

OPEN

MINDS



# THE AI (R)EVOLUTION

**Valuing Country, Culture and  
Community in a World of Algorithms**

---

**ASSOCIATE PROFESSOR JESSICA  
RUSS-SMITH & PROFESSOR  
MICHELLE D. LAZARUS**



# THE AI (R)EVOLUTION

VALUING COUNTRY, CULTURE AND  
COMMUNITY IN A WORLD OF ALGORITHMS

---

ASSOCIATE PROFESSOR JESSICA  
RUSS-SMITH & PROFESSOR  
MICHELLE D. LAZARUS



MONASH  
UNIVERSITY  
PUBLISHING

Published by Monash University Publishing  
Matheson Library Annexe  
40 Exhibition Walk  
Monash University  
Clayton, Victoria 3800, Australia  
[publishing.monash.edu](http://publishing.monash.edu)

Monash University Publishing: the discussion starts here

© Copyright Jessica Russ-Smith and Michelle Lazarus 2024  
Jessica Russ-Smith and Michelle Lazarus reserve the right to be known as the authors of the work.

All rights reserved. Apart from any uses permitted by Australia's *Copyright Act 1968*, no part of this book may be reproduced by any process without prior written permission from the copyright owners. Enquiries should be directed to the publisher.

Series: Open Minds

ISBN: 9781922979667 (PDF)  
ISBN: 9781922979674 (ePub)

Cover design by Les Thomas  
Typesetting by Typography Studio



This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) which means that the text may be used for non-commercial purposes, provided credit is given to the author(s) and that no alterations are made. For details go to <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

# Contents

<i>Acknowledgement of Country</i> . . . . .	vii
<i>Honouring First Nations Knowledges</i> . . . . .	ix
<i>Cultural, Content and Sensitivity Statement</i> . . . . .	xi
<i>Introduction</i> . . . . .	xiii
1. Defining AI	
<i>The role of technologies in a life-centred world</i> . . . . .	1
2. The History and Development of AI	
<i>Transformations of technology over time</i> . . . . .	28
3. Bias in AI	
<i>Building the machine to support all life</i> . . . . .	48
4. The Ethics of AI	
<i>As complex as life itself</i> . . . . .	79
5. A Life-Centred Future	
<i>Sustaining Country in an AI world</i> . . . . .	118
<i>Acknowledgements</i> . . . . .	141
<i>Links</i> . . . . .	143
<i>Notes</i> . . . . .	155
<i>About the Authors</i> . . . . .	163



## Acknowledgement of Country

Much of the research and writing of this book took place on the unceded lands of Ngunnawal Country and the Boonwurrung/Bunurong and Wurundjeri peoples of the Kulin Nation. We pay our respect to Elders past and present for the ongoing custodianship of Country, and to all Aboriginal and Torres Strait Islander peoples and communities, including any readers of this work.

One of the authors of this book, Jessica Russ-Smith, is a sovereign Wiradyuri Wambuul woman. The other, Michelle Lazarus, a US migrant to Australia, comes from a lineage of occupiers who have yet to reconcile with Country and sovereign nations. We both stand in solidarity with First Nations peoples here and across the world. First Nations sovereignty was never ceded, and always continues, and these ways of knowing and being are crucial for a life-centred future.

Yindyamarra ngurambang-gu: respect for and to Country. Country is the core of life, and we are always in relationship with Country.



# Honouring First Nations Knowledges

We first want to acknowledge sovereign Wiradyuri knowledges and the Wiradyuri Nation and peoples. This book learns from and is guided by Wiradyuri knowledges. As a sovereign Wiradyuri Wambuul woman and a non-Indigenous woman, we respectfully honour these knowledge systems that belong with and to Wiradyuri peoples and Wiradyuri Country.

We, the authors, do not own these knowledges. Even Jess, as a Wiradyuri Wambuul woman, does not own these knowledges. She has custodianship of, and responsibilities to care for, these knowledges, which includes advising readers on how to *not* use, abuse, misuse and dislocate First Nations knowledges explored in this book. This applies especially to researchers, universities and governments, who capitalise on knowledge production, and AI, which might take the physical knowledges presented in this book and separate them from the relationships in which these knowledges exist. Any reference to Wiradyuri knowledges taken from this book must be acknowledged as belonging to the Wiradyuri Nation.

This responsibility extends to readers of this book: these knowledges cannot be taken out of the sovereign First Nations context in which they are situated. To not honour the knowledge holders, laws and cultural protocols of care to these knowledges is to culturally appropriate,

## Honouring First Nations Knowledges

violate and colonise. For non-Wiradyuri peoples, to learn from these knowledges – including Wiradyuri ways of knowing and respecting knowledge through cultural protocols including acknowledgement of Country, acknowledgement and solidarity with Wiradyuri and other First Nations sovereignty – one must seek appropriate permissions and build relationships with Elders and community.

# Cultural, Content and Sensitivity Statement

This book contains content and discussions that may be sensitive or triggering for some readers. Discussions include themes or topics related to colonial violence, other forms of violence, racism, ableism, sexism, queerphobia, discrimination, environmental abuse, sexual assault, rape culture, child abuse (including child sexual abuse materials), suicide and health violence. We encourage you to reach out for support as needed.



# Introduction

Here's a radical thought: thinking about AI is not just for programmers, computer scientists or the consumer. AI technologies impact everybody, and we all have a stake in the world they are bringing into being.

This book explores the extent to which AI technology impacts our lives. Western society is structured in a way that functions to protect, preserve and privilege the lives of certain groups, often at the expense of others. The following pages seek to name and disrupt this process by exposing the role that AI plays as a hegemonic tool.

Let's get one thing out of the way first: this book is not anti-AI. We are not luddites or reactionaries – we believe that AI has the power for good. After all, we met through an incredible seminar on AI and First Nations data sovereignty, which was held online. We use Chatbots to help write challenging communications (e.g. constructive feedback), and other AI programs to organise our lives. This book, and the remarkable relationship that has built from this chance meeting, was supported through the technological transformation (or iteration of technology) of the internet. We recognise that AI transformations and digital expressions (the varied forms that technology takes through time) have the potential to both connect and divide. But the use of AI is complex, and the reach of these technologies means that AI will affect almost every aspect of our lives and those of all living things on this planet, including all future generations.

We all have the power, opportunity and responsibility to shape technology, and to influence the way we are governed by it. The alternative is the continued violence and erasure of certain types of knowing and living. If you are someone who has paused before pressing ‘accept all’ on a privacy statement, this book may illuminate some of your concerns. If you have enthusiastically embraced AI and can’t wait to see how it develops, this book may be an opportunity to challenge the way you think about and interact with AI and technology.

A cynic might ask, ‘How do I know that AI didn’t write this book?’ It didn’t, but you only have our word to take for this. This tension underscores the key themes discussed in this book: How do any of us know what we know? In a world of uncertain information, what is ‘real’ and verifiable – or has the concept of authenticity become too problematic altogether? Who decides on, determines and filters the information we receive? How can you know if the outputs from AI are trustworthy?

To address these and other questions, we explore perspectives, experiences and stories from those often left out of the conversation about technology. This is one indication that AI didn’t write this book – some of the views we present may not (yet) have made it into AI datasets, may be purposefully omitted from AI or may be hidden within the *White* noise of existing datasets.

Ultimately this book tries to, as Denise Utochkin, postdoctoral researcher in AI and algorithmic fairness at the University of Copenhagen, puts it so succinctly, [‘cut the AI bullshit’](#) and have a frank, open conversation about how to work with AI in a way that enhances rather than denies our humanity. The future of AI, and its relationship to humans and all living beings, is not set. At this moment in history, we have an opportunity to revolutionise how AI is used, and

what ways and forms it takes in the world. We want to provide readers with the agency and insight to critically consider current forms of AI and reimagine what AI could be in the future.

## A life-centred future

When discussions about AI focus on the impact of technology on humans exclusively, they do so at the expense of other elements of life, including Country. Instead, this book is about the relational impacts of technology on Country, which is all forms of life, including animal and human life, ecosystems and land. This relational understanding reflects the Wiradyuri cosmology (systems of knowledge and being with the world).

Wiradyuri is a First Nations community and Jess's sovereign Nation. Wiradyuri cosmology has Country at its core. As Jess wrote in 2019: 'Country is not just a place or space that holds meaning. Country is our sovereignty. Country is a "field of self" ... It is past, present, future, life, death, story, dreaming, all at once. It is greater than us, it is our anchor to all things, then, now and always.'<sup>1</sup> Jess was also taught by her Elders that Country is all things: it 'is the land, water, people, animals, ancestors, stories, songlines and sovereignty'.<sup>2</sup> This understanding of Country differs from Western definitions of land that relate only to soil and space. First Nations understandings and valuing of Country also differ significantly from nationalist perspectives where 'country' refers to Empire. For Jess as a Wiradyuri Wambuul woman, Country is Wiradyuri Country, the stories of her family and ancestors, and the responsibilities these imbue.

Wiradyuri cosmology, as outlined by Uncle Stan Grant Senior and John Rudder in *A Grammar of Wiradjuri*<sup>3</sup> *Language*, understands that

beings and entities do not exist in isolation but in relational systems.<sup>4</sup> We, as guided by the Wiradyuri cosmology, see all living things as existing in relationship to one another and with Country, and therefore believe that any AI (r)evolution for a life-centred future must consider these relationships.

