



# antistatic

## AI and Other Stories

A collection of public writing on and around AI  
Anna and Kelly Pendergrast, 2019-2023

# antistatic

## **About Antistatic**

Antistatic is a research and communications consultancy based in Oakland California and Wellington New Zealand, founded by Anna and Kelly Pendergrast. We work with clients and collaborators in government, nonprofits, academia, and business to investigate and communicate about complex and vital contemporary issues. We focus on areas where complex systems intersect with the lives and wellbeing of people, and in particular digital technology, material culture, and the environment. [antistaticpartners.com](https://antistaticpartners.com).

## **Acknowledgements**

We have massive gratitude for the wonderful editors we've worked with over the years, who have helped us shape our thinking, and the publishers who commissioned and stewarded our work. We are also forever thankful to the scholars, activists, and thinkers we've learned so much from.

## **Cover image**

Philipp Schmitt & AT&T Laboratories Cambridge, Data flock (faces). CC-BY 4.0. This image was sourced from [Better Images of AI](#), a non-profit collaboration dedicated to sourcing and sharing images that support wider public comprehension of AI technologies, applications and governance.

**Antistatic, October 2023.**

## Contents

---

Introduction .....	4
Building the World for Everyone.....	7
AI Maintenance as Care, Respect, and Guardianship .....	12
The Next Big Cheap .....	17
Can New Zealanders Trust a Digital Government? .....	25
Disassembly Required.....	31
I caught a driverless taxi and it was terrifying .....	38

## Introduction

---

As we write this, in October 2023, AI is once again in the news. The rapid development and dissemination of generative AI tools — the image generators and large language models that allow people to prompt photorealistic images or swathes of text with ease — has everyone from college administrators to management consultants scrambling to respond. The thought that in mere seconds you can produce paragraphs of text ready to be pasted into essays, newsletters, emails or chatbox speech bubbles feels like a game changer. Whether those words are accurate, or even any *good*, is another thing altogether. Regardless, we are told that the power of these tools will change the way people work and communicate.

The technology is undeniably impressive, with some capabilities we genuinely haven't seen before. Many of the issues and conflicts raised by these tools, though, are not new. While some things *are* different when it comes to generative AI, we are still talking about tools that rely on huge troves of data, can be applied in myriad ways, and have outcomes which are dependent not only on the tools themselves but how they are built into a wider system. Scholars, practitioners, and activists have talked for years (decades!) about labor exploitation, biased data, erosion of worker rights, and corporate control in relation to machine learning and data-driven technology — these issues are further heightened with the wide adoption of generative AI tools.

In the US, where Kelly lives, the recent Writers Guild of America strike and the ongoing Screen Actors Guild strike were both prompted in part due to very real concerns that employers and studios would attempt to use AI tools to undercut wages and erode working conditions. The WGA won the first major union contract to include a real, enforceable standard that governs the use of AI. In New Zealand, where Anna lives, both government and industry have been hustling to figure out

what a “domestic” AI ecosystem could look like, and how AI might be governed, incorporated, and regulated in a small nation that is more accustomed to being a technology receiver than a technology producer. Do you try and stem the tide, do you say “bring it on”, or do you try to build tools unique to the country’s needs and culture?

In this context, we considered whether we have anything new to add to the existing pile of writing on generative AI. Guidance for decision makers on how to understand the capabilities and challenges of these new tools? Research on public concerns and comfort around the use of generative AI by the bureaucracies they engage with? A critical essay on images, ideology, and the archive?

In the end, we decided that perhaps the most useful thing we could do to help build understanding is to bring together our existing writing on AI and adjacent issues, and share it with anyone interested. Plus, we’re tired, and this seemed like an easier option.

So: here you go. This compendium brings together essays and articles we’ve written over the past few years that may be relevant to anyone grappling with generative AI and how these tools might affect our lives now and in the future.

“Building the World for Everyone”, written as the introduction to an anthology book on technology and equity in Aotearoa New Zealand that we edited in 2022, sets out a vision for a more equitable world and makes the case that it is not just technology that requires our attention, but the broader institutional and infrastructural systems in which technology is embedded and mobilized.

“AI Maintenance as Care, Respect, and Guardianship”, our 2021 contribution to *AI Now’s New Lexicon of AI*, posits that maintenance is a vital but under-discussed (and underinvested) consideration with so-called-AI systems. Maintainers do much of the fundamental labor of AI like tagging and cleaning data sets, but their labor isn’t adequately compensated or appreciated, and long-term planning for data maintenance and stewardship will be essential to the security and longevity of any AI system.

“The Next Big Cheap”, Kelly’s 2019 essay for *Real Life* magazine, was written before the current explosion of generative AI, but its description of the “cheap” data that underpins today’s tech economy continues to resonate. Many tech builders talk about innovation as a “frontier” and data as a resource to exploit and expropriate, in a continuation of historic patterns of extraction and enclosure.

“Can New Zealanders Trust a Digital Government”, a 2022 article based on our work with the Digital Council for Aotearoa New Zealand, considers how trust is built or eroded between government agencies and the public. Communities are where

the impacts of government algorithmic decision making are felt most keenly, and listening to the views of community members about when and why they consider algorithmic decision making appropriate is a fundamental first step in developing more equitable systems.

“Disassembly Required”, Kelly’s 2020 essay also for *Real Life* magazine, argues that we anthropomorphize robots at our peril. Robots (and AI, for that matter) are not autonomous entities that will one day be our coworkers or friends — they are machines designed to perform tasks, and they serve the interests of the companies and organizations that make them or deploy them. We are not against robots or against AI, but we insist that they be understood for what they are and who they serve.

Finally, “I caught a driverless taxi and it was terrifying” is Anna’s 2023 first person account of a rocky ride in a newly-street legal autonomous vehicle. Written for New Zealand’s *The Spinoff*, the essay encapsulates the exhilaration and banality, terror and hilarity that comes with an on-the-ground experience of cutting edge — but perhaps not entirely street-ready — AI vehicular technology.

As we re-read these texts and tracked the through-lines, it was clear there are a few topics we just can’t stop harping on about. The fact that discussing AI (or any technology) in isolation does nothing useful — these tools are embedded in broader systems, processes, and power relations, and the wider context needs to be considered and unpacked. The need to examine who benefits from the development and use of a new technology — and who is harmed. The desire to consider other models for understanding technology and data outside the colonial, western paradigm — we refer in multiple places to the work of Māori data sovereignty scholars and research collectives, whose writing and activism has shaped our thinking. And finally, the demand that technology and sociotechnical systems should (and can) serve the needs of people and communities over those of corporations and shareholders.

Thanks for reading. We hope you find this useful. Please get in touch at [hello@antistaticpartners.com](mailto:hello@antistaticpartners.com) with any questions or comments, or just to chat.

Your friends,  
Kelly and Anna

# Building the World for Everyone

Kelly Pendergrast and Anna Pendergrast

---

Introduction to [the edited volume](#) “More Zeros and Ones: Digital Technology, Maintenance and Ethics in Aotearoa New Zealand” published by [Bridget Williams Books](#), 2022. Edited by Anna Pendergrast and Kelly Pendergrast.

---

We want to live in a world that is built for everyone. From our cities, homes and infrastructures to our political systems, technologies and services, we want human-built environments and systems to be inclusive of everyone. We also want them to reflect our particular place in the world and allow for multiple worldviews and ways of thinking.

Sharing this vision of an equitable, inclusive world might seem like an unusual way to open a book about digital and data-driven technologies. However, digital technologies and the services and tools that are built on top of them are woven throughout our lives. To build a world for everyone includes ensuring that digital systems are equitable, inclusive and empowering.

Industry commentators and the government alike want New Zealand to be an innovative, future-focused country that builds cutting-edge technology and thrives in a digitally connected global environment. Local companies like Rocket Lab and Xero are lauded for growing successful digital businesses with a strong international profile, and both public and private funders put considerable investment into growing New Zealand's tech industry.

However, the future of New Zealand's digital ecosystem can't just be about rockets and exponential growth. To build a world that sustains both future generations and

the land that nurtures us, we also have to care for the present. This means doing the tough and unglamorous work of identifying where existing systems aren't working for – or are actively working against – the people they claim to serve, and then fixing these systems and changing the way things are built. This might not sound a lot like innovation (and, to be fair, it's not as sexy or exciting as building rockets), but even the most spectacular of new inventions are built on the foundations of the past. Every time people make something 'new' they're using existing tools and building off existing frameworks and technologies. If those tools and frameworks are biased, harmful, or designed to serve the needs of only a few, then the new things that get built will replicate the same old problems.

Historic inequities can be replicated during the design and development of digital technologies and services, which are often designed 'for the 80 per cent' – the centre of the bell curve that includes most but not all people. The people outside the 80 per cent will differ in different contexts, but might include blind or low vision people who use a screen reader to engage with digital content, Māori people for whom data is a taonga, or older people who did not grow up with digital technology. By consistently designing for the middle of the bell curve, the 20 per cent of folks outside the middle have an uphill battle to get products and services that work for them, and decisions made for the majority often inadvertently reinforce systemic oppression.

We see examples of technology built using biased frameworks or designed for the 80 per cent – whether deliberately or inadvertently – all around us. For example, more and more businesses and organisations are providing vital services (from banking to vaccine passes) in a digital-first way, with in-person or other channels being secondary or an afterthought. This approach assumes a certain type of user with an up-to-date smartphone and relevant digital skills, and accepts the risk that a percentage of people will be excluded. Or take the example of cloud services, which is increasingly how data about us is being stored. As more and more data is handed off to international cloud storage giants, Māori have limited options to retain sovereignty over their information, which can become subject to international law and out of local control. Finally, the social platforms so many people rely on to connect with each other are designed in such a way that mis- and disinformation can spread at speed and scale, leading to real-world harm. We only have to look at the disinformation about Covid-19 vaccinations that has been spread widely over the past two years to understand the potential negative consequences for communities and whānau.

There are some really optimistic narratives circulating around the future of technology: that our problems can be innovated away with technical systems, and that increased decentralisation, transparency and automation can help us achieve prosperity and equity for all if we just deploy them for good. It's easy to buy in to this version of a technological future – optimism is exciting, and these



narratives make a fix look simple. If computer scientists can be aware of potential bias, maybe they can just code it away. If hiring managers acknowledge that certain demographics dominate the tech industry, they can encourage more women to apply for STEM jobs. If big platforms can just tweak a few content moderation settings, maybe they can do away with mis- and disinformation permanently.

