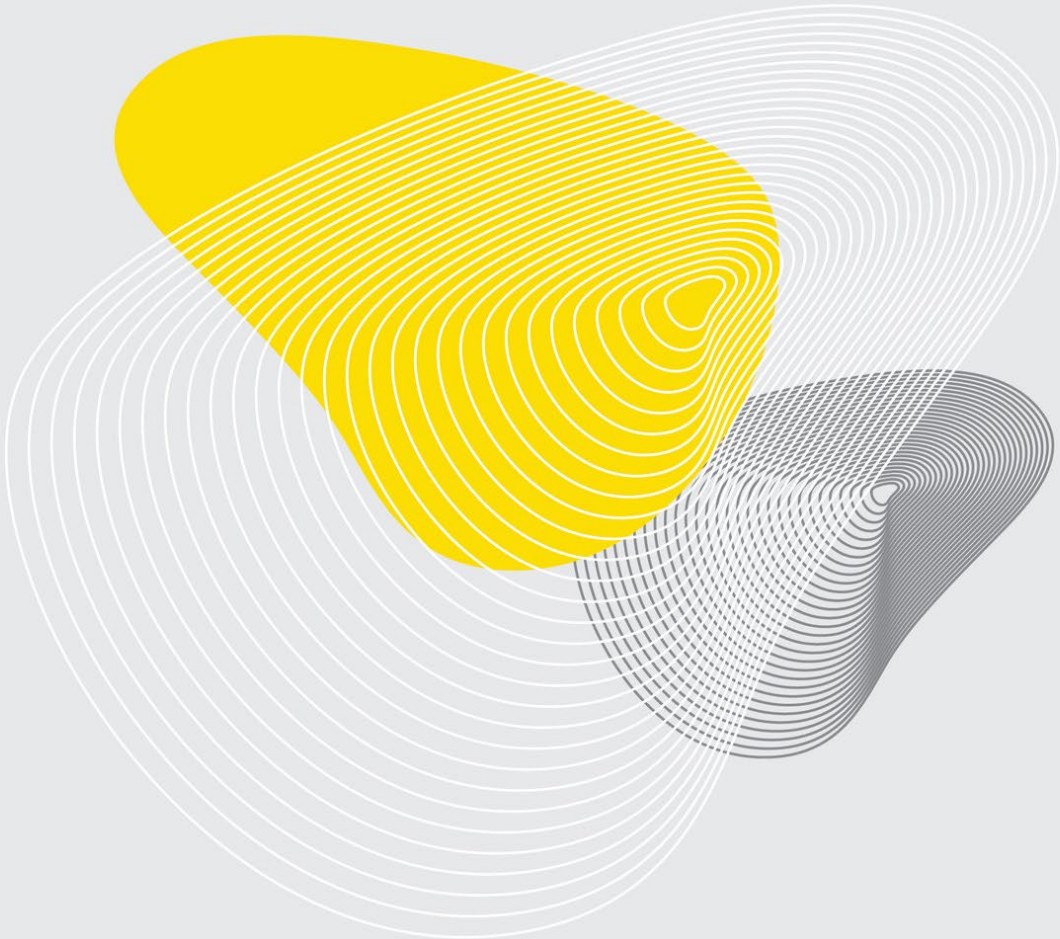
Key Questions for AI in Public Institutions

Prompts and resources to spark critical conversations about AI in universities, social services and beyond

Tanja Dreher and Georgia van Toorn

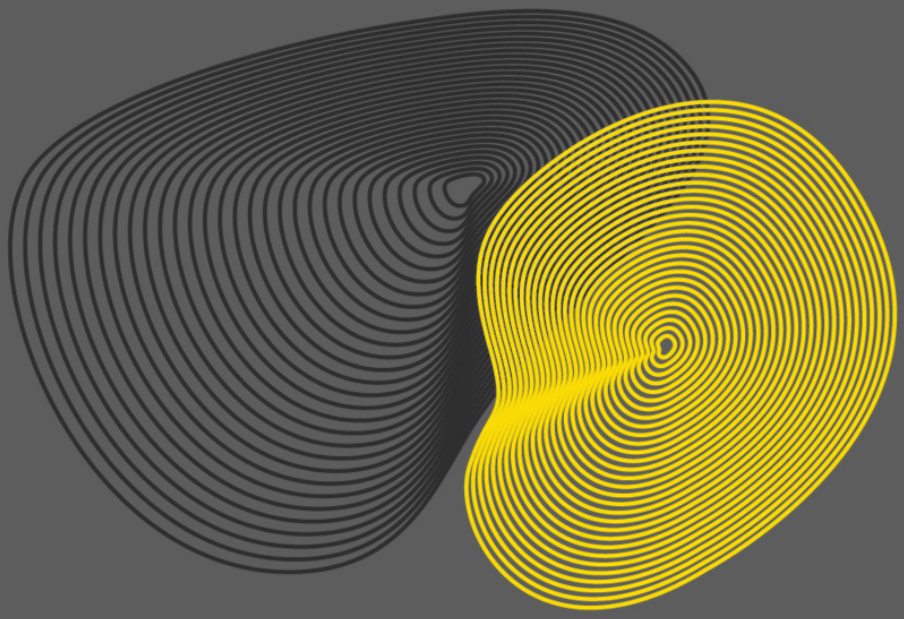
2025



UNSW SYDNEY
MEDIA FUTURES HUB
ADM +S ARC Centre of Excellence for Automated Decision-Making and Society

Contents

Key Questions for AI in Public Institutions



First Published December 2025

Suggested Citation:

Dreher, T., & van Toorn, G. (2025) Key Questions for AI in Public Institutions. ARC Centre of Excellence for Automated Decision-Making and Society.
DOI: 10.60836/tb4d-en57

Acknowledgement of Country

We acknowledge the unceded Country of the Bedegal, Wangal, Dharawal and Wodi Wodi peoples on whose lands we live and work. We pay respects to Elders past and present.

In Acknowledging Country we also acknowledge First Nations scholarship foregrounding the coloniality of AI and questions of Indigenous Data Sovereignty and self-determination. We also acknowledge that scientific experimentation always occurs in place, despite ubiquitous references to "the cloud" and other placeless metaphors that obscure the material sites and ecological environments that make such experimentation possible. We acknowledge Indigenous expertise on histories and current manifestations of surveillance, extraction and exploitation, bias and discrimination and on knowledges for more just futures.

We acknowledge the imperative to centre Indigenous voices in the development and governance of AI. The 2025 UN International Day of the World's Indigenous Peoples focused on the theme: *Indigenous Peoples and AI: Defending Rights, Shaping Futures* stating "While AI can support cultural revitalisation, youth empowerment, and even adaptation to climate chance, it often reinforces bias, exclusion, and misrepresentation towards Indigenous Peoples.... To unlock AI's full potential, Indigenous Peoples must be respected as rights-holders, co-creators, and decision-makers."