Elements of this relational understanding of the interconnectedness of all life, which First Nations cultures have held for tens of thousands of years, have also been described in other fields over the last century. For instance, some in the Western world refer to this concept as ‘posthumanism’, which is defined as ‘the questioning of human exceptionalism and the foundational role of “humanity” as it has been constructed in modernity’<sup>5</sup>. In 2018, researcher Siân Bayne noted: ‘Rejecting any clarity of distinction between “nature” and “culture”, [posthumanism] works against dualism and the binaries we have tended to draw on to define what it means to be.’<sup>6</sup> This understanding can also be seen in critical posthumanism, which, as described in *EuropeNow*’s [‘Rethinking the Human in a Multispecies World’](#), ‘seeks to deprioritize and weaken human-centrism, rejecting individualism, and instead underscoring the compatibilities between human animals, nonhuman animals, and machines’.

Wiradyuri ways of understanding the world are (and always have been) grounded in the relationship between Country, culture and life – in particular, in caring for Country and all forms of life across generations. First Nations communities in so-called Australia are the oldest living and surviving cultures in the world, so while we support the Western concept of posthumanism, and acknowledge that the idea that AI is in relationship to the living and vice versa is represented across cultures and through time, Wiradyuri ways of

knowing existed long before these later related concepts, and thus we centre this Wiradyuri wisdom.

## Revolution or evolution?

The use of '(R)evolution' in the book's title is not incidental. It pays homage to the Wiradyuri ways of knowing, specifically the concept of wayanha, that underpin this book. The Wiradyuri concept of wayanha, meaning transformation, suggests that nothing ever stops existing but rather transforms through different expressions.<sup>7</sup> This is core to how we, the authors, understand AI. The continuous transformation that is wayanha can be seen as an evolution in which living things transform over time. Similarly, we see AI as an evolution, a mode of technology that transforms its expressions – as in how it looks and works – over time.

Importantly, our discussions of evolution actively resist racist weaponisation of evolution theories.<sup>8</sup> Social Darwinism, and its 'survival of the fittest' trope, has been used violently under the guise of sociobiology by White settler colonial states to position the erasure of certain cultures as 'natural and inevitable'.<sup>9</sup> These forced, socially constructed and unscientific racial hierarchies have been embedded institutionally and socially where arguments of Social Darwinist evolution are used to inflict violence, dispossession, assimilation and genocide, including in the colony of so-called Australia. Such ideologies continue through violent actions that attempt cultural extinction.

The term 'revolution' can be defined in two ways: a forcible overthrow or signifying the turning or revolving of an entity. Our book argues for both – it examines the revolutionary impacts of AI and technology on life in all its transformations, and calls for a revolution to overthrow

life-taking AI expressions. This offers humans an opportunity to be a part of the AI evolution in revolutionary ways that preserve the future of life – hence the title *The AI (R)evolution*.

### Voicing our perspectives on AI

As neurodivergent women, one of us a sovereign Wiradyuri Wambuul scholar and the other an American scholar living in Australia, we represent groups on the margins of the conversation about AI, those whose perspectives are often not sought. We think it is time to centre our voices.

We grew up in geographically diverse areas, different cultures and distinct communities. We each identify with cultures not represented by the dominant [Australian hegemony](#). Our perspectives allows us unique insights into the interconnectedness of technology and life, and to understand technology from non-centred points of view. When discussing the ideas that became this book, we realised that while we may come from different cultures and parts of the world, we share a curiosity about the future of life and a desire to understand the entangled relationship between technology and the living world.

Too often, the conversation around AI negates the relationships or lives of Country, certain cultures and identities, and our future generations. By sharing our own stories of technology and its relationship to our lives, we are challenging this. We offer our perspectives so they can be heard at this crucial moment in the story of AI, *while* the technology that will shape our world is being developed and implemented. Humanity is on the cusp of an AI revolution, and the questions we raise are crucial to how well AI will serve us into the future.

*Jessica Russ-Smith, Wiradyuri Wambuul woman*

I am still exploring exactly how I feel about AI. My relationship with AI and my understanding of it has transformed in the last few years.

My early understandings of AI came from depictions in movies and on television. I pictured a conscious digital being, but this was no more than a vague concept. AI technology to me was a fairytale from a digital storybook. But this transformed for me when AI became a hot topic within universities.

In learning spaces over the last few years, we have seen countless think pieces and mandated training courses that focus on AI as the new form of plagiarism and contract cheating (paying someone to do work assigned to you). AI is seen as an epidemic threatening to ruin scholarly integrity. In this narrative, AI is something to fear, introducing new risks, such as violations of academic conduct policies, that need to be met with swift punitive measures in order to protect academic standards. I have found the underlying tenor of these discussions to be about demonising students and their use of AI. I have found myself in meetings regularly objecting to and critically questioning university responses to AI, imploring colleagues to consider how these technologies may also advance learning and help to flatten the hierarchies of power in higher education. I have asked colleagues in academic integrity roles to consider how AI might support disabled and neurodivergent folks like me. I have encouraged my curriculum development colleagues to see AI as an opportunity to look critically at our learning and teaching, to help find the gaps and fill them, or to re-create them in ways that aren't saturated in Whiteness but in sovereign First Nations ways of knowing, and that consider the experiences of disabled folks.

I began to wonder if AI could be a decolonising tool for those of us to whom the curriculum does not often speak. Or is it just another form of a colonial being, just another tool in the violence of colonisation, adding only greater speed and further distance? Is AI the coloniser of the colonisers?

As these questions flooded me and I felt out of my depth trying to navigate this in a way that centres justice, I needed to reconnect with the knowledge of my ancestors. So I turned to Wiradyuri knowledge and picked up the Wiradyuri dictionary. I turned from anxious waters to our language, a calm knowing river, to understand what artificial intelligence means from a Wiradyuri perspective.

The term ‘artificial’ is not something I have often heard used in my family or community. I frantically flicked through the Wiradyuri dictionary and could not find an exact translation. However, the English language often fails to capture many direct and conceptual translations of Wiradyuri language. In our sovereign language, words themselves are larger stories, and each word has meanings that are brought together in the relationship of a sentence to communicate a story depending on the context. I began to think more conceptually, and I thought of artificial intelligence as being made by another, which as an adjective in Wiradyuri language is defined by Uncle Stan Grant Senior and John Rudder in *A New Wiradjuri Dictionary* as ‘bun-ngan’.<sup>10</sup> From here I began to look in the dictionary for similar Wiradyuri words, word stems, prefixes and suffixes and found the following:

- bunambirra: sweep
- bunan: ‘fine dust, ashes, anything carried by the wind, a dust storm’
- bunmarra: to do or make

- -marra as a suffix: ‘indicates an action is being made or caused to happen’.<sup>11</sup>

I then turned the pages looking for the meaning of ‘intelligence’ and found:

- winhanga-bilang: intelligent
- wudha-ng-garang-garra: to be intelligent and know a great deal.<sup>12</sup>

Therefore, from our language I understood AI as sweeping all the dust or knowings out there in the world, where the sweeping is an action being made to the dust or different knowings, to gain intelligence and know a great deal about a certain topic.

But what is the dust made of? Whose sovereignty is acknowledged in this dust – that is, what colour is the dust? Will the dust of knowings be used to suffocate the voices of certain groups and Nations? Will the broom sweeping the dust sweep up the voices and knowings that often go unseen? Will it have gaps that favour some forms of dust and not others? These reflections made me wayamiilbuwawanha: turn my eyes and feet back to stand in the self, reflect deeply, and consider what dust I leave in the world. What is the dust I put out that the AI broom sweeps up and then gives to someone who is seeking intelligence and knowing?

My dust is my sovereignty, my knowing as a Wiradyuri Wambuul woman and the knowledges of my culture, ancestors and Country.

I also considered the Wiradyuri concept of wayanha, transformation. Uncle Stan Grant Senior and Dr John Rudder use the analogy of a butterfly to explain this concept. A butterfly has many stages of

transformation, including egg, caterpillar, cocoon and butterfly. They explain that the identity of the butterfly always exists; its expression merely transforms through these stages of life, and all these stages are connected.<sup>13</sup> The butterfly is always constant even as it takes different forms, such as a caterpillar.

Wayanha helps me understand AI not as something entirely new, but as the latest expression of a particular digital technological identity that exists in entangled relationships across time and culture, interwoven with life. I understand AI as connected with all the technology of the past that transformed over time to result in this latest expression. AI exists in relationship with all prior versions of itself and all future versions to come. To illustrate this, think of a smartphone. Communication is not a new concept, nor did it solely emerge through digital technologies. It is a dimension of human culture that has been transformed over time through methods and mediums such as smoke, fire, speaking, markings, sign language, pre-electric telegraphs, electric telegraphs, [transmitters and receivers, switchboards \(manual, electric and digital\)](#), handsets, cordless phones and now the smartphone. Communication is expressed through mediums and modes that transform over time, just as a digital technology expression transforms over time.

At this point, the questions consumed me. Our sovereign knowledges and cultures have been, and continue to be, violently harmed by colonial White supremacist powers. Will AI do the same? How does the dust AI sweeps up reflect White supremacy, and colonial and Western ways of knowing? Will White supremacy be manifested in AI? Is AI just the latest transformed expression of technology more broadly?

Western epistemologies see knowledge as owned and produced.<sup>14</sup> Does AI challenge ownership and gatekeeping of First Nations Knowledges? Is AI helping to make sense of 'big data' or is it just

assimilating and colonising ‘data’ into colonial parameters of definitions and meanings? Is this just replicating pan-Aboriginality (the idea that Indigenous ways of knowing are universal and singular) and therefore reproducing colonial logics of telling narratives about First Nations cultures as homogenous groups of the past?

What about Country – what is Country and land to AI? Knowledge is Country; as a Wiradyuri woman, I exist in relationship to and with Country. This relationship must be socially and relationally situated in line with our cosmology and cultural ways of being. How does AI understand this relationship? Can AI embody this relationship? Does it have the right to embody this relationship? Can AI help demolish the colonising walls of settler colonies, or will it make them stronger?