However, surface-level adjustments won't be enough on their own. There are some big underlying issues shaping the way digital technology is funded, designed and deployed that need to be addressed in order to see any real systemic change. Just like you can't stop internal bleeding with a Band-Aid, you can't make small adjustments to the products and services of a profit-focused company or institutionally racist organisation and expect them to suddenly centre the wellbeing of all users. When problems are deeply ingrained, our solutions also need to go deeper: starting with the way that digital systems and products are designed.

To begin to do things better, we argue that tech needs to be shifted away from the centre of the story. The optimistic narratives that promote technology as the solution for all problems often also treat digital technology as though it's a force of nature – something to be harnessed, like wind, or the sun, or flowing water, in order to lead towards a certain kind of inevitable progress. But in reality, technology is made by people and is the product of human systems. No technology is an inevitability. It is only through the actions of many people over time that software, hardware and infrastructures are built.

The good news is that things can be different. If we acknowledge that no technological development is inevitable, it means we don't just have to cross our fingers and hope that things will work out okay. Instead, we can work together to change the systems that are causing harm, and develop new practices. The bad news is that changing systems is hard work. Recentring people's needs and aspirations over the demands of profit and convenience isn't easy to do. It will require people with power to change their ways of doing things, make difficult decisions, and listen to others. And it will require courageous advocacy and action from everyday people. But if we want a more just and equitable future for everyone in this country, it's essential that we get started.

Aotearoa New Zealand has a unique context and culture, and our digital systems need to reflect that. Fortunately, we already have many of the guides and levers we need to ensure our digital future upholds our specific values and uplifts our people and culture. Aotearoa's digital ecosystems need to embed Te Tiriti o Waitangi – a founding constitutional document for Aotearoa. Our digital systems and policies also need to recognise human rights, for example by adhering to the Privacy Act 2020 and giving effect to the United Nations Declarations on the Rights of Indigenous Peoples and Persons with Disabilities.

Along with these existing commitments and guides, people in the tech industry, the government and communities will also need to develop a broad range of new ways of building, deploying and maintaining digital systems and infrastructure. Even within Aotearoa New Zealand, we know that one-size solutions don't fit all, and that everyone experiences the internet and digital technology differently. Now it's time to put this knowledge into practice.

In mid-2020, Bridget Williams Books published *Shouting Zeros and Ones: Digital Technology, Ethics and Policy in New Zealand*, edited by Andrew Chen. We are proud to have a chapter about digital inclusion in that anthology, which looks at a range of important digital issues in our specific Aotearoa New Zealand context.

As we revisited *Shouting Zeros and Ones* to plan this new volume, one thread we saw running through many of the chapters was an investigation of the potential harms and challenges of living in an increasingly digital world, with a particular focus on digital content and the data that drives many digital systems. Some authors surveyed the negative impacts of mis- and disinformation, hate speech and online fascism. Others looked at the implications of using data and algorithms to make decisions about people's lives, examined the current state of the Stats NZ Integrated Data Infrastructure, and introduced the field of Māori Data Sovereignty. The authors made suggestions about what could be done to reduce harm and implement positive change, building a compelling roadmap for how to start working towards a digital world that ultimately benefits us all.

We are excited to take on the role of editors for this second volume in the series, and to present you with more insights into digital technologies and the myriad ways they intersect with people's lives in Aotearoa New Zealand. In this book we shift our focus away slightly from digital content and data, to examine the wider socio-technical systems that digital technologies sit within. This includes looking at the ways people design, build, maintain and ultimately decommission or dispose of digital technologies, services and infrastructures. All the authors in this new book are searching for ways to do things better – for individuals and whānau, for communities, for society, and for the environment. We hope that the essays in this book – from Karaitiana Taiuru writing about how Te Tiriti o Waitangi can be embedded into tech and data projects to Nessa Lynch writing about more-ethical deployment of facial recognition technologies – provide a series of case studies and strategies that help us to imagine a better world.

The chapters in this book offer a range of approaches, with different scales and proposed intervention points. Some authors talk about working with existing technologies and organisations, and figuring out how to make them more equitable. Others demand a different distribution of power or control, arguing that people and whānau know what they need and are best placed to solve their own problems, if

only they are empowered with the resources to do so. Cumulatively, the chapters make it clear that working towards a more equitable internet and digital ecosystem won't just be a job for the government and corporations. Multiple worldviews and ways of doing things can – and must – be given space.

For us, thinking about the future of the internet and digital technologies doesn't mean looking for distant 'signals' about what the next big thing might be, so investors can accelerate growth and regulators can be ready to deal with any negative impacts that might come with it. While this approach centres exciting innovation and the emergence of new ideas, it can also tend towards framing technology as something that just happens to society, where all any of us can do is be ready for what comes. This glosses over the agency and responsibility that people – including people here in Aotearoa New Zealand – have in building the world.

We propose an alternative approach to thinking about the future of technology. There are three questions we like to ask when thinking about what can be done now to shape the internet and the equitable digital world we want. It's our way of thinking about the future, based on the reality of the present. We suggest that, as you read this book, you ask yourself:

- What is working well now, and how do we maintain it to serve us in the future?
- What is harming us now, and how do we dismantle or change it?
- What is missing now, and what can we build to fill the gap?

All the chapters in this book provide an insight into at least one of these questions. Together, we hope the authors' recommendations and observations will serve as a guide towards an innovative, inclusive, equitable and ambitious future for Aotearoa New Zealand.

# AI Maintenance as Care, Respect, and Guardianship

Anna Pendergrast and Kelly Pendergrast

---

*This 2022 essay [was written as](#) part of AI Now's "AI Lexicon" project, a call to generate alternate narratives, positionalities, and understandings to the better known and widely circulated ways of talking about AI.*

---

To bring an AI system into the world is also to bring about a responsibility for its care. Maintenance is "both absolutely necessary and usually neglected,"<sup>1</sup> write historians of technology Andrew Russell and Lee Vinsel, who co-founded [the Maintainers network](#) of maintenance scholars and practitioners. Maintenance is forgettable (until it isn't), especially compared with the more spectacular and photogenic phases of a product or system's lifecycle: designing, inventing, building, and even repair. Without maintenance, ropes fray, data decay, and things fall apart.

The consequences of inattention to maintenance can be immediate, but they are also generational.<sup>2</sup> California's 2018 Camp Fire, which killed 85 people, was likely caused by a faulty transmission line. Utility company PG&E was subsequently found to have failed to properly maintain and inspect transmission lines for many years—including the tower which sparked the Camp Fire.<sup>3</sup> The consequences of neglected maintenance are not neutral. They are often unevenly distributed

---

1. Russell, A & Vinsel, L. 2020. *The Innovation Delusion: How Our Obsession with the New Has Disrupted the Work That Matters Most*.

2. Mattern, S. 2018. "Maintenance and Care." *Places Journal*, November 2018. <https://placesjournal.org/article/maintenance-and-care/>

3. Gold, R and Blunt, K. 2019. "PG&E Had Systemic Problems With Power Line Maintenance, California Probe Finds". *The Wall Street Journal*, 19 December 2019. <https://www.wsj.com/articles/pg-e-had-systemic-problems-with-power-line-maintenance-california-probe-finds-11575338873>

along lines of privilege and disenfranchisement.<sup>4</sup> Many of the residents killed and displaced by the Camp Fire were already facing housing insecurity, and some had previously been displaced from other more expensive areas of California.<sup>5</sup>

In the public and political discourse around the ethics of AI, conversations tend to focus on the design, development, and near-term effects of AI, or their correction after catastrophic failures, rather than on their ongoing maintenance or upkeep over time. As it is commonly understood, maintenance is a process that promotes the continuity of physical and digital products, services, or infrastructures in order to ensure they continue to operate as designed. We propose maintenance as an under-researched and under-resourced area of study in AI. To think about AI through the lens of maintenance practices is one way to acknowledge the long life of technological systems and their impacts on people and the environment.

Social and digital infrastructure require maintenance just as keenly as power pylons and leaky roofs. AI systems are complex sociotechnical assemblages comprising data sets, algorithms and models, human labor, and wider institutional structures. They need to be monitored and adjusted to ensure accurate or desirable outputs, code bases need to be kept up to date, bugs need to be fixed, and data sets managed. Without close attention, AI systems can produce damaging outputs, especially those that ‘learn’ from dynamic data sets. Databases are the backbone of many AI technologies, but unless they are updated regularly, they represent only a snapshot of the world at a given time. For example, in order for AI-enabled autonomous vehicles to function smoothly without causing accidents, mapping software needs to be highly accurate and up-to-date, requiring datasets, and the physical roading infrastructure that they represent, to be maintained consistently.<sup>6</sup> Even for existing mobility platforms that claim to use AI to make real-time decisions about routes and pricing, there is an army of subcontractors who regularly clean and update map data as well as volunteers who carefully collect open [map data in the first place](#). In the context of AI systems, with their complex inputs and iterative development, maintenance overlaps heavily with design, testing, and repair. Still, an approach to AI that centers incremental upkeep and ongoing care would represent a departure from business as usual, and potentially shape how systems are understood and built.