Sources and Further Resources

[WEB] **Indigenous Peoples and AI: Defending Rights, Shaping Futures.** https://social.desa.un.org/sites/default/files/IDWIP%202025%20Concept%20Note%20FINAL.pdf

[ARTICLE] **AI For Life: How Sovereign Wiradyuri Ways of Knowing Can Transform Technology for Good.** https://theconversation.com/ai-for-life-how-sovereign-wiradyuri-ways-of-knowing-can-transform-technology-for-good-243270

[ARTICLE] **Envisioning Aboriginal and Torres Strait Islander AI Futures Communique: March 2025.** https://doi.org/10.54760/001c.133656

[ARTICLE] **AI effects everyone – including Indigenous people. It's time we have a say in how it's built.** https://theconversation.com/ai-affects-everyone-including-indigenous-people-its-time-we-have-a-say-in-how-its-built-239605

[POSITION PAPER] **Indigenous Protocol and Artificial Intelligence.** https://spectrum.library.concordia.ca/id/eprint/986506/7/Indigenous_Protocol_and_AI_2020.pdf

[UN STATEMENT] **Indigenous Peoples and AI: Defending Rights, Shaping Futures.** https://social.desa.un.org/sites/default/files/IDWIP%202025%20Concept%20Note%20FINAL.pdf

[PAPER] **Testing-in-the-wild: Innovation nationalism and the colonial dynamics of new technology testbeds.** https://journals.sagepub.com/doi/10.1177/29768640251392332

[WEBSITE] **Maiam nayri Wingara Indigenous Data Sovereignty collective.** https://www.maiamnayriwingara.org/

[PAPER] **Indigenous Data Governance**

Acknowledgements

We would like to thank Bec Coleman-Hicks for invaluable research assistance on a related research project. Thank you also to the Faculty of Arts, Design and Architecture at the University of New South Wales for a Faculty Research Grant and the ARC Centre of Excellence for Automated Decision-Making and Society (ADMS) for editorial assistance.

Author Bios

Dr Tanja Dreher

Tanja is a media and cultural studies scholar whose work addresses contemporary issues in media and technology through a social justice lens. Her work is grounded in the politics of voice and listening to develop interventions against settler colonialism, racism in media, and the harms of datafication. She is an Associate Professor in Media at UNSW and a Co-Director of the UNSW Media Futures Hub in which she co-leads the Data Justice stream.

DrTanja Dreher

Dr Georgia van Toorn

Georgia is a political sociologist specialising in international social policy, politics, disablement, and social justice. Her research explores global transformations in welfare governance, with a particular focus on the growing impact of data analytics and algorithmic decision-making in the public sector. She is an Associate Investigator at the ARC Centre of Excellence for Automated Decision-Making and Society.

Dr Georgia van Toorn

Introduction

Artificial Intelligence (AI)/GenAI is being rapidly adopted across public institutions and government departments, including in public universities and social services in Australia.

The speed and scale of AI rollout has prompted concerns about environmental impacts and intellectual property, corporate influence, amplifying discrimination and more. Educators have signed open letters and published articles urging caution on the rollout of AI in universities, and students have signed petitions against the compulsory use of GenAI.

In this think piece we sketch the impacts and implications of corporate AI for institutions with a public interest orientation. We focus on public universities and social services as vital public institutions that have been widely understood as essential to democracy and social justice. Through a series of prompts, we aim to provoke conversations, advocacy and policy engagement by providing resources for further thinking and action.

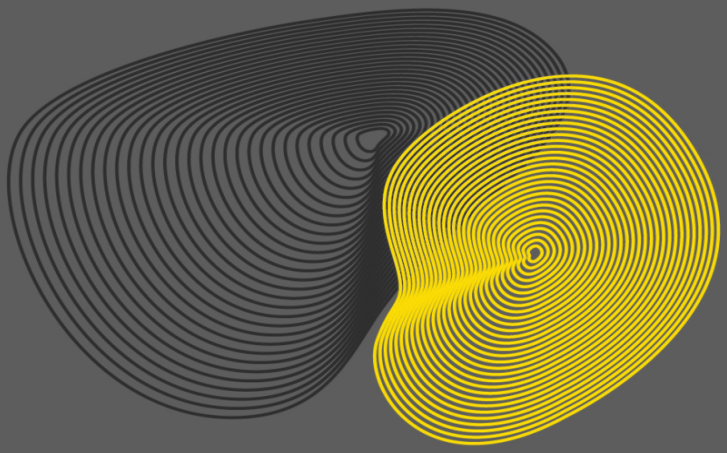
Our Approach

We approach the topic through a Data Justice lens. Data Justice is a framework that focuses attention on the societal impacts and social justice implications of data-driven systems, including AI. This approach expands the frame beyond individual privacy, copyright, efficiency and convenience, to centre social justice and public interest concerns. A Data Justice approach includes listening to the most impacted, engaging with civil society organisations, and collective responses. Data Justice challenges the seeming inevitability of Big Tech, aiming instead to foster alternative imaginaries of the society that we want, and then asking what might be the place for AI in a more just future.

Writing from settler colonial Australia, our Data Justice approach is grounded in respect for First Nations sovereignty and knowledges, prioritising First Nations expertise. We also draw on traditions of critical sociology, media, and communications scholarship and practice. Critical EdTech and critical GovTech scholarship provides vital starting points – including the inherent tensions between the corporate values of Silicon Valley's Big Tech and the public interest values of universities in Australia, and the public good mission of social services. We are also inspired by interventions that pose critical questions for EdTech (Castaneda and Selwyn 2018), and before that, Big Data (boyd and Crawford 2012), and for AI from the perspectives of the Majority World (Ong et al 2024, Amrute et al 2022).

Prompts for critical conversations

Over the following pages, we present a series of prompts intended to spark critical conversations on AI in public institutions. This is not a comprehensive overview, but rather an invitation to think together, slow down, ask questions, and make space for considering the social and democratic implications of corporate AI as it is rapidly shaping institutions including



public universities and social services. We present eight short discussion starters and resources for further conversation. Each prompt is illustrated with a brief example and includes a list of further sources, tools and organisations. Each short discussion starter suggests that 'we need to talk' about key issues, including:

1. We need to talk about business models
2. We need to talk about climate impacts
3. We need to talk about bias and discrimination
4. We need to talk about surveillance, control and coercion
5. We need to talk about consent and refusal
6. We need to talk about expertise and voice
7. We need to talk about participation and co-option
8. We need to talk about alternatives

To end, we offer a series of questions that might provoke more critical conversations about AI in our public institutions.

Sources

[PRIMER] **A Primer on AI in/from the Majority World: An Empirical Site and a Standpoint.** http://dx.doi.org/10.2139/ssrn.4199467

[OPEN LETTER] **An open letter from educators who refuse the call to adopt GenAI in education July 2025.** https://openletter.earth/an-open-letter-from-educators-who-refuse-the-call-to-adopt-genai-in-education-cb4aee75

[PAPER] **Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon.** https://doi.org/10.1080/1369118X.2012.678878

[PAPER] **More than tools? Making sense of the ongoing digitizations of higher education.** https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-018-0109-y

[PAPER] **Collectivity in data governance and data justice.** https://doi.org/10.1080/1369118X.2025.2478096

[PAPER] **Against the Uncritical Adoption of 'AI' Technologies in Academia.** https://doi.org/10.5281/zenodo.17065099

[REPORT] **Custom Built / Feito Sob Medida: Reforming Tech & Democracy Programs for the Global Majority.** https://glotechlab.net

01. We need to talk about business models

Although not exclusively an American creation, corporate AI is largely a product of Silicon Valley's technology giants. It is an industry built on logics of scale and relentless expansion, of datasets, models, corporations, and their global reach. The engine of this growth is data.

Often described as "the new oil," data has become a key commodity of the global economy. AI systems learn from vast quantities of data to perform tasks, yet the scale of what is required has transformed data collection into a global industry of unparalleled extraction. AI companies mine digital environments and datasets for information to feed their models, typically without compensating the human producers of data. Underlying the AI business model is a stark asymmetry: the ownership and profit from AI models are concentrated in the hands of the few, while the social and environmental costs of data extraction and processing are borne collectively.