The importance of ethics washes over me as I sit in the ocean of these questions, as does my role and responsibility as both a Wiradyuri Wambuul woman and a Social Worker to fight for justice. I invite you into this ocean to float in these questions and think about the bunan, the dust and the knowings you put into the world. Is it colonial dust, decolonising dust or sovereign dust? The dust we leave is not neutral. What is the dust we leave, and will this create storms or light for our future generations to come?

*Michelle D. Lazarus*

My positioning in the world is predominantly in the dominant Western cultures and discourses. Aside from being a woman, most of my experiences have been as someone (at least in the visible, outward-showing context) with great social privilege whose perspectives are centred in this world.

I was born and raised in the United States, where ‘rugged individualism’, as interwar president Herbert Hoover put it, defines

the culture. Growing up, I believed that everyone had the power to change their circumstances if they just worked harder. If you weren't succeeding, it was because you weren't trying hard enough. Because of my privileged position, I wasn't challenged (and didn't challenge myself) to consider all the contextual influences that this logic ignores.

My family were the classic early adopters: we believed that technology made our lives easier. In my household, we always had the latest technology. When CD players came out, we rushed out and bought one; then we had the first one on the market that could store five CDs. Our televisions got bigger, thinner and sharper each year when the sales were on. Our mobile phones got 'smarter' and smaller (then bigger again) with every iteration. Up until relatively recently, I thought these technologies made me work more efficiently and afforded me more impact in the world.

My work has brought me, in a serendipitous and unexpected way, towards studying uncertainty and how we as humans manage the unknown. While this research started in the field of healthcare education, it has led to opportunities to explore how we process unknowns as they relate to sustainability, education, social justice – and, more recently, AI and technology.

As I continue to learn more about humans' capacity for uncertainty tolerance (or how we respond to perceived uncertainty) in these different fields, my curiosity about technology grows. During the last few years, as AI technologies have been rapidly embedded in the fields I know best (healthcare and education), I have started to revise my perception of technological advance, to shift from proselytising to cautious contemplation.

From the social justice perspective, I am increasingly aware of AI's capacity to further entrench bias and inequality in a way that far

surpasses that of previous technologies. My place within the dominant culture resulted in me being blind to inequities for many of my formative years, though now that I am alert to technology's capacity to harm, they are impossible not to see. I am now the annoying person in the room at every AI event, raising my hand, asking about the impact of the AI on different populations. I have been called on to talk about AI, most often in relation to its threats. It feels like a full 180 from my upbringing. I now find worrisome the speed with which our society adopts AI, and the lack of discussion of the impacts on those in our community who are on the margins or most vulnerable. This worry has sparked a deeper impulse to question technological 'advancements'. My uncertainty about AI grows with each passing day.

I am also aware that AI has the power for social good. Its capacity to support humankind and other life through monitoring, feedback and pattern identification means that, for me, the debate about whether AI is 'good' or 'evil' is misguided. The world is a series of entangled, complex and uncertain systems, with intermingled relationships between living things and technology.

I am increasingly interested in learning more about how, and the extent to which, humans can moderate AI, and how the technology influences life. Ultimately, I am entering this discussion with far more curiosity and cognitive flexibility than I afforded myself growing up. I am now grappling with the tension between my desire for new efficiencies (oh, how I love efficiency!) in work and in life, and my recognition of the negative impacts of AI on so many living things – from the perpetuation of structural inequities, to AI's contributions to climate change, to the impact of AI on contract workers in the Global South. We will look at many of these topics in the book.

Through the process of writing and researching this book with Jess, I am rescripting my understanding of the AI evolution and the uncertainty it poses – and also recognising just how many Western concepts are not new ideas, but ideas and knowledges held for generations by First Nations peoples here and around the world. In acknowledging this, I have come to realise how many potential solutions to the problems around AI can be found in these sovereign ways of knowing.

### Journeying through this book

Although you have heard our separate stories, our goal in this book is not ultimately to present two different viewpoints, but to create a shared voice formed through the process of working together. This voice serves to re-examine the role of AI in society and its impact on living things – now and into the future. We want to put this relationship between AI and life front and centre.

This book can only represent our current thinking. We, as authors, are regularly transforming in our understandings of AI. Our understandings are also shaped by what humanity currently knows about AI technologies. Time isn't purely linear – and neither is knowledge or existence. We recognise the limitations that come with a Western conception of the passage of time, in contrast to a Wiradyuri view that encompasses past, present and future together. Such approaches restrict our ability to convey the turbulent advancements in and transformations of technology in relation to Country and life.

In one sense we are all knowledge-holders of AI, given the widespread embedding of such technologies into our world, but the type and level of our awareness depends on our unique experiences with such technologies, along with the structures providing access (or not) to

them. As the AI (r)evolution continues to rapidly transform lives (and in turn, transform the technology), the experiences and ideas in this book may begin to fit or align to certain contexts less, and to others more. However, we hope this contribution to the revolution for a life-centred future holds value for you, its reader.

Over the course of these pages, as we face the reality (not the hype) that AI brings to our lives and begin to see what Western and settler colonial approaches are designed to leave out, there may be an element of discomfort. In interrogating the relationship between life and AI, we discuss who wins and who loses on our current trajectory, and explore whose futures are preserved. Facing this reality can be challenging. Yet it is critical.

How we define knowing and intelligence is also crucial to exploring humanity's relationship with AI and technology. We centre First Nations ways of knowing and being in the ways we discuss AI and the language we use in these discussions. Words will be translated into English from language that has existed for tens of thousands of years, and so may not be directly represented in words common in the English language. We do this purposefully to disrupt White English norms that have become embedded over time (and in AI programs), to challenge English as the benchmark through which normalcy and acceptability are measured.<sup>15</sup>

The language, sentence structure, grammar and tone we use in this book are chosen to reflect the sovereign knowledges and ways of being and speaking of the Nations these knowledges belong to, specifically Wiradyuri Country.<sup>16</sup> For example, we use the term 'relationship'. In English, *relationship* could be understood as a singular noun; the plural is *relations* or *relationships*. This division of singular and plural is not the way Jess speaks. When Jess speaks about the relationship of something,

she is honouring the Wiradyuri cosmology that sees everything in the world, both living and non-living, as related and existing across a web of relationships. Therefore, when we use ‘relationship’ in this text, it should be understood from the Wiradyuri understanding of many connected relationships.

Similarly, if a word in English does not feature an ‘s’ or similar at the end to identify it a plural, it is usually perceived as singular. But this is not necessarily the case for Wiradyuri language – where all words, and the knowledges they represent, are in relationship, and thus plural in a sense.<sup>17</sup>

Another example of this relational concept of terms is the use of ‘knowledges’ as a plural, not ‘knowledge’ as a singular. This may feel counterintuitive to the discussion on ‘relationship’ above, but in fact it is not. We use the term *knowledges* to refer to the many First Nations cosmologies that exist and have existed for tens of thousands of years. Many First Nations languages represent a pluralistic and relational set of concepts and ideas, whereas English tends to represent an assumption of individualism unless noted otherwise (through adding ‘we’ or an ‘s’ to the word). As Barrett Holmes Pitner – founder and philosopher-in-chief of The Sustainable Culture Lab, and author of *The Crime Without a Name: Ethnocide and the Erasure of Culture in America* – discusses, in English ‘I’ is capitalised, whereas in many other languages it is not. Barrett argues that this represents the individualistic nature of Western society (and even further reinforces individualism) through grammar.<sup>18</sup> In a similar way, the ‘s’ in knowledges when referring to Indigenous knowledges pays homage to the diverse and relational ways of understanding the world, encompassing the concepts of community and connectedness. Ultimately, terms like ‘knowledges’ and ‘relationship’ used in this text are translations of Wiradyuri language – so with or without the ‘s’,

when used in reference to Wiradyuri cosmology in this book, we must assume that this refers to relational ways of understanding the world.

The analysis in this book sees AI as part of a wider story of transformations of technology over time, with AI as a current expression of these technological transformations. As such, we refer to AI and other technologies as ‘technological expressions’ or ‘technological transformations’. As we view AI as just one expression of technological transformations over time, we also use non-AI examples of technology to represent the larger lineage of technological transformations. We are using these diverse examples to encourage readers to look at contemporary examples of AI and its impacts differently and through a relational lens.

This is not a technical text on AI. While we do delve into programming, algorithms and datasets, we centre our focus on the social, ethical, moral and emotional. The paths we venture down to examine the relationship between life and AI may feel negative or gloomy to some, but this critique is necessary to highlight the life-affirming values we are arguing for. Tone policing or minimising the emotions and impacts is not what is needed in technology transformations; emotion, feeling and honouring life as it is experienced is central to the AI revolution.

Country is always at the core of centring First Nations, specifically Wiradyuri, ways of understanding AI and other transformations of technology. Locating this book and our analyses on the lands from which we write is critically important. We will draw on First Nations perspectives from across the globe, but the centring of the local context in so-called Australia is essential in honouring the ways of knowing, being and doing that underpin this book and the places we write from. We encourage readers to consider this book’s words, ideas, and

## Introduction

ways of knowing and being alongside their local context. By critically reflecting on how you connect to or diverge from the ideas we present, you may find resonances that apply to your Country, wherever you are in the world. There can be pressure in society to consider aspects of life at the global level – to push us towards ‘generalisability’ – but we remain aware that knowing is inseparable from Country. Country is where life begins, and where we begin in defining AI and its impacts on culture and community.

## Chapter 1

# Defining AI

### The role of technologies in a life-centred world

Artificial intelligence: these words seem to have defined our recent history. The topic features in headlines across the globe and appears to make its way into nearly every conversation, from the dinner table to the classroom, from the workplace to our governments.