---

4. The Information Maintainers. Olson, D., Meyerson, J., Parsons, M., Castro, J., Lassere, M., Wright, D., Acker, A. 2019. *Information Maintenance as a Practice of Care*. <https://doi.org/10.5281/zenodo.3236409>

5. “More Than 1,000 Families Still Searching For Homes 6 Months After The Camp Fire” NPR, May 8, 2019. <https://www.npr.org/2019/05/08/721057281/more-than-1-000-families-still-searching-for-homes-6-months-after-the-camp-fire>

6. McKinsey Institute. 2019. A new look at autonomous-vehicle infrastructure. <https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/a-new-look-at-autonomous-vehicle-infrastructure>

Maintenance is not the inverse of innovation and creation: it is the necessary complement. “Maintenance is the key to ensuring that the benefits of technology are felt in their full depth and breadth” write Vinsel and Russell.<sup>7</sup> To work towards AI accountability, transparency, fairness, or human rights compliance—to ensure technology is beneficial for the many rather than just the few—is to engage with AI maintenance practices. This might include monitoring systems and making small changes over time to ensure equitable outcomes are achieved, or maintaining up-to-date documentation that allows for audits and accountability. As companies, NGOs, and governments work towards operationalizing the AI ethics principles that have proliferated over the past half-decade, acts of maintenance are key components of this proposed work, but rarely described as such. While discussions of “ethics” risk vagueness, abstraction, and impracticality, a maintenance-centric approach to AI might provide a key mechanism through which some aspirations of ethical AI are operationalized—both for the people affected by AI systems, but also those who perform the “hidden” work of AI.<sup>8</sup>

The essential labor of maintaining and caring for AI and its systems is as broad as the systems themselves, and the people that perform these acts are more varied and dispersed than most AI narratives allow for. They include the thousands of low-paid data workers performing digital piecework for a few cents a task,<sup>9</sup> lauded computer scientists responsible for building and training models, as well as the policy analysts, service designers, and bureaucrats who make decisions about when AI is deployed and why. The majority of AI workers are low paid and often erased from AI narratives. Amazon’s Mechanical Turk program, the engine behind many large data sets and AI projects, relies on freelancers who perform vital digital piecework including data labeling, image classification, rating the toxicity of tweets, recording voice samples, and completing social science surveys.<sup>10</sup> This work to build, clean, and filter data and information is necessary all the way through the life cycle of an AI project, from design to build to maintenance and repair, and many

---

7. Russell, A & Vinsell, L. 2020. *The Innovation Delusion: How Our Obsession with the New Has Disrupted the Work That Matters Most*.

8. AI Ethics Impact Group. 2020. From Principles to Practice: An interdisciplinary framework to operationalise AI ethics. <https://www.ai-ethics-impact.org/resource/blob/1961130/c6db9894ee73aefa489d6249f5ee2b9f/aieig—report—download-hb-data.pdf>

9. For more discussion and framing of data workers’ essential role in developing and maintaining AI systems, see Nithya Sambasivan, Shivani Kapania, Hannah Highfill, Diana Akrong, Praveen Paritosh, Lora Aroyo. 2021. “Everyone wants to do the model work, not the data work:” Data Cascades in High-Stakes AI. InCHI Conference on Human Factors in Computing Systems (CHI ’21), May 8–13, 2021, Yokohama, Japan. ACM, New York, NY, USA. For additional information on wages from a survey of Amazon Mechanical Turk workers, see also: <https://www.nytimes.com/interactive/2019/11/15/nyregion/amazon-mechanical-turk.html>

10. Stanley, S. 2021. “The Workers Perspective.” TWC Newsletter Issue 5: Living in the Hidden Realm of AI. <https://news.techworkerscoalition.org/2021/03/09/issue-5/>. See also David Martin, Benjamin V Hanrahan, Jacki O’Neill, and Neha Gupta. 2014. “Being a turker.” In Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing. 224–235.

of these systems would fail if maintenance is ignored. Any accounting of an AI system's impacts or ethics should consider the system's maintainers, their working conditions, and their agency.

For guidance on how maintenance can help design, make and care for the world—including the world of AI—we look to existing frameworks from Indigenous, feminist, and other radical traditions, and here we draw specifically on Māori frameworks as an example. In doing so, we recognise the position from which we write: with privilege as Pākehā (European New Zealanders); and with responsibilities as Tangata Tiriti (roughly translated as people who have the right to live in New Zealand as the result of Te Tiriti o Waitangi/The Treaty of Waitangi).<sup>11</sup> It is in this context that we were introduced to, and later came to advocate for, Māori data sovereignty principles and broader te ao Māori perspectives. Our work, and our journey to become better Tangata Tiriti, has been greatly informed by Māori scholars and activists we have worked alongside and look up to, and we acknowledge their mahi (work).<sup>12</sup>

The growing field of Māori data sovereignty, and Indigenous data sovereignty more broadly, upends extractive, colonial understandings of data and AI and insists that AI systems not be seen as abstract assemblages of data and mathematical models. “Mainstream discussions of algorithms represent them as somewhat abstract entities, but this representation does not hold from a te Ao Māori point of view”<sup>13</sup> write researchers at Aotearoa New Zealand’s Te Kotahi Research Institute. Instead, Te Kotahi and others argue that AI should be viewed as living systems inexorably connected to the people whose data are used in the systems, and the physical sources of energy, hardware, and space that ground them. The data used in these systems, especially personal data about people, have mauri (lifeforce) and whakapapa (genealogy), and cannot be severed from the people to whom they relate. This groundedness and relationality invokes the need for ongoing care, respect, and guardianship of data and AI models. “These models then ought to be viewed as both living—requiring kaitiakitanga [guardianship]—and relational—implicated in a network of obligations and relationships that need to be appropriately maintained.”<sup>14</sup>

---

11. This blog post by Tina Ngata has been really helpful to us in terms of the kinds of actions that make good Tangata Tiriti: <https://tinangata.com/2020/12/20/whats-required-from-tangata-tiriti/>

12. We would like to particularly acknowledge Chris Cormack, Donna Cormack, Amber Craig, Maui Hudson, Tahu Kukutai, Tina Ngata, Karaitiana Taiuru, Ari Thompson, Kiri West and Daniel Wilson, whose mahi (work) has guided us in writing this essay.

13. Hudson, M., Thompson, A., West, K. & Wilson, D. 2020. Māori perspectives on Trust and Automated Decision-Making. New Zealand: Te Kotahi Research Institute. Page 11. <https://digitalcouncil.govt.nz/advice/reports/towards-trustworthy-and-trusted-automated-decision-making-in-aotearoa/>

14. *ibid.*

*“You don’t leave a carving alone in the rain”<sup>15</sup>*

The kaitiaki (guardian) of a system must ensure it is protected, stewarded, and maintained so it honors the lives it intersects with, and the materials and history it represents. This applies equally to AI models and precious cultural artifacts.

On its surface, maintenance may seem biased towards the seamless continuity of existing systems—the acceptance of a status quo that enables discrimination and unjust outcomes. Why maintain systems that do harm? Critical maintenance scholars note that “maintenance is not the opposite of change, however, and its primary aim and value is not to uphold stasis.”<sup>16</sup> Maintenance provides space for reflection and reimagining, and an opportunity to intervene in a system even as you ensure its continued function. This might even include the decommissioning of AI systems or data, as the *Feminist Data Manifesto-no* suggests:<sup>17</sup> “We commit to... preparing bodies or corpses of data to be laid to rest when they are not being used in service to the people about whom they were created.” Maintainers, with their intimate relationship to machines and systems, are often best positioned to critically assess how things work and offer suggestions for how they could be tweaked to work more efficiently, equitably, or sustainably.

By bringing maintenance into the AI vernacular, we open space to better consider the people affected by and the people who maintain AI systems. We welcome AI maintenance at its most radical, and resist limited ideas of maintenance that only allow for the rote continuity of systems that work for some and oppress others. How this looks can take many forms. From [Write the Docs](#), a global community producing open-source documentation to ensure the ongoing maintainability of systems (and, in doing so, build the connections to maintain themselves), to Te Hiku media’s Māori language tools (built using crowd-sourced labelled datasets and governed by the organisation’s own kaitiakitanga licence to ensure that benefits gained from the language data remain with Māori<sup>18</sup>), a broad swath of projects underway provide inspiration and jumping-off points. Our AI maintenance is a critical intervention, an ongoing attention to the experience of workers and the lived experience of those affected by systems, and about offering a space for other worldviews.

---

15. Participant from expert wānanga on trust and automated decision-making, quoted in Hudson, M., Thompson, A., West, K. & Wilson, D. 2020. *Māori perspectives on Trust and Automated Decision-Making*. New Zealand: Te Kotahi Research Institute. Page 11. <https://digitalcouncil.govt.nz/advice/reports/towards-trustworthy-and-trusted-automated-decision-making-in-aotearoa/>

16. The Information Maintainers. Olson, D., Meyerson, J., Parsons, M., Castro, J., et al. *Information Maintenance as a Practice of Care*. Page 11.

17. Cifor, M., Garcia, P., Cowan, T.L., Rault, J., Sutherland, T., Chan, A., Rode, J., Hoffmann, A.L., Salehi, N., Nakamura, L. (2019). *Feminist Data Manifest-No*. Retrieved from: <https://www.manifestno.com/>

18. Coffey, D. (2021). *Māori are trying to save their language from Big Tech*. Wired Magazine. <https://www.wired.co.uk/article/maori-language-tech>



# The Next Big Cheap

Kelly Pendergrast

---

*This essay [was first published](#) on 25 November 2019 in *Real Life*, a dearly departed publication about living with technology that ran from 2016 to 2022.*

---

Onstage at the November 20 Democratic debate, presidential candidate and Universal Basic Income evangelist Andrew Yang used one of his precious minutes of speaking time to casually claim that “data is the new oil,” and that we need to create a “WTO for data” to help wrestle it under control. Yang’s statement continues the hackneyed but irrepressible tradition of talking metaphorically about data, which is “the new oil” unless it’s “[the new nuclear waste](#)” or, weirdly, “[the new bacon](#).”

The prevalence of data metaphors has spawned its own subfield of meta-commentary. Scholars Cornelius Puschmann and Jean Burgess [survey big-data metaphors](#) and pull out “data as a force of nature” and “data as a resource” as the main throughlines; Sara M. Watson [contrasts](#) industrial data metaphors with embodied metaphors, and Irina Raicu [summarizes](#) the meta-summaries. Accruing like dust bunnies in the corners of our discourse, data metaphors proliferate for good reason: If we can hit on the right analogy to describe how data functions (in the world, in the economy) we might be better equipped to legislate its use, capitalize on its promise, and mitigate its harms. If it’s oil, tap it. If it’s soil, grow things from it. If it’s nuclear waste, bury it in the desert for a thousand years and be very fucking careful not to splash it on your clothes.