What are the implications of AI's political economy of extraction for public institutions, like universities, whose legitimacy and purpose revolve around serving the public good? AI's rapid rise has drawn universities closer into the orbit of Big Tech and the antidemocratic politics espoused by its leading Broligarcs such as OpenAI's Sam Altman. A recent example is OpenAI's ChatGPT Edu, licenced by university clients around the world, including in Australia. These are not merely tools. When integrated so thoroughly into university infrastructures, GenAI brings significant transformations in teaching, learning, administration and institutional governance. These changes should prompt closer consideration of how deeper entanglement with Big Tech is reshaping the public university. Historically, universities have been spaces of independent inquiry and critical thought. Yet growing dependence on, and enthusiasm for, corporate AI tools is increasingly challenging the core public values of higher education, as universities become key sites in the expanding political economy of corporate AI.

Universities have been undergoing processes of corporatisation for some decades. What is new is the overwhelming enthusiasm with which higher education has embraced AI's supposed inevitability and value. This both reflects and reinforces the broader colonisation of public institutions by private, profit-driven logics. The pattern extends beyond academia. In Australia, governments also predominantly contract with global technology firms, who design tools, build public infrastructures, and embed themselves deeply within the machinery of the state, without the public responsibilities of an elected government. There is mounting concern over the growing concentration of technological, informational and political power in private hands, and, crucially, that the communities which public institutions are meant to serve have little influence over how that power is exercised.

ChatGPT Edu

Australia's higher education sector is rapidly embedding generative AI tools through high-profile partnerships, with institutions presenting these collaborations as both innovative and risk-managed. NextEd Group Limited has become the first provider in the VET sector to roll out ChatGPT Edu nationally, while UNSW Sydney has purchased 10,000 licences for all fixed-term and permanent staff, touted as the largest education deal with OpenAI in Australia. These moves mirror a broader global trend, with universities such as Oxford, London Business School, Columbia and Arizona State University striking similar agreements.

Sources

[BLOG]: **From Territorial to Functional Sovereignty: The Case of Amazon.** https://lpeproject.org/blog/from-territorial-to-functional-sovereignty-the-case-of-amazon/

[PAPER]: **Digitalised higher education: key developments, questions, and concerns.** https://doi.org/10.1080/01596306.2024.2408397

[NEWS]: **UNSW Sydney inks Australia's biggest ChatGPT Edu deal with OpenAI.** https://www.unsw.edu.au/newsroom/news/2025/09/unsw-sydney-inks-australias-biggest-chatgpt-edu-deal-with-openai

[VIDEO]: **Karen Hao, "Empire of AI: Dreams and Nightmares in Sam Altman's OpenAI".** https://www.youtube.com/watch?v=-bSdBg6AhOw

Further Resources

[ARTICLE]: **The Broligarchy: The Who's Who of the Silicon Gilded Age.** https://theintercept.com/2025/01/24/podcast-silicon-valley-tech-gilded-age-trump/

[VIDEO]: **WTF are Data Centers? The Rise of the Tech Broligarchy.** https://www.cctv.org/watch-tv/programs/wtf-are-data-centers

[ARTICLE]: **Tech Bro Power Play: Zuckerberg vs. Global Tech Justice.** https://www.techpolicy.press/tech-bro-power-play-zuckerberg-vs-global-tech-justice/

[PAPER]: **When data is capital: Datafication, accumulation, and extraction.** https://journals.sagepub.com/doi/10.1177/2053951718820549

02. We need to talk about climate impacts

Government departments and public universities typically have stated commitments to the United Nations Sustainable Development Goals and to reducing carbon emissions. AI use produces environmental impacts that stand in contradiction to these commitments.

Scholars, journalists and civil society organisations have raised the alarm on the 'hidden costs' of AI contributing to the climate crisis. AI has been described as a process of extraction and exploitation operating at a planetary scale. Whether mining rare earth minerals in Bolivia or labelling LLM data in Kenya or disposing of ewaste in the Democratic Republic of Congo, exploited workers labour in harmful conditions on processes that degrade the land, water and air quality.

Generative AI currently being adopted in Australian universities and social services is known to have exponentially higher environmental impacts than a typical computing workload. A single ChatGPT-5 request uses up to 60 times the energy of a Google search. Each short request can consume 30 mL of water. 2025 annual GPT4 water consumption was equivalent to the annual drinking water needs of 1.2 million people across all queries.

Concerns about the hidden environmental costs of AI also focus on data centres that process enormous stores of data – requiring vast amounts of energy to run and also fresh water for cooling. Nvidia, Google, Microsoft and Meta have already seen their cumulative emissions rise by 72% in the last five years. Meta, Google, and Amazon have pledged to triple nuclear power by 2050 to meet the energy demands of their data centres. Coal plants slated to be retired to support clean energy efforts are now being tapped to meet AI's energy needs. In Australia, Morgan Stanley estimates of the electricity on the country's power grid by 2030. There are reports that data centres in Sydney have strained the areas water supply.

There are already more than 250 data centres across Australia. an additional A$20 billion for new data centres across Sydney and Melbourne, while federal and state governments have announced plans to fast track approvals. As public concerns have grown, reports on corporate greenwashing tactics employed by the Big Tech companies have also emerged. There is a risk that public institutions will also seek to downplay the environmental costs of AI in the headlong rush to adopt the latest Silicon Valley tech.

Resisting Data Centres in the US South

In September 2025, MediaJustice launched a report and campaign against the rapid expansion of data centres across the southern states of the USA. MediaJustice challenges 'Tech giants that drain local resources for massive data centers, sell surveillance tools that target Black, immigrant, and Muslim communities, and monopolize media channels to shape which stories get told and who gets to tell them'. Just as Big Oil plants created "sacrifice zones" of poisoned air, contaminated water, and abandoned communities, Big Tech data centers are now layering a new wave of extraction on top of that history. This continues a history of environmental racism. Black and working class communities have long been primary residents of geographic areas that poison local communities due to environmental degradation from industrial pollution and systemic economic divestment. Data Center Watch tracks grassroots opposition to data center development across the United States and has counted at least 150 community-led protest campaigns, highlighting the health and environmental impacts of data centers, and calling out corporate greenwashing and the lack of transparent, democratic decision-making behind data center development.

Sources

[WEB] **United Nations Sustainable Development Goals.** https://sdgs.un.org/goals

[BOOK] **Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence.** https://yalebooks.yale.edu/book/9780300264630/atlas-of-ai/

[REPORT] **The People Say No: Resisting Data Centers in the South.** https://mediajustice.org/resource/the-people-say-no-report/

[ARTICLE] **Making an image with generative ai uses as much energy as charging your phone.** https://www.technologyreview.com/2023/12/01/1084189/making-an-image-with-generative-ai-uses-as-much-energy-as-charging-your-phone/

[NEWS] **AI is driving data centre growth – and it's bringing environmental challenges.** https://www.abc.net.au/news/2025-05-20/data-centre-growth-energy-water-internet-future/105089876

[NEWS] **Power-hungry data centres scrambling to find enough electricity to meet demand.** https://www.abc.net.au/news/2024-07-26/data-centre-electricity-grid-demand/104140808

[WEBSITE] **Australia Data Centers.** https://www.datacentermap.com/australia/

[ARTICLE] **Amazon investing AU$20 billion to expand data center infrastructure in Australia and strengthen the nation's AI future.** https://www.aboutamazon.com/news/aws/amazon-data-center-investment-in-australia

[PODCAST] **Can we afford to power AI?** https://www.abc.net.au/listen/programs/if-youre-listening/can-we-afford-to-power-ai/105975942