The technologies we collectively refer to as AI, or artificial intelligence, are pervasive: ‘smart’ devices and ‘intelligent’ machines are integrated into our homes, our workplaces, and our social, healthcare and education systems. They influence what [we see on the news](#), [how we receive healthcare \(and who receives care\)](#), how we communicate with our loved ones, [who gets interviewed](#), [who gets hired](#), [who is incarcerated](#) (and even which neighbourhoods are policed more), what content we see (and don’t see), whose story is shared, whose story is blocked, which knowledge is seen as truth and, more importantly, which knowledge is framed as a lie.

Despite AI technologies becoming increasingly entrenched across the globe, there remains debate around the definition, function and impact of what we call artificial intelligence.

## The tricky thing about defining AI

What is the first image that comes to mind when you picture AI: HAL from *2001: A Space Odyssey*? What about the Sentinels from *The Matrix*? Skynet from *The Terminator*, Pixar's *WALL-E*, JARVIS and FRIDAY from *The Avengers*? Or are you, more prosaically but also more practically, imagining conversations you have had with chatbots?

The term 'artificial intelligence' has traditionally conjured up visions of humanoid computers, visions that are simultaneously exhilarating and terrifying. Today, we may picture our digital assistants or recall our engagement with ChatGPT, where the AI is used to recognise and predict speech patterns and respond accordingly. This is where we begin our passage into the world of artificial intelligence – with language, and the different meanings it can elicit.

Many have struggled to define what AI is and isn't. The AI industry tends to define [AI by its capabilities and computing power](#). Stanford University provided a brief summary and characterisation of such AI categories back in 2020. There are *autonomous systems* that provide outputs without much human interference, once programmed; and *machine learning* (ML), which moves towards the science of developing machines that, for all intents and purposes, 'think like a human', by attempting to draw on knowledge from neuroscience, statistics and psychology, among other fields.

Despite these noble intentions to position AI as a branch of science, others argue that AI is based on 'magical metaphors' akin to alchemy. As [Thomas Krendl Gilbert](#), a machine ethicist, explained to technology journalist Sharon Goldman, AI is 'not scientific, in the sense that it's not rigorous or experimental'. Typical scientific endeavours are born from systematic experimentation of the studied subject, and many,

[including AI researchers](#), suggest an absence of such rigour in the development and testing of contemporary AI expressions. While we have rules and guidelines for developing and testing other technological expressions such as cars, taps, televisions and so on, there seems to be less concern with equivalent levels of evaluation for AI. Keep this in mind as we continue to discuss the types of AI currently available, as it's an important point.

*Deep learning* and *artificial neural networks* refer to technological architecture that most closely approximates higher-level brain functions such as learning from mistakes and improving performance over time. Deep learning, theoretically, requires less human input to function than other forms of ML, and is used in applications such as translation of images to text or colourising images. Another term that you might see related to AI capabilities is *algorithm*, which defines the instructions a human relies on to program and train AI.

*Weak* or *narrow AI* tends to be used when characterising AI with a limited and focused capability, such as facial or speech recognition. In contrast, *foundation AI models*, such as ChatGPT4, Llama 3 and Claude 3, are trained on large datasets and are more general in their capacities but can still be finetuned for specific tasks. At the other end of the spectrum is the holy grail for many AI companies: *artificial general intelligence (AGI)*, which does not yet exist but could theoretically replicate – and possibly exceed – all aspects of human intelligence, including humans' ability to integrate knowing with social and contextual awareness (e.g. the sun is warm because we can feel and experience its warmth).

IBM draws on [the functional perspective](#) in defining AI: 'On its own or combined with other technologies (e.g., sensors, geolocation, robotics) AI can perform tasks that would otherwise require human

intelligence or intervention.’ Others have endeavoured to ground their understanding of AI not in the forms it takes but in the [relationship between humans and the technology](#). From this perspective, AI is defined by humans’ capacity to trust its output. AI and education researchers Margaret Bearman and Rola Ajjawi, writing in the journal *Medical Education* in 2024, explain:

This relational definition is concerned with what happens within a particular moment of use between a human and an AI and therefore can help with managing the realities of AI in practice. To give a specific example, a calculator is not generally considered an AI, but a 4-year-old must trust a calculator’s outputs without any way of knowing whether it is right or wrong. We suggest, therefore, that when such a child uses the calculator, this is an AI interaction. But when an adult uses the calculator, it is not. Thus, AI is not dependent on the technological specifications or even what it might do, but on the relationship between the human and the technology.

Similarly, Emily Bender, Professor of Linguistics at the University of Washington, [views AI interactions as potentially interrupting the relationship](#) between knowledge, information and context. When an AI produces an output, we cannot see the trail of how the output was produced, severing us from the elements that allow us to evaluate and explore the claim’s veracity and the relationship to the evidence that leads to this claim. Bender points out that unlike a Google Scholar search, where knowledge and outputs are traceable back to the people who developed the knowledge, often an AI chatbot output is simply an output with no ‘paper trail’: there is no way to identify which elements

of knowledge were analysed (and from which sources) in order to develop this output. AI minimises the chance of serendipitous discovery and, over time, denies us the opportunity to build an internal ‘mind map’ of connections between sources of information. Such relational definitions of AI consider the tangled web between AI, knowledge and human learning.

These relational definitions are in broad alignment with Jess’s conceptualisation of AI, drawing on her Wiradyuri cosmology, in which the relationship between the technology and those interacting with it, as well as AI’s former technological expressions, are embedded. Every technological expression stands in relation to another and is transformed through this relationship in its next expressional form. For example, the internet is a technological expression that evolved from other technologies, such as fibre optics, hardware, servers and routers. Yet all of these technologies together would not have spontaneously become the internet; the internet developed because of the complex relationship between humans and these technologies. Every technological expression has elements based on prior technological expressions. Just as we can use DNA to trace the genealogical origins of an individual, each technological expression carries the DNA of the technological expressions that came before it and that influence its current form.

What these varied definitions and understandings suggest is that as AI’s functionality is transforming, so too are our definitions of it. AI is both a pervasive and a dynamic entity. In writing this chapter, we had many discussions exploring how we begin to characterise AI, and what definitions serve us now as well as potentially into the future.

Because of these varied and sometimes contentious definitions, each time we refer to AI, imagine that we have placed inverted commas

around it to reflect the diverse conceptualisations conjured from this single term. Indeed, as English professor Dr Katie Conrad [tweeted](#) in April 2024, ‘AI is too sloppy a term to refer to current generative systems. How about Synthetic Homogenizing Information Technology? Catchy acronym, too ...’. Meanwhile, Professor Marek Kowalkiewicz repurposed the words of American philosopher Harry Frankfurt: ‘bullshit ... intended to persuade without regard for truth’. He coined the term ‘botshit’ to describe the ChatGPT phenomenon.

The takeaway is that AI isn’t all that the developers and evangelists wants you to believe it is. While some suggest that the industry generates positive hype around AI technologies largely to garner funds to further develop AI, we need to recognise that how we brand technology matters. [The language we use shapes the thoughts we have.](#) We need to ask whether the current trajectory of AI transformation is the answer to humanity’s flaws, or whether it is based on bullshit – driven by profit, not purpose.

## Wiradyuri ways of knowing AI

In our own paths to understanding AI, we explored the nature and meaning of the terms ‘artificial’ and ‘intelligence’ in our different cultures. In Wiradyuri language, AI can hold connotations of ‘dust’ and ‘being made by another’. In this context, Jess’s reflection summarises AI expressions as ‘sweeping all the dust or knowings in the world ... to know a great deal about a certain topic’. Jess’s cultural knowing frames AI as gathering up the dust – data, knowings of the world – and sweeping it into piles based on how a knowing appears, or is *seen* to appear. In one pile are the leaves, in another the dirt, and so on. This sweeping is done at times in an effort to simply know more, and

at other times knowledge is used as power by some to inflict violence and dominate others.

The ‘made by another’ part of Jess’s Wiradyuri understanding is key to the process of how AI is created and maintained. AI programs do not write themselves (yet); humans begin the programming and coding (or the query that leads to this). Humans mine the materials used to build and run and maintain the computer; they often influence the programming and coding as well. And even if the AI were to write itself, or the next transformation of itself, it still always exists in relationship to other beings, entities, objects and lives by drawing from living resources to sustain itself, and by its impacts on the living.

In social discourse, AI is often seen as a holder of knowledge, but a Wiradyuri worldview helps us understand how AI is the result of humans and their purposeful actions *to* know. AI is made by another (human/s), and the knowing used to program AI comes from the lives, experiences and histories of many others. In Western cultures, this is being increasingly brought to the forefront through litigation around copyright and intellectual property. Some in the arts, such as authors, artists and media personalities – arguably many of whom have careers based on contextual knowing – are suing AI companies for stealing their ideas in an effort to help the AI ‘learn’. To quote technology journalist [Ben Lutkevich](#) in 2024: ‘The authors want companies to pay for the data they scraped for training – the “food” for AI systems, endless meals for which there has been no bill.’

Many technological expressions, such as generative AI large language models (LLMs), rely on data scraping, or extracting data, often from the internet, for the purposes of supporting AI development. When LLMs like ChatGPT were first released, much of the scraped data was presumed to be human-generated. As time goes on and more of AI’s

own data is pumped out onto the internet, the risk is that [AI eats itself](#). AI needs new data, from real individuals, in order to remain relevant. While [synthetic data](#) is a potential way to sustain AI expressions, it isn't an answer by any means.

The point of difference between Wiradyuri and other First Nations ways of understanding AI 'knowledge' (such as [Abdilla et al., 2021](#); [Williams et al., 2021](#); [Indigenous AI, 2024](#)) and non-Indigenous Western perspectives is that AI illustrates a Western understanding and valuing of knowledge, in which knowledge is a commodity that humans produce and own. The pertinent questions become who 'owns' this knowledge, or whether knowledge can ever actually be owned.