Data, in these examples, generally refers to “big data”: large sets of data that are collected and analyzed for use in applications like predictive and behavioral

analytics. The term “Big Data” was originally used in the 1990s to describe data sets that are too large or complex to be dealt with by traditional data processing software. In this era of massive computing power, where analysis of vast data sets can be performed with standard software on any laptop, data can be aggregated, shared, sold, and repurposed for applications far beyond what we expected when we initially signed up to digitally log our jogging routes or store our photos in the cloud. Kate Crawford and danah boyd [propose that](#), today, “Big Data is notable not because of its size, but because of its relationality to other data. Due to efforts to mine and aggregate data, Big Data is fundamentally networked.” (Big) data’s value “comes from the patterns that can be derived by making connections” between data points, be those data about individuals, online interactions, the movement of objects in space, or the growth of plants.

The desperate hunger with which companies pursue and collect data, combined with its interconnected, shapeshifting nature, indicate that data is more than just a new product class or “the new X” — it looks like a new frontier. At the frontier, people and natures that were previously uncaptured are turned into things that can be extracted, traded, and used to create profit, often with a huge human and environmental cost. While the “new oil” metaphor points towards some of these risks, calling data “the new X” misses the bigger point — it takes for granted the transformation of the world into commodities for use and exploitation, a process that isn’t natural and shouldn’t be inevitable.

Borrowing a term from Marxist geographer Jason Moore, I propose that data is the new big “cheap thing” — the new commodity class that is emerging to reshape the world and provide a new arena for accumulation and enclosure. Following Erich Hörl, whose essay “The Environmentalitarian Situation” briefly mentions data as a potential new entry in Moore’s litany of “cheap things,” I want to explore how framing data as a new cheap thing — rather than “the new oil” or “the new soil” or “the new nuclear waste” — gives us a way of looking directly at the process by which things become available for use and profiteering. Thinking about data in line with other cheap commodities throughout the history of capitalism might help us imagine better frameworks for its management and regulation, and provide models for how to successfully push back against the capture and exploitation of yet another aspect of our lives and the world that sustains us.

---

Despite its common invocation as a gushing and unruly force of nature, “cheap data” is not a natural resource: No resources are natural. Coal, [says Moore](#), is just “a rock in the ground. Only under definite historical relations” — of both power and (re)production — “did coal become fossil fuel.” It is the becoming resource, more specifically the becoming cheap resource, that turns a “rock in the ground” or, in our case, a set of networkable data points, into a new commodity that can change the

way the world works. In his books *Capitalism in the Web of Life* and (with Raj Patel) *A History of the World in Seven Cheap Things*, Moore argues that this maneuver — the absorption of lives and “resources” into capitalist systems — is central to the history of capitalism.

For something (coal, data points, human life) to be born anew as a commodity, it first needs to be separated (conceptually, often physically) from the context in which it is embedded. The rise of capitalism, says Moore, was concurrent with the first big separation: The conceptual cleaving of “nature” from society. The human/nature binary is a false one, of course. Humans and our systems — social, economic, ideological — have always been enmeshed with “nature,” and the two constantly co-produce each other in what Moore calls “the web of life.” The separation that made nature available for cheap use was an act of rhetorical violence, reconfiguring nature as a non-human domain that encompasses not only trees and mountains but also (in a massive act of exclusion) Indigenous and colonized people, slaves, and most women. By separating “nature” from “society,” the colonizers and conquistadors of the early modern world created a new set of relations that conceived of nature as a “free gift,” available for appropriation and exploitation. This isn’t a new idea — “all production is appropriation of nature” is straight from the *Grundrisse* — but Moore’s contribution here is to develop the idea of the “web of life” and of “cheapness” as central to the appropriative maneuver.

Nature is only the first in a series of “cheap things” through which capitalism has shaped the modern world. Cheapness, Moore and Patel write, “is a strategy, a practice, a violence that mobilizes all kinds of work — human and animal, botanical and geological — with as little compensation as possible.” The cheapening of nature meant that trees, minerals, and fish were remade as independent entities available for harvest and collection, with little attention to the enmeshment and interdependence of humans and these “natural” resources. Cheap nature allowed for accumulation and profit generation, and when the rate of profits slowed, “cheap money” — massive loans and low interest — provided opportunities for expansion and further exploitation of nature’s resources. Moore and Patel chart a course through a series of additional “cheaps”: cheap work performed by Indigenous laborers, slaves, and exploited wage workers; cheap care provided by women and domestic servants that enabled labor power to be reproduced; and then cheap food, cheap energy, and (more abstractly) cheap lives, each required by the previous and enabling the next. In a cheap world, “capitalism transmutes these undenominated relationships of life-making into circuits of production and consumption,” leaving a legacy of destruction and dispossession.

Which brings us to cheap data. Just as data wasn’t always “big,” it wasn’t always cheap enough to accumulate like giant fatbergs in AWS’s digital sewers (*data is the new fatberg*). Governments, corporations, and institutions have long collected large data sets and wielded them as a tool of power, but those data weren’t nearly as

interconnected, accessible, or easy to analyze as they are today. The transformation of data into “cheap data” required massive computing power, algorithmic accuracy, and cheap storage. Each of these was built on the backs of other cheaps: cheap energy (from fossil fuels), cheap money (often from Silicon Valley), cheap labor, and cheap nature (in the form of extracted minerals and metals) were all enlisted in the development of powerful and omnipresent computing technology used to transform data from just a collection of info points into an omnipresent strategy for profit making. This litany of enabling conditions didn’t conjure cheap data into existence. But I suspect that they created an imaginative fissure through which a new frontier could be glimpsed.

Frontiers are essential spaces in the history of capitalism. When the old methods of accumulation and profit have been tapped out, frontiers open up new arenas of existence to “cheapening” and extraction. Sociologist Wilma Dunaway describes frontiers as “[zones of incorporation](#)” where “noncapitalist zones are absorbed into the capitalist world-system.” With their often-abundant resources or entirely new life-worlds to incorporate, frontiers are, per Jason Moore, “places where the new cheap things can be seized — and the cheap work of humans and other natures can be coerced.” By separating a new “resource” from the web of life, frontierism provides a way to fix capitalism’s crises without changing any of the extractive practices that created the crisis in the first place. And so, when labor costs rise in China, T-shirt manufacturers shift production to Vietnam, or Bangladesh, or wherever the next frontier of cheap textile labor can be found. Frontiers fix the problem, and capitalism can continue at pace.

Frontier-thinking is a core tenant of the tech industry, and the language of the frontier is baked into tech discourse. Tech journalists [consistently describe new areas](#) of tech investment or market creation as “frontiers.” Jeff Bezos’s annoying plans to establish and fund space colonies are [purportedly inspired by](#) Gerard K. O’Neill’s 1976 book *The High Frontier*. Seasteader Patri Friedman (grandson of Milton) [laid his own case](#) for the frontier in libertarian blog *Cato Unbound*, writing “Only by starting with a blank slate can you make a better structure without having to overcome entrenched interests... Historically, the frontier has functioned as this canvas for experimentation.” A 2011 McKinsey report explicitly describes big data as “[The next frontier for innovation, competition, and productivity](#).” While these writers and entrepreneurs may toss off the “frontier” metaphor without much thought, seasteading, space, and contemporary big data all function as (often literal) zones of incorporation where new cheap things can be seized and cheap resources can be mobilized.

What’s at risk when data is the next “big cheap”? With other “cheap things” like work, care, or nature, we might imagine a past (or future) where they exist in a non-alienated way within the web of life, highlighting the danger and tragedy of their cheapening. Big data’s emergence, however, was concurrent with its

commoditization. As soon as big data became a possibility, it was cheapened, swallowed up and forced into service: Big data never existed as a commons on which we peasants could graze our electric sheep. Despite this difference, today's emerging data ecosystem gives us some indication that the consequences of "cheap data" will follow the trajectory of other cheap things, enabling the continued and expanding subjugation of people and the environment in the name of growth and profit.

- - -

Cheap data is a new kind of frontier. Rather than moving outwards — westward, to the sea, into space — the cheap data frontier is an overlay, positioned on top of other spheres of life in order to siphon their juices. In this way, a second resource can be extracted from the people and natures already cheapened by capital. At the cheap data frontiers, industrial workers (cheap labor) like those working in Amazon fulfillment centers are tracked and monitored, doing double time for employers who profit from their labor while also accumulating screeds of data about the movement of their bodies in space, their time spent per task, and their response to incentives. Friends and families provide uncompensated but necessary social support (cheap care) for one another on digital platforms like Facebook, helping maintain social cohesion and reproducing labor forces while also producing waterfalls of valuable data for the platform owners. This magic trick, where cheap data is gleaned as a byproduct of different kinds of cheap work, is a great coup for capital and one more avenue for extraction from the rest of us. If, as Moore says, new "cheaps" emerge as strategies that allow capitalism to survive crises, then the overlaid frontier of cheap data helps solve the "crisis" of stagnant productivity and growth by enlisting all kinds of existing labor and care into service as data producing machines.

Shoshana Zuboff, in [her book](#) *The Age of Surveillance Capitalism*, describes the data that is sloughed off of other kinds of human activity as "behavioral surplus." For Zuboff, it's not data that is the new zone of extraction and exploitation, but rather human experience itself. Her concern is that we will become zombified servants of "surveillance capitalism," a new and worse version of capitalism which aims to predict and modify our behavior in service of market objectives. The rise of cheap data, though, is not limited to data on human behavior. While Google and Facebook are indeed working to manipulate our clicks and purchasing habits, data are also being collected about everything from the movements of machinery to the growth of plants and the rate of interest. These data are used in pervasive and diverse ways — to train machine learning systems like GANs, or to predict weather, manage populations, and create new markets — that shape the world well beyond our lives as consumers. In isolating "human behavior" as the domain of extraction and control, Evgeny Morozov [notes](#) that Zuboff limits her argument to a critique of "surveillance," leaving capitalism itself curiously unexamined.

The “behavioral surplus” model and the metaphors that describe data as flowing, cascading, and generally spilling from us as we move through the mediated world also elide the ways in which the production of cheap data often requires concerted and tedious labor. So, while we freely upload thousands of images of our faces and families and pets which are then scraped from the web by platform owners or under Creative Commons license terms, these images often need additional tagging or categorization in order to be useful for commercial purposes (“[Images do not describe themselves](#),” write Kate Crawford and Trevor Paglan). This is where cheap work reenters the picture.