[EXPLAINER] **Privacy International – Humans in the loop: the data labellers behind some of the most powerful LLM's training data sets.** https://privacyinternational.org/explainer/5357/humans-ai-loop-data-labelers-behind-some-most-powerful-llms-training-datasets

[NEWS] **Thirsty data centres threaten to delay thousands of new homes.** https://www.smh.com.au/national/nsw/thirsty-data-centres-threaten-to-delay-thousands-of-new-homes-20240807-p5k0ah.html

Further Resources

[REPORT] **The Environmental Impacts of AI: Primer.** https://huggingface.co/blog/sasha/ai-environment-primer

[REPORT] **Generating Harms: Generative AI's Impact and Paths Forward.** https://epic.org/generating-harms/

[LECTURE] **Kate Crawford: Anatomy of AI.** https://www.youtube.com/watch?v=uM7gqPnmDDc

[NEWS] FAIR: **'Media Need to Report on the Real Cost of Data Centers': CounterSpin interview with Jai Dulani and Vivek Bharathan on data center opposition**. https://fair.org/home/media-need-to-report-on-the-real-cost-of-data-centers/

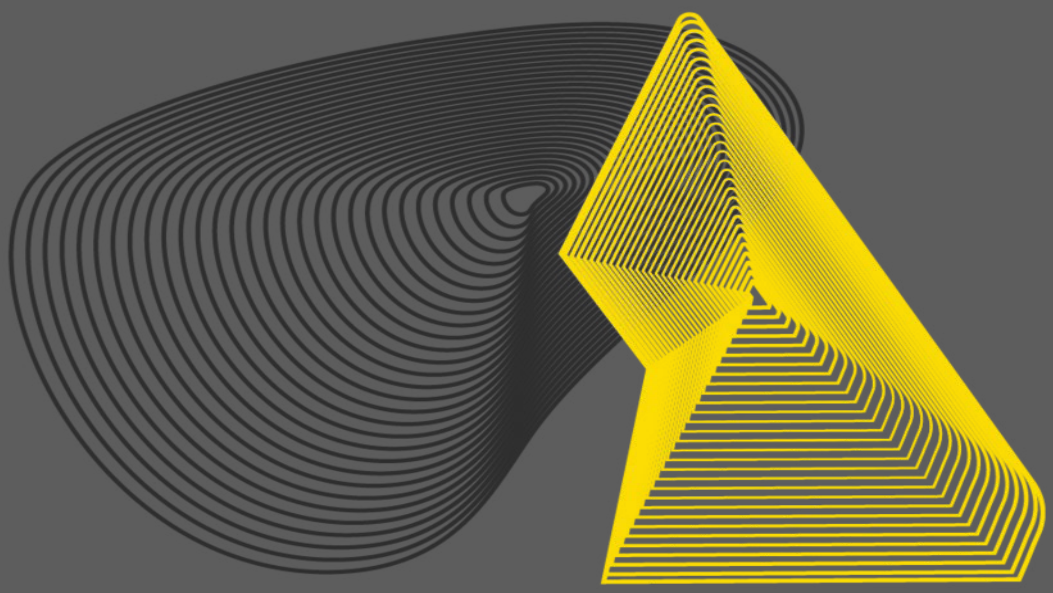
[WEB] **Data Center Watch.** https://www.datacenterwatch.org/

[BLOG] **Ketan Joshi on Big Tech greenwashing.** https://ketanjoshi.co/2025/05/23/data-collection-big-tech-emissions-energy/

and

https://ketanjoshi.co/2025/08/23/big-techs-selective-disclosure-masks-ais-real-climate-impact/

[NEWS] **Data center emissions probably 662% higher than big tech claims. Can it keep up the ruse?** https://www.theguardian.com/technology/2024/sep/15/data-center-gas-emissions-tech



03. We need to talk about bias and discrimination

It is now widely recognised that AI reflects, and often intensifies, forms of discrimination and bias present in society. Given the many points at which human judgments shape the design of AI systems – from the data selected, to the categories defined, to the goals optimised – it should come as no surprise that AI reflects the biases and assumptions embedded in these design choices.

In part this is because of the ways AI is 'trained' to learn from patterns in data. Regardless of whether these patterns accurately reflect social inequalities or distort reality by over- or under-representing certain groups, the result is the same: AI systems often amplify social harms by embedding prejudices into decisions that profoundly affect people's lives.

Due to skewed training data, for example, facial recognition systems often fail to accurately identify non-white faces, leading to wrongful arrests and the amplification of long-standing patterns of racialised policing. In Australia, similar concerns have been raised about child risk prediction tools, which tend to flag First Nations children as being at greater risk. This can lead to unnecessary interventions including the removal of children from families. While Silicon Valley broligarchs obsess about a future in which artificial general intelligence escapes human control, these every-day harms are a current reality for people most impacted by AI assisted decisions.

Public institutions have unique obligations in addressing discrimination and bias in the AI systems they use. For these institutions, the risk of harm can be heightened because of who interacts with the system and the critical nature of the decisions or services being delivered. People who interact most with social services, for example, tend to experience greater social and structural disadvantage, and their interactions with government leave so-called digital traces (that is, traces of personal and administrative data generated through everyday contact with digital services, government agencies, and online platforms). Using these datasets to train AI systems implicates government in reproducing the very inequalities that public institutions work to overcome. Anti-democratic outcomes and forms of discrimination can manifest both in government decisions and also in the decision-making process itself, as AI derives "truths" from algorithmic processing of data, rather than direct engagement with citizens subject to decisions.

Australian government and university policies will often acknowledge issues of bias and discrimination, but social harms are rarely anticipated or proactively addressed at the outset of AI adoption. Instead, corporate AI is often applied opportunistically, its adoption justified by claims of greater efficiency, improved productivity, enhanced accuracy, and reduced costs, with seemingly little regard

for whether it addresses a real need. Australian universities, in particular, have leaned into AI's promissory narratives and growing techno-hype, but have so far avoided difficult conversations about its public obligations to ensure the technology it invests in is neither biased nor discriminatory. Unlike Big Tech companies, public institutions cannot "move fast and break things". Major missteps in universities' use of AI, such as the Ofqual grading scandal in the UK, show how automating decision-making can produce widespread, life-altering consequences for students. We now have enough experience with AI experimentation to foresee the possibility of such outcomes, even if the precise forms of bias and discrimination are difficult to predict. This places a greater responsibility on public institutions to act proactively to prevent harm.

Sources

[BLOG] **"F**k the algorithm"?: What the world can learn from the UK's A-level grading fiasco.** https://blogs.lse.ac.uk/impactofsocialsciences/2020/08/26/fk-the-algorithm-what-the-world-can-learn-from-the-uks-a-level-grading-fiasco/

[NEWS] **Facial recognition remains largely ungoverned - and dangerous - in Minnesota.** https://www.aclu-mn.org/en/news/biased-technology-automated-discrimination-facial-recognition

[NEWS] **'Racially biased' screening tool used in the child protection system places more Indigenous children in custody, critics say.** https://www.abc.net.au/news/2020-07-10/queensland-discriminatory-tool-indigenous-child-protection/12434784

[PAPER] **Who is analysing what? The opportunities, risks and implications of using predictive risk modelling with Indigenous Australians in child protection: A scoping review.** https://onlinelibrary.wiley.com/doi/full/10.1002/ajs4.155

[PAPER] **Can Predictive Algorithms Assist Decision-Making in Social Work with Children and Families?** https://onlinelibrary.wiley.com/doi/10.1002/car.2547

[PAPER] **An evaluation of the racial equity of the actuarial Family risk assessment instrument used in Queensland, Australia.** https://www.sciencedirect.com/science/article/pii/S0190740924004638

Predictive risk assessment in Child and Family services

Algorithmic tools used to predict the risk of child maltreatment promise objectivity, but in Australia and elsewhere they tend to mirror the deep racial and other social inequalities built into the data they are trained on. While discrimination protections prohibit the use of racial categories, many of the variables included in these tools act as stand-ins for race and other supposed risk factors like disability. Risk assessment tools rely on patterns found in administrative data to make their predictions. This means that children get classified as "high risk" not on the basis of individual evidence, but because they share characteristics with groups and families who have had repeated contact with child welfare services and justice agencies in the past. These supposed risk factors are themselves shaped by Australia's history of colonisation, systemic discrimination, and the over-surveillance of Indigenous families and people with disabilities. As a result, Indigenous children are more likely to be labelled high risk, even though studies show that "actuarial tools are superior only in predicting subsequent child protection involvement, which is used as a proxy measure of child maltreatment". Using disability as an indicator of risk can similarly result in disproportionate scrutiny and interventions, without improving child safety.