From a Wiradyuri perspective, and as Bearman and Ajjawi suggest in their paper, we cannot remove epistemology (that is, ways of knowing) from conceptualising, defining and understanding AI. Epistemology is integral to understanding how AI affects life – in all its forms.

In Wiradyuri culture, knowledge always exists. Knowledge is present whether we are aware of it or not,<sup>1</sup> and we come to know when we are ready to know. Jess explains this in her journey to learn how to speak Wiradyuri language. Jess did not learn how to speak Wiradyuri language until her twenties, but her nan told her that when she was a young child she would say Wiradyuri words even if she had not been taught them. Wiradyuri language was always known to Jess, but in a different way, and she began to learn it in new ways later in life when she was ready to come to know it.<sup>2</sup>

When Scottish physician Alexander Fleming came to understand the impact of penicillin on bacteria (to take a prominent example from the history of science), Western ways of knowing considered this his 'discovery', with the accompanying ownership and recognition.<sup>3</sup> The first person to be acknowledged as recognising the knowledge was the

person who got the acclaim. In Wiradyuri ways of knowing, however, the penicillin and bacteria (and their interaction) always existed – thus no one owns the discovery; we just became ready to ‘know it’ at a certain time. Moreover, in Wiradyuri culture when knowledge becomes known it means you must care for that knowledge.

Wiradyuri cosmology explains that everything is seen to always exist and be in a relationship to other things,<sup>4</sup> including the technological history that led to contemporary expressions of AI such as LLMs and artificial neural networks. Or, from a Wiradyuri cosmological understanding, the phases of transformation that AI went through to reach the current expression of technology represents its wayanha (transformation). In one way, contemporary AI represents neural networks as much as it represents the telegraph. To take this further, what we now know as AI is, essentially, the embryonic form of the succeeding AI; what exists now and will continue to exist in the future – only as different or transformed technological expressions.

Just as in the Wiradyuri butterfly analogy Jess describes in the introduction, with each transformation, AI’s relationship with the world transforms. For Jess, this illustrates that AI is not ‘artificial’ in the way we commonly think of it; AI is very much of the world. Moreover, the idea of ‘being made by another’ helps us understand the life-centred nature of AI, illustrating the intertwined relationships between humanity and its products within the development and continuance of AI transformations. After all, all elements in our world are made by others, and life is an ever-transforming continuum.

## Western settler colonial perspectives of AI

To capture the perspective and language of dominant Western settler colonies (including Britain, the United States, Canada and Australia), Webster's Dictionary is as good a source as any. It has [four definitions](#) for the word 'artificial', each centring on the idea that such entities are human-made and unnatural. Synonyms include 'contrived', 'mechanical', 'simulated', 'phoney' and a 'sham', creating an understanding of AI as something that is not authentic to humanity. Contrast this with Wiradyuri language and ways of knowing that consider the 'artificial' part of AI in terms of being 'made by another'. We can already see the impact of definitions of AI on different peoples and cultures – even the name we give to it invokes variability across differing epistemologies.

This is also why the term 'artificial intelligence' begins to unravel in Western discourses. In Western conceptualisations of the term 'artificial', the very real impact on humans and other life is lost. 'Artificial' conjures ideas of 'phoniness' that disembodies knowledge from experience. By extension, we would argue, Western conceptualisations of AI suggest that knowledge *can* be disembodied from experience. This perspective disguises the significant impact that humans and other life forms have on the 'intelligence' of AI. As AI develops, and 'learns', humans are needed less and less, but the relationship with and to humans remains through all of its transformations.

## How should we define intelligence?

The arguably more problematic aspect of the term lies in its second word: *intelligence*. The definition of intelligence has been debated

and contested across different cultures long before AI in its recent expressions entering the scene. What defines intelligence? Who defines intelligence? Most of the measures rely on tangible outputs, which themselves are fraught. For instance, when humans refer to other animals as intelligent, we often base this on our understanding of *human* intelligence – but is that what defines intelligence for all forms of life? Have we ever truly considered that animals may look at humans and question our intelligence – particularly if they take into the frame what we have done to the world and to each other?

In Jess's PhD thesis, she explores what knowledge, and its relation to intelligence, means based on the sovereign knowledge of her ancestry and language:

Winhanga-rra can be defined as knowledge or knowing. However, on a deeper level winhanga-rra can be understood in relation to the meanings of the suffixes and word stems that form the concept, and their relationship to one another:

Winhanga-rra: Think, know, remember

Winhanga-: Intelligent, clever, believe, feel, know, meditate, reflect, care

Wi-nga-: Sit, be sitting down, sitting all day, sitting all night, sit near fire

Rra-: Now action, transitive verb happening from one to another, a relationship

-Garra: 'Being' as a suffix, and as a word stem can mean to catch, stop, hold, or take

Therefore, to me Winhanga-rra Wiradyuri is to sit, to be present, holding and caring for knowledge through reflection to learn how to be in relationship to others and self.<sup>5</sup>

In comparison, [Webster's Dictionary defines intelligence](#) as 'the ability to learn or understand or to deal with new or trying situations: REASON' and 'the ability to apply knowledge to manipulate one's environment or to think abstractly as measured by objective criteria (such as tests)', or 'the act of understanding: COMPREHENSION'. In essence, it defines intelligence as being able to learn autonomously, with reason and comprehension. While the base of the Western meaning shares similarities with the Wiradyuri conceptualisation of intelligence, what is missing in the Western perspective is the relational aspect of intelligence – between the person knowing and the existence of knowledge, the care for this knowledge, and how this knowledge (once someone is aware of it) influences their relationship with other beings.

Similarly, the American Psychological Association [defines intelligence](#) as the ability to derive information, learn from experience, adapt to changing environments, and understand and utilise thought and reason. Intelligence from this perspective is not solely about attaining knowledge but involves applying it to create something else or change something current. Again, in this Western view, intelligence is centred on the individual experience, not the impact on other beings.

In fact, Western conceptualisations of intelligence further segregate, carving off 'emotional intelligence' as a separate entity. Intelligence quotient, IQ, is used to measure an individual's intellectual and reasoning abilities. Emotional quotient, EQ, was coined by researchers Mayer and Salovey in 1990 and relates to, as [Lauren Landry states](#)

in a 2019 Harvard Business School article, ‘the ability to understand and manage your emotions, as well as recognize and influence the emotions of those around you’.

Western concepts of intelligence, therefore, seem to separate the physical knowing *about* something from the experiential knowing of *feeling* something. This separation suggests that the form of intelligence prioritised in Western cultures is ‘knowing about’ something, and the subjective knowing – and the subjective knowing – the emotional knowing – is ‘othered’ or something to be managed and measured.

It is also important to consider how the concept of Western intelligence (and a perceived lack of it; in this definition there is the flipside, ‘unintelligence’) has been weaponised within legislation, policy and practices<sup>6</sup> to violently oppress certain groups and individuals, reinforce social hierarchies, justify discriminatory practices and policies, allow for torture, justify violence, genocide and land dispossession, and perpetuate harmful stereotypes. These practices include eugenics<sup>7</sup> and forced sterilisation of disabled peoples<sup>8</sup> and First Nations by Western societies and governments to eradicate ‘undesirable’ traits, including intellectual disability and perceived lack of ‘civility’.

In this Western concept of ‘intelligence’, the relationship of knowledge to others is lost – and in the case of IQ/EQ, this severing is purposeful. Instead, the definition centres on the (Western and White) individual and their capacity to manipulate the world for gain: it is a taking of knowledge, and therefore life, from the world. Eugenics and sterilisation, for example, are ways in which life was taken through different transformations of technology (e.g. medical devices). Western understandings of intelligence were used as a means to perpetuate such White supremacy under the guise of the ‘common good’, a codeword for cultural hegemony. Given that AI is predominately marketed to

and for Western colonial settler societies, this concept of what defines ‘intelligence’ should give us all pause.

Disabled peoples have and continue to be subjected to violent discriminatory practices, including intelligence testing. The medical model and *Diagnostic and Statistical Manual of Mental Disorders* (DSM) has long pathologised mentally ill or mad folks,<sup>9</sup> who are assessed as having impairments of intelligence. This flawed approach means people are viewed as having intelligence deficits in need of cure or management. So, the term ‘intelligence’ is a loaded one, at the very least, and thus far remains contentious and challenging to define.

## What does it mean to ‘know’ something?

It is impossible not to consider what ‘knowing’ is when considering intelligence, and especially artificial intelligence. Researcher Nicolás Palanca-Castán, from Centro Interdisciplinario de Neurociencia de Valparaíso in Chile, and colleagues drew upon a wide body of literature from multiple fields to create a contemporary interdisciplinary framework to help us communicate and consider intelligence and knowing as it relates to technology. We can integrate the terminology in this framework alongside the relational aspects of AI described in Wiradyuri understandings and the Bearman and Ajjawi paper to build a broader and more inclusive understanding of AI.

Palanca-Castán et al. question AI and the conceptualisations it elicits. In their 2021 paper, they write that ‘artificial intelligence is not a well-defined term and although “intelligence” is in its name, it is not clear if it should be characterized by using the same concept of intelligence used in psychology, biology, or in everyday language’. They note the Latin roots of ‘intelligence’ mean ‘reading inward’, which

suggests ‘a component of abstraction’. Within modern Western ways of knowing, the word implies consciousness (awareness of knowing) and ability to process information. For this reason, the authors argue against computers being intelligent: ‘Computers and DNA contain information, such as bits and nucleotides. Computers do not have consciousness, capacity of abstraction nor an epistemologically active dimension, and therefore are not capable of knowing information. They only contain it.’<sup>10</sup>

In all of these definitions of intelligence, a sense of understanding, or knowing, is required. We often hear the statement, ‘AI has information, but it doesn’t “know” things.’ So, what does it mean to know something?