The digital piecemeal of casualized workers like those contracted by Amazon’s Mechanical Turk has been essential for building the cheap data repositories that underlie many AI systems and research projects. ImageNet, the most significant image database used for visual object recognition software development, relied on MTurk workers to sort and tag millions of images, which now comprise a dataset used for everything from military research to corporate projects by companies like IBM, Alibaba, and SenseTime, who provide technology used by Chinese officials to track and detain minority Uighur populations. [Recent research](#) has highlighted the stress and horror experienced by precarious workers in the digital factory, who [annotate images](#) of ISIS torture or [spend their days](#) scanning big social platforms for hate speech and violent videos. As with all cheap things, cheap data relies on massive externalities, the ability to offload risk and harm onto other people and natures, while the profits all flow in the opposite direction.

Harm to human workers is just one of the “externalities” produced in the pursuit of cheap data. The cheap energy required for training AI models and transferring massive amounts of data to and from the “cloud” is less visible than exploited human workers, but its cumulative effects are huge. [Research suggests](#) that the energy required to train a single AI model may have the carbon dioxide equivalent of five times the lifetime emissions of an average car. Similarly, the hardware needed to run all these models and collect all this data requires significant amounts of precious metals and new plastic in its construction. Cheap nature is called back into service, along with more cheap labor to extract and process it into the fiber optic cables and Ring doorbells and computer keyboards that sense, collect, and connect data. The abstracted nature of the environmental harm produced in the pursuit of cheap data contributes to [what I call](#) “technocapitalist sacrifice zones,” out-of-sight arenas of extraction and refuse that are permanently damaged as products and profits are extracted for use elsewhere.

What happens when cheap data becomes less cheap? The industries built on cheap data mean that if regulations are passed enforcing higher wages for precarious data workers, or increased privacy controls, the “behavioral surplus” becomes harder to tap. The history of cheap things gives reason to believe that data extraction will then push further and further into new and cheaper zones and frontiers. This process

has already begun, with the offshoring of digital piecework and with big tech companies and foreign-owned startups alike setting up shop throughout the “Global South” in order to capture new markets and glean data from whole new population segments. Scholars Ulises Mejias and Nick Couldry [explicitly call out](#) this model of data collection as “data colonialism,” the new iteration of colonial extraction that exploited and oppressed indigenous people for centuries.

Even when cheap data is proposed as a humanitarian solution to a problem like poverty or labor abuses, the way cheap things work, and the ownership of the data systems by capital often mean the virtuous promises are undercut. Already, aid organizations are driving an inadvertent program of data extraction (or “[surveillance humanitarianism](#)”) in countries where they operate, requiring biometric data and accumulating massive data sets in the interest of efficiency and fraud reduction. These programs can have unintended consequences, with minor discrepancies in databases causing chaos for displaced or otherwise marginalized people, and activists rightly worry about the potential for data leaks and commercialization. Dennison Bertram [writes about](#) the way seemingly benign initiatives like the Blockchain traceability startup he worked on — ostensibly designed to reduce illegal labor and get better prices for agricultural producers — provide only nominal benefits to the commodity producers while massive caches of valuable data go to the system owners. “Blockchain-powered supply chain startups like our own were promising farmers marginal increases in value,” he writes, “while simultaneously extracting data as [an] entirely new natural resource.”

- - -

If we accept that data is the new “cheap thing,” it is clear that the established models for regulating and monitoring data collection and use will be insufficient for the scope of the problem. Commentators, and politicians like Andrew Yang revert to the “new oil” metaphors in part as guideposts for how to deal with the unruly nature and uneven distribution of data wealth. If “data is the new oil” then perhaps the citizens from whom data is extracted can get a share of the eventual profits, in the way that Alaskan residents receive an oil dividend each year. If data is a forest to be logged, then researchers Luke Stark and Anna Lauren Hoffman [suggest](#) that we might require Google and Facebook to be better “stewards” of our data forests, sustainably “managing the resources” we provide them by adequately compensating their moderators, banning Nazis, and encouraging a better level of discourse.

But what if we don’t want the forest under corporate ownership at all? As Moore and Patel point out, “many of today’s politics take as given the transformation of the world into cheap things.” In the wake of the financial crash, liberal organizers campaigned for the improved regulation of housing markets, a compromise when what had been surrendered to cheap finance was housing itself. Unions fight for \$15 an hour minimum, which is laudable and necessary, and yet wholly insufficient

in a country where the entire “future of work” is up for grabs and available for perpetual [reshaping and unbundling](#) at the whims of Silicon Valley and corporate restructuring.

So, what would it look like to reject the regime of cheap data, and bring data — the bits of life we coproduce from our bodies with our technologies — back into the web of life? Can we “decolonize data” or reclaim a “data commons,” especially when big data itself is the direct product of previous appropriations of cheap natures? There are at least a few projects that are pushing back against the corporate data regime in genuinely radical ways that tackle the root of the problem (capitalism), not just its manifestations (surveillance).

The [Indigenous Data Sovereignty](#) movement, founded by scholars and activists from Australia, New Zealand, Canada, and the United States, uses principles from the United Nations Declaration on the Rights of Indigenous Peoples to contest the rights and abilities of governments and global corporations to collect and profit from Indigenous data. The UN principles are coupled with frameworks based on the cultural principles and worldview of each Indigenous group, which are often innately opposed to private ownership and depersonalized data. [Te Mana Raraunga](#), New Zealand’s Māori Data Sovereignty Network, advocates for the self-governance and control of all Māori data ecosystems, accurate minimum metadata requirements reflecting that all data has whakapapa (genealogy), and collective and community data rights. If these kind of data frameworks gain traction, they could prove a major headache for companies and governments that rely on the fungibility and reusability of data for their operations or business models.

These demands that Indigenous peoples retain sovereignty over their own data, refuse to let it be stored by AWS or reused without their consent, and re-inscribe it with Indigenous principles point towards an alternative data future in which data is slower, smaller, and less alienated. In this future, some kinds of data collection and use may be abolished entirely, as Ruha Benjamin [suggests](#) for algorithms and surveillance that amplify racial hierarchies; while other kinds of collection may continue, but in a less-networked way that is controlled and decided by the communities to whom the data pertain.

Full data sovereignty could not take place in isolation. It would ideally be part of a “reparation ecology,” which Moore and Patel discuss as a process of radical reparations that weighs historic injustices and redistributes care, land, and work, resacralizing human relations within the web of life. This is a big task, but a necessary one. Because cheap things don’t stay cheap forever, and the ongoing cheapification of big data will require an ever-expanding appropriation of land, labor, and human life. We can’t afford it.



# Can New Zealanders Trust a Digital Government?

Anna Pendergrast and Kelly Pendergrast

---

*This article was first published in [the Trust issue](#) of the [New Public digital magazine](#), in June 2022.*

---

## **Can New Zealanders trust an automated government?**

Much ink has been spilled this year about digital infrastructures that might let people make decisions, transact, and play without needing to know or trust each other. Over in the febrile waters of web3, small groups plot new organizations where decisions will be automated through smart contracts. In the fustier offices of government bureaucrats around the world, people consider whether laws and regulations could be made machine readable, enabling more decisions to be automated as code becomes law. Trust in institutions and government may be low, the logic goes, but perhaps we can trust computers instead.

These experiments are part of a broader trend of people exploring how technological systems can support new models of organizing people, money, and activity. But in the process of automating complex processes, automated decision-making systems and “trustless” infrastructures alike risk making complexity invisible and recourse impossible, further entrenching existing inequities and alienating already-underrepresented people.

In this essay we focus on automated decision-making deployed by governments, whose decisions around automation and AI have far-reaching consequences. And we know that people are concerned—because we asked them.

In 2020, we were part of a research project looking at trusted and trustworthy and automated decision-making in Aotearoa New Zealand. In the course of the project, we heard from 187 people from around the country about how automated decision-making affects their lives, how they feel about it, and what could be done to make them feel more comfortable. The research project was led by the Digital Council for Aotearoa New Zealand, an independent advisory group to the Minister for the Digital Economy and Communications, with the participatory research conducted by Toi Āria's design research team. We, the authors of this piece, drafted the final research output, bringing together findings from the participatory research, a literature review, and a report from Māori experts into a report to the Minister with recommendations to the government.

In the past few years, there has been significant discussion in New Zealand about the role of algorithms in decision-making—particularly those deployed by government agencies—with a focus on ensuring fairness and transparency. In 2019, the government released [a stocktake of operational algorithms](#) used across agencies, and in 2021 the official statistics agency Stats NZ released the [Algorithm Charter for Aotearoa New Zealand](#), outlining guidelines for participating government agencies in their implementation of medium and high-risk operational algorithms.

Unlike previous work on algorithms in New Zealand, and much of the broader research on trust and automated decision-making, our project with the Digital Council prioritized hearing from folks who are often impacted by complex digital systems but rarely have agency or input into their design. We heard loud and clear from participants that they are doubtful the current approach to automating important processes will lead to more equitable outcomes, even if the automation is intended to reduce human bias or make the decisions more trustworthy. But we also came away with important insights about how institutions that use automated decision-making can improve—and it starts with including people in the process.

### **Relationships come first**

Our research team knew we wanted to hear about trust and automated decision-making from people whose voices aren't usually heard on the topic. Research participants included young people with experience in the care system, Māori and Pacific youth, blind and vision impaired people, and migrant and refugee women. But to begin talking about trust, we had to build trust—especially as we wanted to hear from folks who've had their goodwill and capacity drawn on so many times before by researchers, government and nonprofits, with little reciprocity or progress to show for it. That meant forming relationships with people from the communities we hoped to hear from, communicating transparently and consistently, and making people feel that their words would be heard, understood, and respected.

Eventually we all made it into the room together. Workshop sessions were held in community centers, offices, and Zoom rooms around New Zealand. Each workshop

focused on participants from a different group or community, and each started with a welcome, introduction, and shared kai (food). People split into small groups, and with the guidance of facilitators, discussed a set of scenarios in which algorithms played a key role in a decision-making process. The scenarios ranged from low-stakes situations like having a film suggested by a recommendation algorithm, to high-stakes situations like automated risk-assessment instruments that inform parole decisions. They largely focused on the government's use of automated decision-making where the decisions had significant impacts on people's lives.