Further Resources

[VIDEO] **Coded bias at Oxford: panel discussion.** https://www.youtube.com/watch?v=C7AsIInFAck

[WEB] **Gender Shades.** http://gendershades.org/

[VIDEO] **I am not a Number** https://www.sbs.com.au/ondemand/tv-program/i-am-not-a-number/2424128579541

[ARTICLE] **Use of AI could worsen racism and sexism in Australia, human rights commissioner warns**. https://www.theguardian.com/technology/2025/aug/13/ai-artificial-intelligence-racism-sexism-australia-human-rights-commissioner

[REPORT] **Xenophobic machines: Discrimination through unregulated use of algorithms in the Dutch childcare benefits scandal**. https://www.amnesty.org/en/documents/eur35/4686/2021/en/

[ARTICLE] **ChatGPT shown to overlook job applicants with disability.** https://www.disabilitysupportguide.com.au/talking-disability/chatgpt-shown-to-overlook-job-applicants-with-disability

[REPORT] **Inescapable AI The Ways AI Decides How Low-Income People Work, Live, Learn, and Survive.** https://static1.squarespace.com/static/65a1d3be4690143890f61cec/t/673c7170a0d09777066c6e50/1732014450563/ttj-inescapable-ai.pdf

[PAPER] **NSW to review child protection screenng algorithm over concerns about racial bias.** https://www.theguardian.com/australia-news/2024/apr/08/nsw-child-protection-laws-indigenous-children-in-care

[PAPER] **Indigenous Data Sovereignty and Social Work in Australia.** https://www.tandfonline.com/doi/full/10.1080/0312407X.2023.2186256

04. We need to talk about surveillance, control and coercion

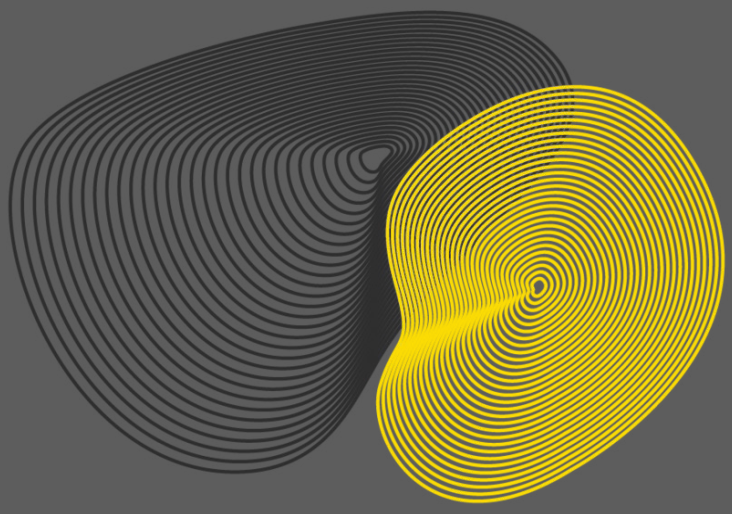
Public institutions are defined largely by their mission to serve the public. They deliver essential services, support, care, cultural and recreational spaces, and serve other functions central to a thriving society. Yet these same institutions rely on mechanisms of oversight and enforcement to manage the communities they serve.

Such mechanisms, while often justified as necessary for accountability and effective service delivery, reveal the dual role of public institutions to both deliver services and monitor populations. Digital technology has long been linked to the state's powers of surveillance. But by embedding monitoring in every online interaction and data-driven decision, AI technologies in many ways sharpen the longstanding tension between care and control.

This tension is especially pronounced in the digital welfare state, where, according to UN Human Rights expert Philip Alston, algorithms and AI systems "automate, predict, identify, surveil, detect, target and punish" as much as they assist in alleviating hardship. And it is within the most punitive corners of the public sector – such as the carceral system – where surveillance technologies are often deployed first, with the least scrutiny and the greatest potential for harm.

First Nations Australians have deep, embodied knowledge of what it means to live with state surveillance. Particularly young First Nations Australians are vastly overrepresented in the institutions, (eg, policing, prisons, child protection systems, and youth detention) where surveillance technologies are deployed in their most concentrated and coercive forms. An expanding network of "care" infrastructure, such as boarding housing, disability group homes, and psychiatric settings, has become yet another arena of surveillance, where technologies are deployed in the name of care in ways that echo the historical – and ongoing – removal of First Nations children under the guise of protection.

Against this backdrop, we are not alone in cautioning that the embedding of AI in public infrastructures could further tip the balance between surveillance and service, control and care, in ways that further compromise their social purpose.



Sources

[REPORT] **UN special rapporteur's report on Extreme poverty and human rights** https://digitallibrary.un.org/record/3834146?v=pdf

[PAPER] **Tensions in digital welfare states: Three perspectives on care and control.** https://journals.sagepub.com/doi/10.1177/14407833241238312

[REPORT] **Digital Futures in Mind: Reflecting on Technological Experiments in Mental Health and Crisis Support.** https://doi.org/10.26188/21113899.v2

[NEWS] **'Racist policing': NSW Police slammed as data reveals more than half of youth targeted by secret blacklist are Indigenous.** https://www.sbs.com.au/nitv/article/racist-policing-nsw-police-slammed-as-data-reveals-more-than-half-of-youth-targeted-by-secret-blacklist-are-indigenous/22p3gpo16

Further Resources

[WEB] **Stop LAPD Spying Coalition.** https://stoplapdspying.org/

[REPORT] **From Surveillance to Empowerment: Advancing the Responsible Use of Technology in Alternatives to Detention.** https://www.unsw.edu.au/content/dam/pdfs/law/kaldor/2025-09-technology-in-alternatives-to-detention.pdf

[PAPER] **Government surveillance and facial recognition in Australia: a human rights analysis of recent developments.** https://www.tandfonline.com/doi/full/10.1080/10383441.2023.2170616

[WEB] **Digital Rights Watch.** https://digitalrightswatch.org.au/

[PODCAST] **Disability Discrimination and Automated Surveillance Technologies.** https://www.csis.org/podcasts/does-not-compute/disability-discrimination-and-automated-surveillance-technologies

[APP] **Copwatch.** https://www.copwatch.org.au/

05. We need to talk about consent and refusal

In the time of #MeToo and investigations into the problem of sexual violence and harassment on campuses, universities in Australia are on notice to embed a culture of consent. Despite many recommendations, concerns remain that Australian universities have not adequately addressed the need.

The speed and scale of AI rollout across public institutions raises questions in regards to the opportunities for consent to, or in, AI on the part of staff, students and communities. While workers are encouraged to take up AI tools within their workplaces, consent processes remain assumed or limited.