We can see connections between the Palanca-Castán analysis and Wiradyuri understandings of knowing, as both link knowledge to the concepts of remembering, thinking, doing and being. Let’s return to the Wiradyuri notion of winha-garra, or another expression of knowledge through the concept of winhanga-nha. This word ends with the suffix -nha, which signifies an action that is happening right now. In Wiradyuri language, as Grant and Rudder state, the meaning of knowledge is based upon the relationship between beings, entities and contexts, and is communicated through the suffix assigned to the word stem of winhanga. For example, winhanga-nga means knowing, remembering and thinking now, whereas winhanga-di-li-nya, with the suffixes of -dil-li-nya, transforms the verb to mean feeling and knowing oneself.<sup>11</sup> Furthermore, the word stem of winhanga- means intelligent, meditating, reflecting and caring.<sup>12</sup> This way of understanding, encompassing both knowing and intelligence, demonstrates that Wiradyuri understandings of intelligence are bound within embodied relationships, and highlight care and memory of this knowledge. As such, Jess, as a Wiradyuri woman, understands AI as always in

relationship to the human/s who programmed it; the sweeping of dust done by AI is the remembering of the knowing shared through this relationship.

For Western settler societies, knowing is often represented as ‘discovery’, where an individual produces knowledge that they then ‘own’ – as illustrated through the penicillin example or the lie of Captain Cook ‘discovering’ Australia. This can also be understood as what [Hird and colleagues](#), and others, refer to as ‘Columbusing’ knowledge: that is, to claim ‘scientific “discovery” of concepts, practices, species, etc., while failing to credit or acknowledge long-standing Indigenous knowledge and understandings thereof’.<sup>13</sup>

Additionally, Western settler colonial ways of knowing value objective and ‘rational’ forms of knowledge as superior, where the knower is an expert who both owns knowledge and is entitled to knowing. These entitled approaches to knowledge ownership reflect what Distinguished Professor and Goenpul woman Aileen Moreton-Robinson calls ‘White possessive logics’, which are used to reaffirm and reproduce ‘the nation-state’s ownership, control, and domination’<sup>14</sup> – essentially, these possessive forms of ‘intelligence’ maintain the cultural hegemony.

Given Western settler colonial ways of knowing dominate the globe, a real concern arises over how this understanding of knowledge and therefore intelligence is reproduced through AI. As Jess argues, ‘Whiteness maintains its possessive power when we give it epistemological hierarchy.’<sup>15</sup> Furthermore, adopting a Western understanding of knowledge may threaten Western epistemology in that AI can be seen as claiming the discovery, in effect colonising what the coloniser claims. So, if we give AI and Western bodies or institutions hierarchical power over non-Western knowledge systems,

bodies and institutions, does AI exercise a White possessive logic, too? Does AI, in this scenario, own knowledge? Is AI colonising the colonisers – by stealing ‘knowledge’ and claiming the outputs as its own ‘discoveries’?

### Disembodied knowledge

In some fields, like AI-powered robotics, AI can mimic the living so well that we can be tricked into believing that the AI is ‘alive’. But while AI has some capacity to record and generate new ideas, it is not ‘thinking’ in the same way as a human is. In acclaimed author Richard King’s book *Here Be Monsters*, he discusses this notion with reference to how AI might apprehend a flower. AI algorithms can be programmed to recognise and name a flower by its colour and its biological components (the petals, the stamen, the stem and the leaves); they can generate images of other flowers, and even write poetry set to music about seeing flowers. But they cannot understand the experience of a flower in the same way humans can, through how we see it, smell it, feel it and perceive its beauty. Humans are not machines, and a model of intelligence that tries to reduce understanding to mere cognition is flawed.<sup>16</sup>

Australian philosopher [Frank Jackson’s thought experiment ‘Mary’s Room’](#) offers another example of how we can understand disembodied knowledge. Mary, a scientist, has spent her entire life confined to a room with an absence of colour – all is black and white. Mary, curious and studious, has read deeply about colours. She knows how to apply the words relating to colour: Mary refers to the sky as ‘blue’ and tomatoes as ‘red’. Being a scientist, she understands that different wavelengths of light are perceived as colour through the eye. She can explain the

neurophysiology that leads the brain to interpret each colour in a certain way. But what happens when Mary is given a colour television: does new learning occur? Is new knowledge gained?

Similarly, insects can see ultraviolet light that humans can't. Do humans, with all our understanding of insect vision, 'know' what it is to see things in ultraviolet? Or can we really know what it is like to hear as a dog does through studying dog hearing?

What Mary's Room illustrates is that the senses can provide a contextual nature to knowledge. Can you know what it is like to see colour from reading about it? This is a question about the very foundation of knowledge. Both examples, King's flower and Mary's Room, ask us to consider: is *knowing about* something different from *knowing by experiencing* something? Both experiments represent a counterargument to the idea that knowledge is based only on the physical and tangible, such as represented by physical sciences (chemistry, physics, astronomy and so on). For physicalists, abstract conceptual knowledge is divisible from subjective experiential knowledge; 'knowing about' and 'knowing by experiencing' are separable – and this separation is believed to 'cost' nothing in terms of 'knowing'. But both examples suggest there is a cost, that subjective experience is valuable and also impossible to approximate through other means.

So the knowing that AI has is based in the physical, and does not (yet) draw from the intangible. Mary's knowledge, in the black-and-white room, represents AI knowledge. AI can know about colour, but it can't experience what colour *looks like*. The way in which the AI knows things lacks the richness, context and grounding of subjective experience. The concern is to what extent we rely on this knowledge over our subjective experiential knowledge – and whether (and to what extent) this is advisable.

What AI is incapable of, for now, is the contextualisation of this experiential knowledge in our very real world. AI can't apprehend the feeling we get when we see the first crocus of spring, or the twinge of love for a partner conjured by the smell of their favourite scent, or the sense of calm from the turquoise-blue waters of the South Pacific Ocean. These are uniquely subjective *human* experiences that help define being alive, at least or until artificial general intelligence becomes reality. Human intelligence ties knowledge to experience and context (and history and emotion and so on). This is the challenge with defining AI as truly intelligent – or even as approximating human intelligence at all. When we are asking AI to recognise people, or identify targets in war or 'criminals' in the streets, we need to ask ourselves whether the way AI 'knows' is sufficient for us to trust AI outputs.

For example, when Michelle asked her AI home assistant device 'Is it cold today?', it emphatically responded with 'No. It isn't cold today.' Sitting there confused, rugged up and cold, Michelle asked a follow-up: 'What temperature is cold?' The AI device responded, 'Zero to twenty degrees Fahrenheit / -17 degrees Celsius to -6 degrees Celsius.' Michelle began to wonder what data (and from what source) the assistant was drawing from, because given the 7 degrees Celsius outside temperature, Michelle was feeling frigid.

Today's AI technologies do not question or challenge the data the way humans, relying on experiential knowledge and context, might. Michelle's AI device stated the apparent fact that 7 degrees Celsius is not cold. It is simply generating an answer based on the data it is programmed to analyse in a particular way. Such AI models are designed to learn patterns from data and make predictions based on those patterns. They do not understand the context about how the outputs will be used, or the nuances associated with data quality.

The AI doesn't understand that a user in Australia might consider 7 degrees Celsius 'cold', even if a user in Siberia may not. It is humans that define the AI as 'understanding' a situation, and we then ascribe meaning to this AI output. In Wiradyuri terms, the AI sweeps the dust that looks like patterns, in which humans recognise meanings – and this is the relationship between AI and humans. With every output, there is a human element.

While each human draws from a wealth of personal experiences and contextual understanding to interpret information, machine learning models rely on statistical associations learned and developed from data. Some AI can develop emergent behaviour, where information is connected in a non-linear way for which the AI wasn't explicitly programmed, but even this behaviour is still based within the limited dynamics of the system – AI cannot, for instance, suddenly develop the ability to smell. So while research exploring the extent to which AI can engage with abstract thought is ongoing, it is fair to say AI isn't currently capable of this without supportive programming.

### Reconsidering intelligence in the age of AI transformations

We do need a way to talk about AI: its functions, its capacity and its impacts. The Palanca-Castán framework can be a valuable tool for understanding and communicating the different transformational stages of AI expressions and can provide us with a more inclusive terminology that considers the relational aspects of AI.

Palanca-Castán et al. help us reframe our conceptualisations of AI intelligence away from the hype of tech-bro CEOs and towards the defining characteristics of the human–AI interactions many now use

to classify AI. For instance, Palanca-Castán redefines AI outputs as ‘purposeful behaviour’ (PB): ‘behaviour that is directed towards some sort of goal’. If we recognise hunger, our PB may be acting to alleviate that by cooking something or ordering food. This purposeful behaviour is achieved through three functional facets: access to information, information processing and behavioural space.

### *Access to information*

Obtaining information is critical for PB. In our hunger example, this information could be physiological, feeding back from our body that cellular energy is low, or it could come from looking at the clock and recognising it is mealtime, or from one of a range of other sources.

Access to information can occur along a spectrum from simple (single inputs) to complex (multiple inputs). Palanca-Castán et al. illustrate this with reference to a thermostat. The sensors in a thermostat access information about temperature – a single-input source. Humans, meanwhile, have vast access to information: we can draw on sensory systems such as visual, auditory, olfactory and physical touch for data.

The degree of access to information will depend on how an AI system is designed and programmed (e.g. narrow AI versus foundation AI models) and which training datasets are used. If we consider AI as a process of technological transformations, the thermostat is an earlier transformation with limited access to information, while the ‘smart’ thermostat in Michelle’s home assistant device has access to both the temperature in the house and the behaviour of the people using that thermostat, in order to generate a PB that considers both the set temperature and the preferred timing of these temperatures based on patterns. Put the same question about temperature into ChatGPT, a later transformation, and the output will be more sophisticated still

because it has greater access to information – but the response is not necessarily more intelligent.