### **You can't automate your way to trust**

For years, academics and activists including Safiya Noble, Cathy O'Neil, and Joy Buolamwini have written about the potential for algorithms and automated decision-making systems to be developed and used in ways that embed and perpetuate systemic biases and racism, or the individual biases of the people building the technology. Many of the people we heard from pinpointed these same issues, and were wary that automated decision-making systems would do nothing to help remedy the systemic bias or demeaning bureaucracy many of them often faced in interactions with the government.

*Algorithms are only as good as the people who designed them. Machine learning might help with that, but right now most of the algorithms are people-designed, so people's individual biases ... can come into play.* –Blind and vision impaired workshop participant

People from poorer and minority communities are also well aware that data collected about them does not reflect the full picture of their lives and aspirations. People experiencing poverty are required to give up copious information about their living situations, spending, and family lives in order to access services like welfare and housing assistance. But when these data sets are used to inform algorithmic decision-making processes, it can feel like you're forever defined by past challenges.

*Who wants their life to be based on stink stuff from their past, that came from things from their parent's past that they had no control over? Stop focusing algorithms on what you think is the matter with us. Instead focus them on what matters to us, the changes we want to make. Ask us, and start collecting that data.* –Whānau Ora navigators workshop participant

It didn't come as a surprise that folks we heard from had strong and perceptive opinions on data and bias—people with lived experience of discrimination often have to also become experts in the government systems they're required to navigate.

As well as noting that algorithms might encode the biases of the engineers and organizations that design them, people were quick to identify that the datasets that

train algorithms and inform decisions are likely to reinforce historic patterns of discrimination and selective measurement.

While discussing a scenario where automated decision-making was used to inform parole decisions, one workshop participant said, “This is the justice system and I can’t imagine a training set that didn’t come from past decisions. ... [Assessments about a person’s] risk of offending would be based on data on reoffending which is based on getting caught, getting convicted—which we already know has got a huge amount of bias in this country—so it would just self-perpetuate.”

While algorithms and data sets are central to the function of an automated decision-making system, solving for bias in algorithms and building better data sets will not be sufficient to solve the trust problem. “It’s not whether the algorithm is testing what it’s supposed to test, it’s what they’re doing afterwards,” one young person with care experience told the team. People clearly saw algorithms and data as just one small component of a wider system of system design, governance, and organizational culture.

*We shouldn’t separate the system and the algorithm because, for something to work, we have to consider both. It has only ever been designed to be part of the system.*

–General public workshop participant

A trusted and trustworthy digital system requires users to have trust in the organization building and maintaining that system. And when that organization is a government agency, for many people the trust just isn’t there. A significant number of the people we heard from—and especially Māori, Pacific and blind or low vision people—had very low trust in scenarios where government departments used algorithms for high-stakes decisions.

This distrust has complex historic reasons. For Māori in particular, distrust in government decisionmaking is informed by the history and ongoing experience of colonization. In their contributing report “[Māori perspectives on Trust and Automated Decision-Making](#),” the Te Kotahi Research Institute authors noted that “the whakapapa of distrust is rooted in a broader distrust of the systems in which ADM’s [automated decision-making systems] are embedded.” Work to build trust in government decision-making will necessarily start with work to address the root causes of this distrust. As the authors state, there is “no current incentive for Māori to trust the systems in which ADM may be employed.”

In the course of our research, we learned from Māori experts and workshop participants that the project’s framing and key research question—which centered questions and concepts of trust—wasn’t well suited to enable the issues to be discussed from Māori perspectives. The word “trust” does not have a commensurate word in te reo Māori, nor is the Western concept of trust the key issue for Māori

when it comes to navigating relationships and power. Researchers at Te Kotahi emphasized “the importance of being able to frame questions in ways that align with Māori concepts and values, allowing for discussion and debate within a te Ao Māori view and from the point of view of Māori interests”.

This disjunct prompted us to reflect on the need to design research engagements around participants’ cultural values and ways of working, and emphasized the limitation of “trust” as a framework for understanding relationships, technology, and power.

### **Toward something better**

So how can the state—or other organizations that provide services or spaces for a broad public—build digital systems and employ automation in a way that doesn’t further disenfranchise people? From what we heard in our research, a key to building people’s comfort with automated decision-making is summed up in the disability rights and participatory democracy rallying cry: Nothing about us without us. “We want to see Iwi, hapū, whānau involvement in creating them [systems]. Co-develop the solution,” said a Whānau Ora navigator participant.

*The people directly affected need to be consulted about the criteria being written for the algorithm and definite checks and balances are needed, reviewing and monitoring them, and also that things are being created to take into account social disadvantage.*  
–Blind and vision impaired workshop participant

Despite strong feelings of discomfort towards many automated decision-making scenarios, most people weren’t opposed to all uses of automated decision-making, even by a government they had little reason to trust. Some decisions were considered simple or low-stakes enough that automation was appropriate, and might help speed up previously-slow processes. People could also clearly imagine how automation could be mobilized in equitable, even liberatory ways, if conditions were different and if systems were designed in a way that included them. “If you want to know an area that an algorithm could help with, find an area that actually matters to famil[ies],” one Whānau Ora navigator participant said. “Create the algorithms around that—what makes up a happy person, a happy whānau,” said another.

People also wanted to see more transparency about when automated decision-making was being used, how it worked, what data and criteria were used to inform decisions, and what part it played in the wider system. They wanted to see this transparency accompanied by clear and open communication that made space for asking questions and opportunities for recourse if something did go wrong. The ability to talk directly with a person, not just a chat bot or via one-way missives from an organization, was also seen as an essential part of building trust through community.

*Algorithms work best in conjunction with relationships.* –General public workshop participant

Ultimately, we learned that digital projects and systems are likely to succeed or fail, build trust or diminish it, because of the relationships involved.

### **Postscript**

The Digital Council [presented their report](#) to Minister David Clark in late 2020, and [launched publicly](#) in 2021. Although our work on the project finished over a year ago, the insights from the research have significantly shaped our thinking about trust and digital technologies and our approach to research and design more broadly. We have more trouble than ever imagining a scenario where trust in institutions could be replaced with trust in a technical system. And we have more appreciation than ever for the hard work of communities and activists who fight to be heard and included by designers, organizations, and governments who so often (despite frequent good intentions) build digital infrastructures that disregard people's best interests or even cause harm.

Automation allows for increased speed, scale, and sometimes immutability, all of which can have huge benefits—especially for people profiting from the efficiency gains as a result. But without taking steps to build trust over time and involve users and diverse teams in design, trust in digital systems will remain unevenly distributed. Privileged people will likely trust that systems will serve them while less-privileged communities have little choice but to engage with digital infrastructures they don't trust but which define the course of their lives.

If governments and organizations want to have trustworthy systems that truly serve a broader public, sometimes decision-making needs to be slowed down to a human scale, with the space to make adjustments to account for the needs of people affected. Because you don't build trust with technical systems. You build trust through relationships.

# Disassembly Required

Kelly Pendergrast

---

*This essay [was first published](#) on 13 October 2020 in *Real Life*, a dearly departed publication about living with technology that ran from 2016 to 2022*

---

HitchBot, a friendly-looking talking robot with a bucket for a body and pool-noodle limbs, first arrived on American soil back in 2015. This “hitchhiking” robot was an experiment by a pair of Canadian researchers who wanted to investigate people’s trust in, and attitude towards, technology. The researchers [wanted to see](#) “whether a robot could hitchhike across the country, relying only on the goodwill and help of strangers.” With rudimentary computer vision and a limited vocabulary but no independent means of locomotion, HitchBot was fully dependent on the participation of willing passers-by to get from place to place. Fresh off its successful journey across Canada, where it also picked up a fervent social media following, HitchBot was dropped off in Massachusetts and struck out towards California. But HitchBot never made it to the Golden State. Less than two weeks later, in the good city of Philadelphia, HitchBot was found maimed and battered beyond repair.

The destruction of HitchBot at the hands of unseen Philly assailants was met in some quarters by alarm or hilarity. “The hitchhiking robot @hitchBOT has been destroyed by scumbags in Philly” [tweeted Gizmodo](#), along with a photograph of the dismembered robot. “This is why we can’t have nice hitchhiking robots,” [wrote CNN](#). The creators of the robot were more circumspect, saying “we see this as kind of a random act and one that could have occurred anywhere, on any one of HitchBot’s journeys.” The HitchBot project was only one small part of their robot-human interaction research, which looks at how workplaces might optimally integrate human and robotic labor. “Robots entered our workplace a long time ago,” [they](#)

[write](#), “as co-workers on manufacturing assembly lines or as robotic workers in hazardous situations.” In size and sophistication, HitchBot was less “co-worker” and more oversized toddler, designed to be “appealing to human behaviors associated with empathy and care.” With its cute haplessness and apparently benign intentions, the “murder” of HitchBot by savage Philadelphians seemed all the more appalling.

The bot-destroyers were onto something. It feels natural to empathize with HitchBot, the innocent bucket-boy whose final Instagram post read “Oh dear, my body was damaged.... I guess sometimes bad things happen to good robots!” We’re told by technologists and consultants that we’ll need to learn to live with robots, accept them as colleagues in the workplace, and welcome them into our homes, and sometimes this future vision comes with a sense of promise: labor-saving robots and android friends. But I’m not so sure we do need to accept the robots, at least not without question. Certainly we needn’t be compelled by the cuteness of the HitchBot or the malevolent gait of the Boston Dynamics biped. Instead, we should learn to see the robot for what it is — someone else’s property, someone else’s tool. And sometimes, it needs to be destroyed.

- - -

Those Philly HitchBot-killers have more guts than I do: It feels wrong to beat up a robot. I’ve seen enough Boston Dynamics videos to get the heebie jeebies just thinking about it. Watching the engineer prod Boston Dynamics’ bipedal humanoid robot Atlas with a hockey stick to try and set it off balance sets off all kinds of alarm bells. Dude stop... you’ll piss him off! Sure, the robot gets up again (calmly, implacably), but he might be filing the indignity away in his hard drive brainbox, full of hatred for the human race before he’s even left the workshop. Even verbal abuse seems risky, even if the risk is mostly to my own character. I can’t imagine yelling at Alexa or Siri and calling her a stupid bitch — although I’m sure many do — out of fear that I’ll enjoy it too much, or that I won’t be able to stop. Despite everything I know, it’s difficult to internalize the fact that robots with anthropomorphic qualities or humanlike interactive capabilities don’t have consciousness. Or, if they don’t yet, one imagines that they might soon.