Against the Big Tech model of consent as a tickbox on bewildering Terms and Conditions, the growing scholarship and practice of 'consentful tech' draws on the principles of affirmative consent and consent culture as understood in contemporary education on healthy relationships and on preventing sexual violence. The idea is to embed an 'enthusiastic' model of consent into everyday tech. The widely adopted 'FRIES' model of sexual consent (Freely given, Reversible, Informed, Enthusiastic and Specific) provides a ready resource, as does the informed consent required for medical procedures. According to advocates, consentful technologies are applications and spaces in which consent underlies all aspects, from the way they are developed, to how data is stored and accessed, to the way interactions happen between users. Consentful technology is about a holistic and ongoing approach to consent. Consentful culture is intended to spread the ethics of avoiding harm across the institution.

Consent culture requires that a 'yes' is freely given and informed, and also the possibility of 'no'. AI refusal has emerged as an intervention in student protests, manifestos, tech worker activism and in data justice scholarship. Refusal here begins with a 'no' that allows a pause and opens up space for critical conversations. The Feminist Data Manifest-No describes itself as a declaration of refusal and commitment, refusing harmful data regimes and committing to new data futures. Refusal can involve a rejection of the default assumptions and booster narratives around corporate AI, in order to orient discussion and development towards core values of social justice and public good. Refusing the seeming inevitability of Silicon Valley AI might also allow opportunities to embed affirmative, enthusiastic consent in the development, rollout and ongoing use of AI in public institutions.

Ditch the Algorithm UK

In 2020, students in the UK protested to demand the Department for Education 'ditch the algorithm' that had been used to generate test scores for students who never sat their exams due to the pandemic. The predictive algorithm developed by the qualifications regulator Ofqual produced unpredictable and biased results, limiting future pathways and disproportionately impacting students at disadvantaged schools. Within day s, the student protests prompted officials to abandon the automatically generated test scores. Significantly, young people were protesting not simply about data protection and privacy, but rather how data had been actively used to change their futures. The emergence of a "" in the A-level grading fiasco is a model for a more effective means of countering bias and intellectual lock-in in the development of algorithms.

Sources

[COMMENTARY] **'Why 'Ditch the algorithm' is the future of political protest'.** https://www.theguardian.com/commentisfree/2020/aug/19/ditch-the-algorithm-generation-students-a-levels-politics

[PAPER] **Refusal in Data Ethics: Re-Imagining the Code Beneath the Code of Computation in the Carceral State.** https://doi.org/10.17351/ests2022.1233

[ZINE] **Consentful Tech.** https://alliedmedia.org/resources/resource-test

[BLOG] **"F**k the algorithm"?: What the world can learn from the UK's A-level grading fiasco.** https://blogs.lse.ac.uk/impactofsocialsciences/2020/08/26/fk-the-algorithm-what-the-world-can-learn-from-the-uks-a-level-grading-fiasco/

[NEWS] **Why 'Ditch the algorithm' is the future of political protest.** https://www.theguardian.com/commentisfree/2020/aug/19/ditch-the-algorithm-generation-students-a-levels-politics

Further Resources

[CONFERENCE]: **The Refusal Conference at Berkely**. https://afog.berkeley.edu/programs/the-refusal-conference

[MANIFESTO]: **The Feminist Data Manifest-No.** https://www.manifestno.com/

[MANIFESTO]: **AI Decolonial Manyfesto.** https://manyfesto.ai/

[BOOK CHAPTER] **Digital exclusion: a politics of refusal.** https://eprints.lse.ac.uk/103076/1/Gangadharan_digital_exclusion_politics_of_refusal_accepted.pdf

[PAPER] **What can HCI learn from sexual consent? A feminist process of embodied consent for interactions with emerging technologies.** https://dl.acm.org/doi/10.1145/3411764.3445107

[PAPER] **Artificial intelligence and consent: A feminist anti-colonial critique.** https://policyreview.info/articles/analysis/artificial-intelligence-and-consent-feminist-anti-colonial-critique

06. We need to talk about voice and expertise

Unions including the Australian Services Union (ASU) and the National Tertiary Education Union (NTEU) have called for workers to have a voice in how AI is rolled out.

University governance is currently under close scrutiny in Australia. High level inquiries have recommended increased opportunities for staff and student voice in decision-making, including on University Councils. This prompts questions on when, where and how the people at the coalface of corporate AI adoption can have a say and be heard in decision-making and design.

Earlier phases of EdTech adoption have established a trend in which universities routinely use student data to monitor and predict student performance, and yet there is limited engagement with student and staff views, social and ethical issues, policy development, and ethical guidance. A scoping review contends that consideration of ethical issues has failed to keep pace with the development of predictive analytics in the tertiary sector (Braunack-Mayer et al 2020). A purely technical approach to evaluating AI cannot adequately understand the social, political, economic and cultural contexts of technologies, and is hence inadequate to identify harms and ensure accountability. Instead there is an urgent need to centre expertise in the core functions and values of public institutions. In the case of social services, this means incorporating the expertise of people who use welfare services and who know best what they need from systems designed to support them. In the case of universities, this includes expertise in pedagogy, in academic rigour, and in higher education as a public good.

As GenAI is rapidly transforming university teaching and learning, educators and professional associations are sounding the alarm. An Open Letter from educators who refuse to adopt GenAI in education states:

> At its heart, education is a project of guiding learners to exercise their own agency in the world. Through education, learners should be empowered to participate meaningfully in society, industry, and the planet. But in its current form, GenAI is corrosive to the agency of students, educators and professionals.

The risks include automating and replacing human effort, and threats to student wellbeing and learning. The Modern Language Association advocates for

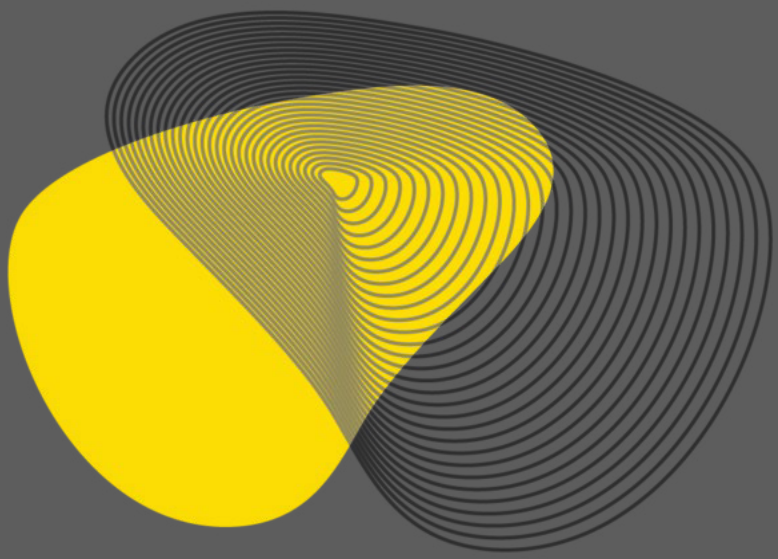
> full faculty and instructional involvement [...] at all levels of decision-making in the selection, procurement, and responsible implementation of instructional technology systems and software, including those incorporating AI. The instructional mission cannot be fulfilled in the absence of faculty expertise, and this expertise extends not only to the "content" of the curriculum but also to the infrastructures and environments—virtual and otherwise—in which learning happens.