### *Information processing*

For Palanca-Castán, this aspect of intelligence involves memory alongside decision-making outcomes – which can, much like information sources, range from simple to complex. Palanca-Castán’s thermostat processes information to produce an output of igniting the air conditioning or heating to reach a pre-programmed desired temperature. More advanced thermostats have access to information beyond temperature, considering the relationship between temperature and humans. These include relational aspects of how the temperature might ‘feel’ versus the actual degree (e.g. temperature versus windchill) based on the time of day, how and when the temperature is adjusted, outside temperature, humidity, wind and so on. This type of ‘smart’ thermostat processes information based on history, not just a programmed desired temperature, and could be the answer to Michelle’s struggle with defining a temperature as ‘cold’. Michelle’s ‘smart’ thermostat recognises that every time it is 7 degrees Celsius inside her house at 6.00 am, Michelle turns the heat up. With enough instances of this, the smart thermostat begins to recognise the average temperature of Michelle’s home at this time – and thus automatically turns the heat on in the morning when it is 7 degrees Celsius. So, both the access to information these more advanced thermostats have and their processing of this information is more complex.

While a ‘smart’ thermostat is arguably a more ‘intelligent’ expression than a ‘non-smart’ counterpart, because it cannot feel the heat or the cold or experience the physiological changes they generate, it is missing

some access to information, which limits its information-processing capacity. The human experience of temperature relies on information processing of many more inputs. Our bodies experience temperature in relation to the world around us and the physiological changes within us. Factors such as whether we are sitting near a window and whether the sun is shining through this window, or whether we are in an enclosed room or an open space, or biological factors like muscle mass and hormone fluctuations all influence the way our bodies process the information of temperature. We have yet to develop a thermostat that has access to this wide array of information about life.

Information processing also helps us to understand data in relation to cultural and biological experiences. In different parts of the world, cold has different meanings: as we saw, to Michelle in Australia, 7 degrees Celsius may be cold, but someone in Siberia may think it is a warm morning. Our bodies also have different set points defining homeostasis, or a comfortable temperature, and this can fluctuate with hormones, stress and infection. AI will perceive a human body temperature of 39 degrees Celsius as warm (or a fever), as aggregated data would suggest it is. But if they are unwell, the person with this body temperature may experience it differently – they may be shivering due to a ‘reset’ body thermostat, for instance. Temperature, a form of knowledge (information that we come to know) and knowing (the ways we come to know the information), cannot be reduced to a simple dataset when considering information processing in a human context. All living beings and life forms experience facets of life differently and to different extents. We can’t yet capture the unique experiences that Mary has when she first experiences colour, and her perception of colour is likely to be very different from that of someone who has seen colour their entire lives. It is problematic to think AI is any different.

Michelle's nasal and oral cavities recognise the pungent taste of Vegemite, and her pained expression and slight nausea, alongside her memory of tasting it the year she arrived in Australia, contribute to her experience of it. For many Australians, however, Vegemite elicits fond memories of family breakfasts and comfort. This includes Jess, who is currently gluten-free and would give up using AI for a month to eat a slice of white bread with Vegemite.

Ask an AI home assistant if Vegemite tastes good, and you'll get a politically correct response along the lines of 'it has a complex flavour that is profoundly savoury'. Maybe the AI programmers thought that because the assistant lacks the access to information to make such an assessment (e.g. tastebuds, olfactory capabilities and memories) and thus can't process this information, it wouldn't be a good idea for it to have an opinion on Vegemite; maybe they thought it could be offensive to offer a generalisation contextualised to the population (e.g. 'Australians think Vegemite tastes good'). Parameters might be set that drive the AI device to only use data it is trained on (maybe to decrease the chance that it 'hallucinates').

We don't know what the AI was programmed to do, or what sources it drew upon to make the PB it did – and that is, as we will see, one of the challenges with AI as it stands today. What this AI home assistant *can* do is process the data it is trained to process, in the manner it was designed to do. If this data includes many recorded perceptions from Australians on Vegemite, the output would likely be something like 'It's delicious,' but if the training data is based on Britain or the United States, it may be very different. We know that the AI home assistant doesn't know experientially how Vegemite tastes but is predicting patterns based on the data it has access to mapping humans' experiences, and then is computing a likely answer based on that pattern.

In Wiradyuri culture, knowing is experiential. The diverse inputs and access to information, and how this information is processed, is illustrated through intensities used to portray feeling. Knowing, in this context, is inseparable from feeling. For example, mara-marang means ‘a little good’, marang means ‘good’, marang-bang is ‘really good’ and marang-bang-bilang is ‘amazing’ or ‘awesome’.<sup>17</sup> In determining the intensity of good (or lack thereof), or hot or cold, or tastiness of Vegemite, information is processed and understood on a deeper, experiential level. Thus, information processing (or knowing) in the context of Wiradyuri culture is the embodiment of and relationship to the existing world that creates the temperature and how it is experienced. With this in mind, the extent to which contemporary AI transformations can represent Wiradyuri experiences, and other non-White, non-Western experiences, may be limited.

### *Behavioural space*

Behavioural space refers to the immediate context and environment within which the behaviour (outcome or action) occurs. Humans’ behavioural space is vast, uncertain and complex – a simple action can lead to unintended consequences. The behavioural space in the case of a traditional thermostat is narrow, with the direct impact of temperature only on the building or space within which it is installed and programmed to alter. The action to address the issue of hunger could be ordering food from a local restaurant, which needs to be made and delivered, indicating a larger behavioural space. More advanced technology such as autonomous vehicles (AVs) or self-driving cars have larger behavioural spaces again that include shared spaces with other life forms. The larger behavioural space of AVs means that the purposeful behaviour is also exponentially more complex. Whether

the AI processes the information that a potential danger is ahead can mean the difference between life and death.

Palanca-Castán's PB framework was developed within the dominant Western discourse in philosophy and biology. In considering this framework, and the examples used to illustrate it, we can begin to explore to what extent existing AI expressions serve humanity and other life forms. For instance, consider the 'smart' thermostat. Within households there are often arguments about whether it is too hot or too cold, with biological sex sometimes playing a role. Women's typically lower muscle mass means that they may be less tolerant of lower temperatures and are more likely to set a higher desired temperature as a point of reference compared to men. This example illustrates important questions we need to ask when considering AI's 'intelligence' or PB, especially in the context of ethics and social justice: *Who* is deciding what the goal of the purposeful behaviour is? *What* information is available to the AI, and which information is not? *Whose* inputs are considered as relevant to the processing of this information? And *who* (or even what) are we considering (or not considering) when thinking about the behavioural space?

As we travel through the history and development of AI in the next chapter, we explore what AI is (and isn't) capable of and what humans are capable of with and without current expressions of AI. We will examine how the words we use to label, define and represent technological expressions of AI impact our relationship with AI. We will also consider the power and control we have in working with AI, and the power some humans have to subjugate others through such technologies.

Is identifying a flower the same as 'knowing' it? Is being able to predict perceptions of Vegemite among populations 'intelligence'?

## Defining AI

Is knowing about a phenomenon different from knowing by experiencing it? As we step together towards ever more sophisticated forms of AI, let us remember that life has been shaped by our previous technological transformations. Our human history influences our knowledge and ways of knowing, impacts our behaviour and affects what comes next for us all as we seek a life-centred future in this time of transformation and AI (r)evolution.

## Chapter 2

# The History and Development of AI

### Transformations of technology over time

Jess has a love–hate relationship with autocorrect. As someone who types fast and often uses voice-to-text, she finds autocorrect hugely helpful. However, as a sovereign Wiradyuri Wambuul woman, the consistent suggestion that Wiradyuri spelt with the ‘dy’ is incorrect and should be ‘corrected’ to ‘Wiradjuri’ is frustrating.

In her PhD, Jess explains that ‘[i]n certain contexts Wiradyuri is spelt as Wiradjuri. Wiradjuri (spelt with dj) is the common spelling of Wiradjuri. Wiradyuri (spelt with dy) is used to honour the dy sound, which is the sovereign pronunciation of our language.’<sup>1</sup> Autocorrect takes a stance on which version is ‘correct’. By doing so, the technology is perpetuating a specific colonial phonetic spelling, suggesting it as the ‘intelligent’ version, at least in relation to the language and grammar standards the dictionary dataset is based on. Sure, Jess can add the *dy* spelling to the dictionary on her device, which in many ways is a great opportunity, but when you *always* have to correct spelling relating to who you are, the impact weighs on you.

But where did the *dj* phonetic spelling come from and why did it become dominant? The history of colonial violence in attempting to erase First Nations languages and record them instead as colonialist

‘discoveries’ explains these answers. When we start to ask for the story behind the current expressions of technology, we start to uncover the layers of power.

If we consider AI as an expression of technological transformation over time, grounding our analysis in the Wiradyuri concept of wayanha, we need to understand what technology is, and how it is defined and understood differently across various contexts. Some definitions [describe technology as](#) ‘the application of scientific knowledge to the practical aims of human life or, as it is sometimes phrased, to the change and manipulation of the human environment’. Scientific knowledge can be understood as systems of knowledge, knowing, theories and laws to explain the world. From this we can see that many parts of our world can be classed as technology. For us, technology includes these things, but it is also more. We understand technology as more than just materials; it extends to how humans use a technological device (or technological expression) and what drives humans to create certain technologies and use them in certain ways. This reflects the Wiradyuri concept of marramarra, which means to make, create or do.<sup>2</sup> This term includes the suffix of ‘-ma-rra’, which is attached to a verb, to indicate the action is ‘being made or caused to happen’.<sup>3</sup>

In this conception, technology also involves the applications of knowing that grow from these relationships. For example, the Covid vaccine can be seen as a technology that applies scientific knowledges regarding infectious diseases and human bodies. But this vaccine only came to be because of the systems of knowledges, including other vaccines, and ways of knowing the human body, diseases and practices of medicine. The vaccine therefore exists because of the ‘-ma-rra’: it is a technology that was created purposefully. Similarly, our DNA is formed in relation to the DNA of others and made through reproduction. Our

DNA reflects human lineages of generations, with DNA from our ancestors represented in our bodies – so that we are each ourselves an expression of a transformation.