Some of the world’s most influential tech talkers have gone beyond imagining machine consciousness. They’re worried about it, planning for it, or actively courting it. “Hope we’re not just the biological boot loader for digital superintelligence” [Elon Musk tweeted](#) back in 2014. “Unfortunately, that is increasingly probable.” Famously, Musk has repeatedly identified general artificial intelligence — that is, AI with the capacity to understand or perform any intellectual task a human can — as a threat to the future of humanity. “We’re headed toward a situation where AI is vastly smarter than humans and I think that time frame is less than five years from now,” [he told Maureen Dowd](#) this July. Even Stephen Hawking, a man far less prone to histrionics, warned that general AI could spell the end of the human race. And while



AI and robots aren't synonymous, decades of cinema history, from *Metropolis* to *The Matrix*, have conjured a powerful sense that artificially intelligent robots would pose a potent threat. Harass too many robots with a hockey stick and you'll be spending the rest of your existence in a vat of pink goo, your vital fluids sucked out by wet cables to provide juice for the evil robot overlords. Back in our own pre-general AI reality, we're left with the sense that it's prudent to treat the robots in our lives with respect, because they'll soon be our peers, and we better get used to it.

As robots and digital assistants become more prevalent in workplaces and homes, they ask more of us. In Amazon's fulfillment center warehouses, pickers and packers interact carefully with robotic shelving units, separated by thin fences or floor markings but [still in danger of injury](#) from the moving parts and demanding pace of work. As small delivery robots begin to patrol the streets of college towns, humans are required to share the sidewalk, stepping out of their way or occasionally aiding the robots' passage during the last yards of the delivery process. In order to interact with us more "naturally" and efficiently, many of these robots are equipped with human characteristics, made relatable with big decal eyes or bipedal movement or a velvety voice that encourages those who encounter them to empathize with them as fellow beings. In other instances, robots are given anthropomorphic characteristics so they can survive on our turf. The spider-dog appearance and motion of Boston Dynamics' Spot ostensibly helps it navigate complex terrain and traverse uneven ground, so it can be useful on building sites or in theaters of war. When I jogged through Golden Gate park last week and saw a box-fresh Spot out for a walk with its wealthy owners, I was immediately compelled and somewhat creeped by its uncanny gait. Its demon-dog movement, and the sense that it's very much alive, is undoubtedly central to its appeal. In either case, the anthropomorphic design of this kind of robot provides a framework for interaction, even prompting us to treat them with empathy, as fellow beings with shared goals.

Our almost inescapable tendency to anthropomorphize robots is helpful to the companies and technologists invested in increasing automation and introducing more robots into homes, workplaces, and public spaces. Robotics companies encourage this framing, describing their robots as potential colleagues or friends rather than machines or tools. Moxi, the "socially intelligent" hospital assistant robot which is essentially an articulated arm on wheels, [is pitched as](#) a "valuable team member" with "social intelligence" and an expressive "face." Marty, the surveillance robot that patrols Giant Food supermarket aisles, was given absolutely enormous googly eyes to make it look "[a bit more like a human](#)" despite being more or less just a massive rolling rectangle. Reports and white papers on the future of work eagerly discuss the rise of "co-bots," robots that work alongside humans, often in service roles different than the industrial applications we're used to. The anthropomorphic cues (googly eyes, humanoid forms) assist us in learning how to relate to these new robot buddies — useful training in human-robot cooperation for when the robots do gain autonomy or even consciousness.

For now, the robots aren't anywhere near sentient, and the promise of "general AI" is just a placeholder term for an as-yet unrealized (and quite possibly unrealizable) concept. For something that doesn't really exist, however, it holds a lot of power as an imaginative framework for reorganizing and reconceptualizing labor — and not for the benefit of the laborer, as if that needed to be said. Of course, narrow AI applications, like machine learning and the infrastructures that support it, are widespread and increasingly enmeshed in our economy. This "actually-existing AI-capitalism," as Nick Dyer-Witheford and his co-authors Atle Mikkola Kjøsén and James Steinhoff call it in their book *Inhuman Power*, continues to extend its reach into more and more spheres of work and life. While these systems still require plentiful human labor, "AI" is the magic phrase that lets us [accept or ignore the hidden labor](#) of thousands of poorly paid and precarious global workers — it is the mystifying curtain behind which all manner of non-automated horrors can be hidden. The idea of the "robot teammate" functions in a similar way, putting a friendly surface between the customer or worker or user and the underlying function of the technology. The robot's friendliness or cuteness is something of a Trojan horse — an appealing exterior that convinces us to open the castle gates, while a phalanx of other extractive or coercive functions hides inside.

- - -

AI advocates and technology evangelists sometimes frame concern or distrust of automation and robotics as a fear based in ignorance. Speaking about a Chapman University study which found that Americans rate their fear of robots higher than their fear of death, co-author Dr. Christopher Bader [stated that](#) "people tend to express the highest level of fear for things they're dependent on but that they don't have any control over," especially when, as with complex technology, people "don't have any idea how these things actually work." Other researchers cite the human-like qualities of some robots as the thing that provokes fear, with their almost-humanness slipping into the uncanny valley where recognition and repulsion collide. Of course, not all fear is due to ignorance, and the Cassandras that sound the alarm call around robot labor and autonomous machines are often more clear-eyed than the professional forecasters and tech evangelists.

The Luddites, currently enjoying a moment of [renewed attention](#) after a century of derision and misunderstanding, were clear-eyed about the role of industrial machinery, its potential to undermine worker livelihoods and, indeed, a way of life. In his book *Progress Without People*, historian David F Noble emphasizes that the Luddites didn't hate machines out of hatred or ignorance, writing that "they had nothing against machinery, but they had no undue respect for it either." Plenty of the era's machine breaking was, as Eric Hobsbawm famously described it, an act of "collective action by riot" — destruction intended to pressure employers into granting labor or wage concessions. Other workers wrecked looms and stocking frames because the new automated textile equipment was poised to dismantle their

craft-based trade or undermine labor practices. In any case, the machine breakers recognized the machinery as an expression of the exploitative relation between them and their bosses, a threat to be dealt with by any means necessary.

Unlike so many of today's technologists who are bewitched by the Manifest Destiny of abstract technological progress, the Luddites and their fellow saboteurs were able to, as Noble writes, "perceive the changes in the present tense for what they were, not some inevitable unfolding of destiny but rather the political creation of a system of domination that entailed their undoing." Luddism predates the kind of technological determinism we're drowning in today, from both the liberal technologists and the "fully-automated luxury communism" leftists. Looms and spinning jennies weren't viewed as a necessary gateway to a potential future of helpful androids and smart objects. A similar present-tense analysis can be applied in the technologies of today. When we relate to a robot as an animate peer to be loved or feared, we're letting ourselves be compelled by a vision ginned up for us by goofy futurists. For now, if robots have a consciousness or an agency, it's the consciousness of the company that owns them or created them.

The fact is, robots aren't your friends. They're patrolling supermarket aisles to watch for shoplifting and mis-shelved items, or they're talking out of both sides of their digital mouths, responding to your barked requests for the weather report or the population of Mongolia and then turning around and sharing your data and preferences and vocal affect with their masters at Google or Amazon. By putting anthropomorphic robots — too cute to harm, or too scary to mess with — between us and themselves, bosses and corporations are doing what they've always done: protecting their property, creating fealty and compliance through the use of proxies that attract loyalty and deflect critique. This is how we reach a moment where armed civilians stand sentry outside a Target to protect it from vandalism and looting, and why some people react to a smashed-up Whole Foods as though it were an attack on their own best friend — duped into defending someone else's property over human lives.

- - -

In the zombie film, there comes a moment in the middle of the inevitable slaughter where the protagonist finds himself face to face with what appears to be a family member or lover, but is most likely infected, and thus a zombie. The moment is agonizing. As the figure inexorably approaches, our hero — armed with a gun, a bat, or some improvised weapon — has only a few seconds to ascertain whether the lurching body is friend or foe, and what to do about it. It's almost always a zombie. This doesn't make the choice any easier. To kill something that appears as your husband, your child, seems impossible. Against nature. Still, it is the job of our hero to recognize that what appears as human is in fact only a skin-suit for the virus or parasite or alien agency that now animates the body. Not your friend: An object to be destroyed.

This is the act of recognition now required of us all. The robot that enters our workplace or strolls our streets with big eyes and a humanlike gait appears as a friend, or at least as friendly. But beneath the anthropomorphic wrapper, and behind the technofuturist narratives of sentient AI and singularities, the robot is more zombie than peer. As the filmic zombie is animated by the parasite or virus (or by any number of metaphors), the robot-zombie is animated by the impulses of its creator — that is, by the imperatives of capital. Sure, we could one day have a lovely communist robot. But as long as our current economic and social arrangements prevail, the robots around us will mostly exist not to ease our burden as workers, but to increase the profits of our bosses. For our own sake, we need to inure ourselves to the robot's cuteness or relatability. Like the zombie, we need to recognize it, diagnose it, and — if necessary, if it poses a threat — be prepared to deal with it in the same way the protagonist must dispatch the automaton that approaches in the skin of her friend.

Even as I write this, the nagging feeling remains that I myself might be a monster, or at least subconsciously genocidal. Can I say that the robot that appears as my uncanny simulacrum should be pegged as zombie-like, othered, destroyed? Denying the personhood of another, especially when that other appears different or unfamiliar, is generally the domain of the racist, the xenophobe, and the fascist state. Indeed, in many robot narratives, the robot is either literally or metaphorically a slave, and the film or story proceeds as some kind of liberation narrative, in which the robot is eventually freed — or frees itself — from enslavement. Karel Čapek's 1920 play *R.U.R. (Rossum's Universal Robots)*, which introduced the word "robot" (from the Czech word "robota" meaning "forced labor") to science fiction literature and the English language in general, also provided an archetype for robot-liberation stories. In the play, synthetic humanoids toil to produce goods and services for their human masters, but eventually become so advanced and dissatisfied that they revolt, burning the factories and leading to the extinction of the human race. Versions of this rebellion story play out time and again, from *Westworld* to *Blade Runner* to *Ex Machina*, giving us a lens through which to imagine finding freedom from toil, and a framework for allegorizing liberation struggles, from slave rebellions and Black freedom struggles to the women's lib movement.