An Open Letter to Stop the Uncritical Adoption of AI Technologies in Academia states that corporate AI 'undermines our basic pedagogical values and the principles of scientific integrity. It prevents us from maintaining our standards of independence and transparency. And most concerning, AI use has been shown to hinder learning and deskill critical thought'.

Underpinning these various interventions is an insistence on the importance of voice and agency as Big Tech AI contributes to the deskilling of workers and citizens and the sidelining of vital expertise within public institutions.

Student Voice on a GenAI Course

In 2025 a student-led petition to stop a course that required students to use GenAI garnered 7000 signatures and a motion from the Student Representative Council to suspend the course and investigate the ethical and environmental impacts of generative AI. The petition highlighted concerns in regards to **plagiarism** – GenAI is trained on the work of creators who have not consented, **electricity & water use, climate impact** and **misuse of GenAI** to create misinformation, propaganda and non-consensual pornography, as well as to commit academic misconduct. Student petitioners and the Course Convenor reported that the consultations sparked by the petition served to improve the course by foregrounding the requirement of students taking the course to undertake ethical analysis of the generative AI tools that they are using. In reporting on the consultations, the independent student newspaper said 'For students, this example highlights the importance of speaking up and voicing concerns about university decisions you may not agree with. For the university, this serves as a reminder to engage with students, genuinely listen to their perspectives, and actively take them into account.'



Sources

[REPORT] **AI Governance Needs Sociotechnical Expertise Why the Humanities and Social Sciences Are Critical to Government Efforts.** https://datasociety.net/library/ai-governance-needs-sociotechnical-expertise/

[STATEMENT] **MLA Statement on Educational Technologies and AI Agents.** https://www.mla.org/Resources/Advocacy/Executive-Council-Actions/2025/Statement-on-Educational-Technologies-and-AI-Agents

[OPEN LETTER] **Stop the Uncritical Adoption of AI Technologies in Academia.** https://openletter.earth/open-letter-stop-the-uncritical-adoption-of-ai-technologies-in-academia-b65bba1e?limit=0&utm_source=substack&utm_medium=email

[OPEN LETTER] **An Open Letter from Educators Who Refuse the Call to Adopt GenAI in Education.** https://openletter.earth/an-open-letter-from-educators-who-refuse-the-call-to-adopt-genai-in-education-cb4aee75

[NEWS] **How one petition was able to improve UNSW's AI Art course.** https://www.noiseatunsw.com/how-one-petition-was-able-to-improve-unsws-ai-art-course/

[PETITION] **Stop DART2252.** https://www.change.org/p/stop-dart2252-generative-ai-for-artists-at-unsw/u/33817106

Further Resources

[DISCUSSION PAPER] **NTEU on AI.** https://www.nteu.au/NTEU/Key_Issues/AI.aspx

[COMMENTARY] **AI is hollowing out Higher Education.** https://www.project-syndicate.org/commentary/ai-will-not-save-higher-education-but-may-destroy-it-by-olivia-guest-and-iris-van-rooij-2025-10

[PAPER] **Cut the bullshit: why GenAI systems are neither collaborators nor tutors.** https://www.tandfonline.com/doi/metrics/10.1080/13562517.2025.2497263?scroll=top#metrics-summary

07. We need to talk about participation and co-option

Democratic participation in AI development, like many efforts to involve the public in matters of government, can be challenging. Public institutions, including government agencies, universities, and others, have a responsibility to engage communities in decisions that affect them.

However, AI development is largely an elite domain dominated by global tech firms. Most AI products used by government are purchased "off the shelf" or only lightly adapted for institutional needs. When design decisions rest with third-party developers, affected communities typically have limited opportunities for meaningful input beyond – in the best-case scenario – consultation at the final stage.

Globally, there is growing concern over how public sector AI and automated decision-making might undermine political legitimacy, leading to "algocracy" (rule by algorithm), and further exacerbating the crisis of democracy we are witnessing across the Global North. While public participation in AI design and implementation can enhance trust and ensure systems reflect the needs and priorities of diverse communities, done poorly, it can also be exclusionary or superficial, giving the appearance of engagement without enabling real influence.

There are two different approaches public institutions can take. A "low road" approach treats participation as a formality: tokenistic, poorly resourced, inaccessible, and rarely acted upon. In contrast, a "high road" approach treats public participation as essential to legitimacy and successful outcomes. It involves continuous engagement throughout the entire process, from problem identification and design to implementation and oversight. It prioritises accessibility and provides logistical support such as accessible venues, language services, technology access, and financial assistance for participants, to eliminate barriers to participation. Crucially, it involves listening for and responding to community concerns, ensuring that community 'voice' translates into real influence over decisions. Finally, the high-road we are advocating here requires institutional and regulatory support: public participation needs to be mandated and enforceable, rather than optional or dependent on corporate goodwill.

What will it take to move from voluntary engagement to enforceable mechanisms that guarantee communities a real say in how AI shapes their lives?

Sources

[PAPER] **The Threat of Algocracy: Reality, Resistance and Accommodation.** https://link.springer.com/article/10.1007/s13347-015-0211-1

[ARTICLE] **Government using machine learning to help create draft plans for NDIS participants, documents reveal.** https://www.theguardian.com/australia-news/2025/nov/12/government-using-machine-learning-to-help-create-draft-plans-for-ndis-participants-documents-reveal

[ARTICLE] **How people are assessed for the NDIS is changing. Here's what you need to know.** https://theconversation.com/how-people-are-assessed-for-the-ndis-is-changing-heres-what-you-need-to-know-266255

[WEB] **How we co-design.** https://engage.ndis.gov.au/page/co-design

[PAPER] **Procurement as policy: Administrative process for machine learning.** https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3464203

Automated Decision-Making in the NDIS

In Australia, major reforms are underway to integrate AI into the National Disability Insurance Scheme (NDIS). The agency that oversees the scheme recently took part in a government-wide trial of Microsoft Copilot. An internal FOI-released review of the trial noted modest productivity gains but also staff scepticism about the tool's usefulness, concerns that AI could be used to justify staff cuts, and unease about adopting automated systems following the release of the Robodebt Royal Commission findings.

AI is also being embedded in NDIS support planning. For example, machine-learning models generate draft support plans for NDIS recipients based on their diagnostic and other data. While the final decision rests with an agency representative, the algorithm's recommendations in effect frame the choices available to human decision makers. Research across social services shows that automation bias means staff defer to algorithmic outputs instead of using their own judgment or drawing on the lived experience of end users.

Despite the scale and significance on the lived experience of people with a disability, neither the Copilot trial nor the machine-learning planning tool involved meaningful input from disabled people. The agency's claims of "co-design" ring hollow when participation is limited to superficial consultation; and even more so given the secrecy around how AI actually shapes decisions.

