (Web Page)  
<<https://www.sapowernetworks.com.au/about-us/customer-and-stakeholder-engagement/>>.
- xiv. IBM, AI Ethics (Web Page)  
<<https://www.ibm.com/impact/ai-ethics>>.
- xv. James Vincent, 'The problem with AI ethics' The Verge (Online, 4 April 2019)  
<<https://www.theverge.com/2019/4/3/18293410/ai-artificial-intelligence-ethics-boards-charters-problem-big-tech>>.



## Putting people at the centre of AI – impacted communities and missing voices

### AI governance snapshot #2

- xvi. Where ‘project’ can include: new policy proposals, legislation, new or changed programs or activities, implementing or expanding IT systems or databases, new or changed methods for service delivery or information handling, implementing AI technologies and automated decision, restructures or changes to business-as-usual activities (Office of the Australian Information Commissioner, When do agencies need to conduct a privacy impact assessment? (Web Page)  
**<<https://www.oaic.gov.au/privacy/privacy-guidance-for-organisations-and-government-agencies/government-agencies/australian-government-agencies-privacy-code/when-do-agencies-need-to-conduct-a-privacy-impact-assessment>>.**
- xvii. Office of the Australian Information Commissioner, Guide to undertaking privacy impact assessments (Web Page)  
**<<https://www.oaic.gov.au/privacy/privacy-guidance-for-organisations-and-government-agencies/privacy-impact-assessments/guide-to-undertaking-privacy-impact-assessments>>.**
- xviii. Office of the Australian Information Commissioner, Guide to undertaking privacy impact assessments (Web Page)  
**<<https://www.oaic.gov.au/privacy/privacy-guidance-for-organisations-and-government-agencies/privacy-impact-assessments/guide-to-undertaking-privacy-impact-assessments>>.**
- xix. Office of the Australian Information Commissioner, ‘Undertaking a PIA’, Guide to undertaking privacy impact assessments (Web Page)  
**<<https://www.oaic.gov.au/privacy/privacy-guidance-for-organisations-and-government-agencies/privacy-impact-assessments/guide-to-undertaking-privacy-impact-assessments#undertaking-a-pia>>.**
- xx. United Nations, Guiding Principles on Business and Human Rights, Principle 15  
**<[https://www.ohchr.org/sites/default/files/Documents/Publications/GuidingPrinciplesBusinessHR\\_EN.pdf](https://www.ohchr.org/sites/default/files/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf)>.**
- xxi. United Nations, Guiding Principles on Business and Human Rights, Principle 18  
**<[https://www.ohchr.org/sites/default/files/Documents/Publications/GuidingPrinciplesBusinessHR\\_EN.pdf](https://www.ohchr.org/sites/default/files/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf)>.**
- xxii. The Danish Institute for Human Rights, Human Rights Impact Assessment Guidance and Toolbox (Report, 2020)  
**<[https://www.humanrights.dk/sites/humanrights.dk/files/media/document/DIHR%20HRIA%20Toolbox\\_Welcome\\_and\\_Introduction\\_ENG\\_2020.pdf](https://www.humanrights.dk/sites/humanrights.dk/files/media/document/DIHR%20HRIA%20Toolbox_Welcome_and_Introduction_ENG_2020.pdf)>.**
- xxiii. Australian Securities & Investments Commission, Product design and distribution obligations (Regulatory Guide 274, December 2020)  
**<<https://download.asic.gov.au/media/dlpccdof/rg274-published-11-december-2020-20220628.pdf>>.**
- xxiv. Safe Work Australia, Safe Work Method Statement for High Risk Construction Work (Information Sheet, December 2014)  
**<<https://www.safeworkaustralia.gov.au/system/files/documents/1703/information-sheet-safe-work-method-statement.pdf>>.**
- xxv. Worksafe ACT, Safe Work Method Statements – High Risk Construction Work (Guidance Note)  
**<[https://www.worksafe.act.gov.au/\\_\\_\\_data/assets/pdf\\_file/0010/2187874/SWMS-Guidance-Note.pdf](https://www.worksafe.act.gov.au/___data/assets/pdf_file/0010/2187874/SWMS-Guidance-Note.pdf)>.**
- xxvi. Ada Lovelace Institute, Going public Exploring public participation in commercial AI labs (Discussion Paper, December 2023)  
**<[https://www.adalovelaceinstitute.org/wp-content/uploads/2023/12/2023-12\\_ALI\\_Going-public\\_Discussion-paper.pdf](https://www.adalovelaceinstitute.org/wp-content/uploads/2023/12/2023-12_ALI_Going-public_Discussion-paper.pdf)>.**