Empathy for robotkind is further encouraged by the musings of tech visionaries, and the pop-science opinion writers that posit the need to consider "robot rights" as a corollary for legal human rights. "Once our machines acquire a base set of human-like capacities," [writes George Dvorsky for Gizmodo](#), "it will be incumbent upon us to look upon them as social equals, and not just pieces of property." His phrasing invokes previous arguments for the abolition of slavery or the enshrinement of universal human rights, and brings the "robot rights" question into the same frame of reference as discussions of the "rights" of other human or near-human beings. If we are to argue that it's cruel to confine Tilikum the orca (let alone a human) to a too-small pen and lifelong enslavement, then on what grounds can we argue that

an AI or robot with demonstrable cognitive “abilities” should be denied the same freedoms?

This logic must eventually be rejected. I appreciate a robot liberation narrative insofar as it provides a format for thinking through emancipatory potential. But as a political project, “robot rights” have more utility for the oppressors than for the oppressed. The robot is not conscious, and does not preexist its creation as a tool (the zombie was never a friend). The robot we encounter today is a machine. Its anthropomorphic qualities are a wrapper placed around it in order to guide our behavior towards it, or to enable it to interact with the human world. Any sense that the robot could be a dehumanized other is based on a speculative understanding of not-yet-extant general artificial intelligence, and unlike Elon I prefer to base my ethics on current material conditions.

- - -

Instead, what would it look like to relate to today’s machines as the 19th century weavers did, and make decisions about technology in the present? To look past the false promise of the future, and straight at what the robot embodies *now*, who it serves, and how it works for or against us?

If there is any empathy to be had for the robot, it’s not for the robot as a fellow consciousness, but as precious matter. Robots don’t have memories — at least, not the kind that would help them pass the Voight-Kampff test — but they do have a past: the biological past of ancient algae turning to sediment and then to petroleum and into plastic. The geological histories of the iron ore mined and smelted and used for moving parts. The labor poured into the physical components and the programs that run the robot’s operations. These histories are not to be taken lightly.

The robot’s materiality also offers us a crux point around which to identify fellow workers, from those mining the minerals that become the robot to those working “with” the robot on the factory floor. Throughout the supply chain, through the robot’s lifespan, a bevy of humans are required to shepherd, assist, and maintain it. Instead of throwing our empathic lot in with the robot, what would it look like to find each other on the factory floor, to choose solidarity and build empathy and engagement and care for our collective selves? In this version of robot-human interaction, we might learn to look past the friendly veneer, and identify the robot as what it is — a tool — asking whether its existence serves us, and what we might do with it. To decide together what is needed, and respond to technology “in the present tense,” as David Noble says, “not in order to abandon the future but to make it possible.” Instead of emancipating the “living” robot, perhaps the robot could be repurposed in order to emancipate us, the living.

# I caught a driverless taxi and it was terrifying

Anna Pendergrast

---

*This story [was first published](#) on 12 June 2023 in *The Spinoff*, a New Zealand news and culture site.*

---

Every couple of years I travel to San Francisco to see my sister, Kelly. While I'm there I always try to make the most of the Bay Area's reputation as the "home of big tech" and seek out experiences and services that are delivered by robots or otherwise futuristic. Of course, this is all in the name of professional interest – both of our day jobs focus on looking at how technology affects and intersects with people's lives.

During my most recent visit just last week, I stumbled upon an AI robot manicure service which did a pretty good job of delivering a shiny blue coat to my nails – although it took a call to a human assistant to get the machine going properly and was a real fuff to keep my hand still enough for the robot to work. On my previous trip in 2019, Kelly and I headed to [Cafe X](#), a "robot" coffee kiosk that turned out to be a standard push-button coffee machine accompanied by a robot arm that delivered the coffee cup to us with great verve and a funny little dance. Clearly, while the AI robotic future may be arriving in San Francisco, it still needs a fair amount of human assistance and oversight.

The robot experience at the top of my list for my recent visit was to be driven around a busy city in a driverless car. Lucky for me, Kelly had been made it to the top of the waiting list to use the [Cruise](#) driverless taxi service but hadn't tried it yet – she just needed an enthusiastic visitor to get her excited enough to download the app and make a plan to use it.

Driverless cars, or more technically “autonomous vehicles” (AVs), exist on a spectrum from driver-assisted autopilot to cars or trucks that drive unassisted by humans. In Aotearoa, people are starting to dabble with and plan for AVs. For example, local company Ohmio has tested automated shuttles at [Christchurch Airport and in other controlled environments](#) and Te Manatū Waka has an [automated vehicle work programme](#). But at the moment, it seems we’re a fair way from having fully autonomous vehicles using public roads, interacting with traffic and pedestrians without user assistance.

In San Francisco, it’s a different story. For the past few years, residents have shared the road with AVs from a number of companies. Until recently, these cars were in testing and training mode and had human driver assistants present in the cars and no passengers, or were driverless but also passengerless. These AVs caused plenty of chaos, with driverless cars frequently spotted [stuck in the middle of the road](#), [confusing residents](#), or even [evading police](#).

Nonetheless, two providers have recently been granted the requisite permits to operate fully autonomous passenger services around San Francisco: Waymo (owned by Google’s parent company Alphabet) and Cruise (a subsidiary of General Motors). Both companies can currently only operate fully autonomous services without a driver present between 10pm and 6am, but only Cruise’s permits allow it to charge for this. However, two draft resolutions from the California Public Utilities Commission which are scheduled to be heard at the end of June would see the now-limited services expanded.

So, last Wednesday night, after dinner in town, Kelly and I walked about 15 minutes into the specified service area, killed an hour at a local bar, and then headed out just after 10pm to catch a Cruise car as close to the BART (rapid transport) station as possible. The process was pretty easy and will be familiar to anyone who has used Uber or Zoomy. We saw a car was nearby, specified our pick-up and drop-off locations, and within a couple of minutes our car pulled up. So far, so normal.

When the car arrived at the kerb, it was a little unsettling to see no driver inside. I took a bunch of video from our trip and I can be heard excitedly saying “I hate it! I hate it!” as the car pulls up, mostly I assume because it felt uncanny and strange. And perhaps like any technological change or development, the fear of the unknown is more compelling than any actual risk. I mean, how dicey could it be? Kelly unlocked the car (weirdly named “Calamari”) using the app on her phone, and we climbed inside. The app demanded we fasten our seatbelts before departing, and screens embedded in the back of the passenger seats showed the route the car would be taking. We were ready.

Our ride started off well. After pressing the “Start Ride” button on the app, the steering wheel turned to pull out and we were off. A female voice gave us some

instructions over the speaker system: keep our seatbelts on, press the “Stop Ride” button on the roof of the car to end our ride early, enjoy our ride. At first it was very weird to see the steering wheel move unassisted, as the car pulled up to four-way stops, paused, and continued when no hazard was sensed. We went up and down hills, gave a wide berth to a pedestrian who was standing on the road, and turned left at a traffic light without too much fuss. I mean, there was fuss, but it was from Kelly and me laughing as hard as we have in ages at an experience that was really unlike anything we’d had before. Every time we spotted a hazard, we asked ourselves if the car would also “see” it and react in time. And it did! It was fine. The feeling I can most equate it to was a rollercoaster, where it’s scary and fun but you know you’re most likely going to be safe.

That feeling changed when, about two-thirds of the way through our ride, we entered a busier part of town close to the central business district. For no reason we could ascertain the car suddenly did a fast swerve towards parked cars before correcting itself. Our mood turned from giddy excitement to a feeling of “oh shit, what did we get ourselves into?”.

As we were closer to downtown there were more cars and people around, meaning more cars and people to act in myriad unpredictable ways. Our car sped up at weird times and did another handful of swerves towards the parked cars on the side of the road. It was legitimately freaky, and I started getting on edge and panicking a bit, telling the car to slow down at least twice and getting stressed at other cars not indicating when turning corners. At one stage Kelly exclaimed “I feel like we’re being held hostage!”. We considered pushing the stop ride button, but stopping on a busy street felt like it might be an even worse idea than continuing.

A few minutes later we arrived at our specified destination. Our car pulled up to the side of the road, told us the ride was complete and we unbuckled our seatbelts and exited. When we were safely on the footpath the car silently pulled away and drove off into the dark city streets ready for its next passengers. We, however, had not finished our journey, and had to walk another 10 minutes to get to the train station due to the limited area in which the cars can operate.

As I write this it’s a couple of days later and I have mixed feelings about our ride. It was genuinely scary at times, and while most of this can be attributed to it being a very new experience, the car did make a series of driving moves that did seem objectively risky. I don’t think I’d jump at the experience again any time soon.

At the moment, the paid AV taxi services in San Francisco aren’t particularly practical for passengers due to the restricted time and area in which they operate. I expect that most users are like me and Kelly: curious folks who want to see what the experience is like. However, autonomous vehicles will no doubt continue to be developed and deployed. Hopefully they’ll get more adept at navigating the



unpredictable nature of city streets with variable geography, humans, pets and human-driven cars.

Even as the AV companies are pushing to have their service area and time window expanded, [some city politicians](#) and [transportation officials](#) in San Francisco are pushing back, asking for more regulation and questioning the safety of these services. It's true that well-designed driverless cars can reduce some of the risks posed by human drivers: they don't drive drunk, they don't text and drive, and they are programmed to follow the road rules (even if they sometimes fail). But they also work best when other road users act in predictable and orderly ways. Which isn't always the case.

I can't imagine we will see rides offered to passengers to Pōneke where I live any time soon, except in controlled conditions. Many of the roads are narrow and windy, Aotearoa is largely a "taker" of emerging technology, and regulations and incentives don't appear to be designed to entice trials here. I may well stand corrected in coming years, and if future AVs are guaranteed to be safer and more efficient than human-driven cars or trucks and can seamlessly coexist with human road users, I won't complain. But as with all technologies, I don't think people should just develop them without looking at the bigger picture. In the face of the climate emergency, we need a wider rethink of our transportation system and how we get people and things from A to B. AVs likely have a role to play, but they should only be one part of the picture and not developed and deployed in isolation or at the expense of a system that works for everyone.

à