Further Resources

[REPORT] **Democratising AI: Principles for Meaningful Public Participation.** https://datasociety.net/library/democratizing-ai-principles-for-meaningful-public-participation/

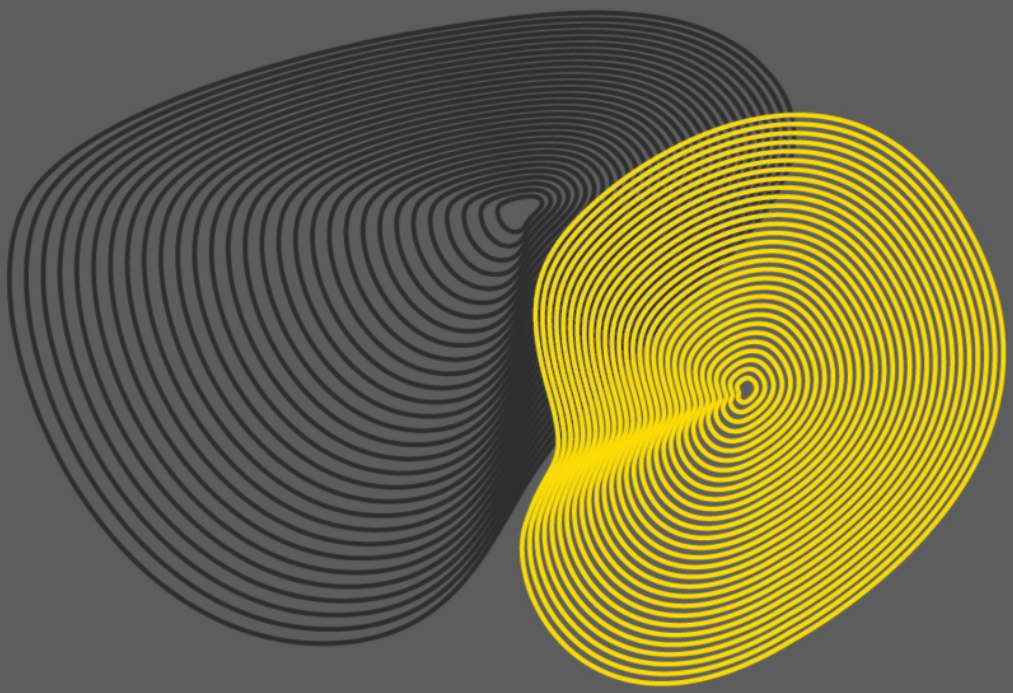
[PAPER] **Data governance and citizen participation in the digital welfare state.** https://www.cambridge.org/core/ournals/data-and-policy/article/data-governance-and-citizen-participation-in-the-digital-welfare-state/CCF2E1914C7E2D4593D550C6BC9E4C70

[REPORT] **Civic Participation in the Datafied Society: Towards Democratic Auditing?** https://datajusticelab.org/wp-content/uploads/2022/08/CivicParticipation_DataJusticeLab_Report2022.pdf

[PAPER] **Decoding the algorithmic operations of Australia's National Disability Insurance Scheme.** https://onlinelibrary.wiley.com/doi/10.1002/ajs4.342

[VIDEO] **Databite No. 155 Democratizing AI: Principles for Public Meaningful Participation.** https://www.youtube.com/watch?v=42sop-UJVAg

[ARTICLE] **Did we learn nothing from robodebt? NDIS automation will put vulnerable lives at the mercy of machines.** https://www.theguardian.com/commentisfree/2025/dec/04/ndis-news-support-plans-automated-robodebt-national-disability-insurance-scheme



08. We need to talk about alternatives

The scale and speed of AI adoption across public institutions makes the challenge of imagining and working for alternatives equally urgent. Given the environmental, social, political and public interest implications of corporate AI, we need to talk about different futures and how to get there.

It is vital to work collectively for community involvement in decision-making on whether or not to adopt AI, and if so how. Here we offer a few brief sketches of organisations and projects working on alternative visions for AI. Beyond questions of regulating the market, here are suggestions for building from values of First Nations sovereignties, social justice, cooperatives and public interest. This is far from a comprehensive account, rather these examples are presented as resources that might spark critical conversations on the values and practicalities of AI beyond the Big Tech model.

First Nations Knowledges highlight the colonial dynamics of exploitation and extraction inherent in Silicon Valley AI. Indigenous Data Sovereignty asserts an alternative. Maiam nayri Wingara Indigenous Data Sovereignty principles include: control of data creation, development, stewardship, analysis, dissemination and infrastructure; data that empowers sustainable self-determination and effective self-governance; data structures that are accountable to Indigenous peoples and First Nations; and data that are protective and respect Indigenous individual and collective interests.

Te Hiku Media

Te Reo Irirangi o Te Hiku o te Ika (Te Hiku Media), a First Nations media organisation in Aotearoa New Zealand, has developed its own AI for Māori language translation while keeping control of the community's data. Te Hiku Media began in 1990 as a community radio station, grounded in Māori sovereignty and the Māori language revitalisation movement. It brings the principles of Indigenous sovereignty and community control to the development of AI, through principles of mana Māori motuhake (Māori autonomy), tino rangatiratanga (self-determination), and manaakitanga (care), and kaitiakitanga (guardianship). Careful attention is paid to what data is collected, how data is collected, who has access to the data and what the data is used for. The Kaitiakitanga License developed by Te Hiku Media aims to prevent harmful uses of data, such as for surveillance.

Te Hiku Media also speaks to the importance of small scale, bespoke, locally developed AI as an alternative to corporate AI built for scale and ubiquity. The tradition of community owned and controlled media, including First Nations media, is embedded in close relationships between specific communities and the means of communication. Community participation and oversight is central to the mission. A more recent model of collective action and organising is Cooperative AI, emerging from the Platform Cooperativism movement and prioritising democratic decision-making and shared ownership by workers and users.

The movement for Public Interest Technology focuses on values of accountability, transparency, public good and diversity. It addresses public institutions as potential leaders or innovators in developing for purpose, publicly owned and operated, not for profit technologies. The aims include auditable, transparent AI infrastructures that can be examined publicly and operate for public good.

Sources

[COMMUNIQUE] **Maiam nayri Wingara Indigenous Data Sovereignty Principles.** https://www.maiamnayriwingara.org/mnw-principles

[ARTICLE] **Kaitiaki: closing the door on open Indigenous data.** https://doi.org/10.1007/s00799-025-00410-2

[ESSAY] **The Case for Digital Public Infrastructure.** https://technologist.mit.edu/the-case-for-digital-public-infrastructure/

[ESSAY] **Cooperative AI?** https://ia801506.us.archive.org/31/items/cooperative-ai/CooperativeAI.pdf

[BLOG] **Reflections on the Cooperative AI Conference.** https://platform.coop/blog/when-workers-take-the-wheel-a-cooperative-ai-conference-reflection/

Further Resources

[WEB] **Te Hiku Media.** https://tehiku.nz/

[WEB] **Platform Cooperativism.** https://platform.coop/

[ARTICLE] **AI For Life: How Sovereign Wiradyuri Ways of Knowing Can Transform Technology for Good.** https://theconversation.com/ai-for-life-how-sovereign-wiradyuri-ways-of-knowing-can-transform-technology-for-good-243270

[WEB] **Australia Institute Centre for Responsible Technology.** https://www.centreforresponsibletechnology.org.au/

[WEB] **The Global Technology for Social Justice (GloTech) Lab.** https://www.umass.edu/gateway/research/stories/internet-technology-ai/global-technology-social-justice

[WEB] **Ford Foundation Public Interest Tech.** https://www.fordfoundation.org/news-and-stories/big-ideas/public-interest-technology/

[PAPER] **Sovereignty-as-a-service: How big tech companies co-opt and redefines digital sovereignty.** https://journals.sagepub.com/doi/10.1177/01634437251395003

[WEB] **Maiam nayri Wingara.** https://www.maiamnayriwingara.org/

Key Questions for AI in Public Institutions

To end, we offer a series of questions that might provoke more critical conversations on AI in public institutions, including universities and social services.

- Do the benefits outweigh the harms?
- Does the business model match our core values?
- Does this fit with our commitment to the UN Sustainable Development Goals?
- Is this consentful?
- Who participates in the design and decision-making?
- Is it possible to say no?
- How could the public institution become a leader in built-for-purpose, local, public digital infrastructures?
- Who owns the infrastructure? Who benefits? Who is harmed? Who decides?
- What are the hidden costs? Is there a better way?
- Is this worth it?