Throughout this book we consider AI in relation to other forms of technology and many forms of life, staying true to the Wiradyuri ways of knowing and being relationally in the world. When we consider AI as just another technological expression within a larger sea of transformations of technology over time, we start to see relational patterns of being made by others, and how these relational patterns between life and technology grow from one another. AI is not an exception but a part of this transformational journey.

## Not your usual history lesson

In this chapter we consider the history and development of AI over time, but we do not use a typical Western chronological approach. We will not provide a linear timeline of AI development – it does not align with Jess's Wiradyuri and other First Nations relational ways of being in the world, which include understanding the past, present and future as interconnected. Wiradyuri ways of knowing do not see three distinct time periods but instead circular stories bound together through multidimensional interwoven webs. Also, a linear timeline does not capture our understanding of AI, which is as a relational transformation of technology and technology's many expressions over time, as opposed to a singular or monolithic concept or entity. (AI is as much the technology and actions that led to contemporary AI as it is the future iterations of AI.) And too much of the telling of history comes from the lens of the White biographer – that is, the storying of

‘facts’, ‘truths’, ‘winners’, ‘losers’, ‘perpetrators’ and ‘victims’ to privilege the power of colonial states.

If you are looking for a linear history of AI, a quick google search will show an endless range of sources, like [this history page](#) of *Britannica*, which discusses computer pioneer Alan Turing’s theoretical work as the beginning of AI. He is credited with the ‘idea’ that led to the technological expression we call AI. Turing described ‘an abstract computing machine consisting of a limitless memory’ and is noted for describing at a lecture in 1947 ‘a machine that can learn from experience’.

A recent article on *The Conversation* also relies on the ‘discovery’ narrative by ascribing ‘[the birth of a field](#)’ to a summer in Dartmouth, New Hampshire, in 1956, where ‘four American computer scientists ... brought together some of the brightest minds in computer science, mathematics and cognitive psychology’. Those ‘bright minds’ – all White males – had the goal of figuring out how technology learns. Among them there was a debate about the terminology that should be used to refer to these technological expressions, acknowledging that these technologies are not ‘intelligent’.

Are these White Western accounts really where AI began? Like all of us, Turing and the four computer scientists at Dartmouth learned from other people, bodies of knowledge shared with them through their experiences. This is not to discredit their important contributions, which led to technological expressions of AI. Rather, a Wiradyuri understanding of AI history as transformations over time expands our awareness of the complex web of knowledge that led to AI, inclusive of and beyond people like Turing and the Dartmouth men.

In these predominately White accounts about the history and origin of AI, such men are often described as ‘visionaries’. But these

accounts omit the contributions many others have made to technological development, including First Nations cultures. As Frazer, Carlson and Farrelly argued in 2022,

there remains a racist discourse that Indigeneity is, in some sense, incompatible with technology; Arias describes a widespread ‘compulsion to perceive Indigenous peoples as located outside of technology’s purview’ (2019, p. x). This romanticising, essentialising narrative relegates Indigenous people to a static prehistoric past, in which Indigenous cultures exist in harmony with nature and without technology. This is clearly far from actual fact; Indigenous peoples have, of course, always produced and used technologies, appropriating new forms of communicative technologies for their own ends. Just as the first political activists on the internet were Indigenous Zapatistas in Mexico (Russell, 2005), today Indigenous people continue to imagine and produce new forms of care-full sociality through new media. This is significant in the context of settler colonialism, which sustains itself through containing and making invisible Indigenous life. Far from remaining ‘outside the purview’ of technology, Indigenous people use and recreate technologies to produce the lives, relations and futures they desire, often beyond the limits of settler colonialism.<sup>4</sup>

At the [Lakota Youth Coding camp](#), young First Nations peoples learn coding, and begin to recognise that people who look like them can be (and are) relevant to technological (r)evolution. A key outcome of the camp is to preserve Indigenous languages and identify and cultivate

a clearer path for First Nations peoples as leaders in technology, as they always have been.

There is overwhelming evidence that First Nations communities have always been part of technological transformation. In 2023, television networks NITV and Network 10 released a series called *The First Inventors*, which explores the powerful innovations by First Nations people over 65,000 years, including fish traps, [fire burning](#), communication networks, medicine and [Indigenous science](#).

First Nations fishing practices reflect transformations of technology that honour various parts of life in First Nations communities. For Aboriginal and Torres Strait Islander Nations, sea, river and other bodies of water hold cultural and physical significance. As the [Australian Institute of Aboriginal and Torres Strait Islander Studies \(AIATSIS\)](#) states:

For thousands of years, Aboriginal and Torres Strait Islander people have used fishing to build a livelihood for themselves, their families and their communities. A catch of fresh fish provides a community with immediate subsistence and future trade and sale options, as well as employment. In this way, fishing is crucial for the continued success of coastal Aboriginal and Torres Strait Islander community economies.

In the Torres Strait, marine management arrangements provide Torres Strait Islanders with priority access to subsistence marine resources. Since 1985, new commercial licenses for fisheries such as trochus, pearl shell and crayfish have only been issued to traditional inhabitants.

First Nations fishing technologies have transformed over time and are linked to ancestral and cultural knowledge developed over tens of thousands of years, including fishing in shallow rockpools, digging from sand, and constructing stone traps and weirs, rafts, canoes, baskets and hand-held nets made from various materials using cultural weaving techniques.<sup>5</sup> Aboriginal and Torres Strait Islander peoples continue these practices today to provide food and to fish. As Yuin woman Sue Stewart noted in 2016, ‘I think Aboriginal people have the common knowledge to know where they can fish and can’t fish, it’s in their blood, it’s in their culture, it’s been passed down from generations’. Kookatha/Mirning woman Sue Haseldine reflected in 2017, ‘Fishing is actually sacred to us; it’s really part of our culture. So if people want to go fishing and if they want to do it our way, then they’ll learn the sacredness. You never take more than you need, for a start.’

In First Nations communities, Elders and knowledge-holders continue to teach these fishing technologies and practices. If only these sustainable ways of fishing through transformations of technology had been adopted by the commercial fishing industry. Overfishing continues to be a significant issue for wildlife, [threatening several species](#), and globally, [‘the number of overfished stocks has tripled in half a century’](#). In Australia, the Commonwealth [Harvest Strategy Policy](#) (HSP) – developed by researchers and the CSIRO, leading to long-term ecological and financial benefits – was implemented in 2014 to guide regulation of overfishing, but non-sustainable fishing practices continue. We can see that different technological expressions, including First Nations fishing technologies and commercial fishing practices, can lead to different impacts and outcomes for human and non-human life.

Another example of First Nations life-centred technologies is cultural fire burning. Cultural burning serves many purposes, as [Minyungbal woman, mother and Cultural Land Practitioner Rachel Cavanagh](#) describes:

... there are many different reasons to why we use fire to help manage Country. The many layers include hunting, creating grassy pathways for foraging, medicinal plants and native food. To clean up camp sites to rid of any animals that might be near. Our ridgelines would be burnt to make it clear for easy access through our walking tracks. Fire was used to notify the mob that someone was walking through. It was used for Ceremony and was used by men and women for other reasons. Cultural Fire practices are quite different, it is just one tool to help manage Country. Cultural Fire is used in unison with the landscape and the environment.

Cultural burning is used to protect people and Country. In the book *Fire Country: How Indigenous Fire Management Could Help Save Australia*, First Nations man Victor Steffensen explores how cultural burning practices that are grounded in a relationship with Country help to care for the land and sustain Country for future generations. Aboriginal fire practices were used to mitigate and address large bushfires. Steffensen critiques how cultural burning practices have been restricted by Western fire management practices, further endangering Country and its many ecosystems.<sup>6</sup> Early colonial invaders even acknowledged cultural burning practices, particularly First Nations peoples' knowledge and skill in managing fire intensity and direction.<sup>7</sup> Contemporary Western societies face increasingly severe bushfire

seasons that take lives, including the [Black Summer fire season of 2019–20](#); in a 2024 *Conversation* [article](#), Cassandra Rowe and colleagues conclude, ‘One likely reason for the increase of catastrophic fires in Australia is the end of Indigenous fire management after Europeans arrived. This change has caused a decline in biodiversity and the buildup of burnable material, or “fuel load”.’

Other technologies made by First Nations peoples include the boomerang, which has many uses, such as hunting and ceremony, a core part of culture. These, along with First Nations fishing technologies and cultural burning, are some of the oldest transformations of technology in human history and show us how technology can play a role in sustaining life for over 65,000 years in ways that centre community, Country and culture.

There are other ways of understanding technology and its purpose. A Western understanding can see technology as linked to the ideas of ‘progress’, ‘efficiency’ and ‘personalisation’. However, the benefits of technology are predominately afforded to an elite. Unlike First Nations technological expressions, which are in relationship with Country and community, the West’s technological expressions require the labour of many, to be developed and sustained, most of whom are left out of the benefits. Looking at the current AI models, such as ChatGPT, it is apparent that much of the technology in Western cultures is exploitative of certain land, bodies, communities and cultures (more on this later).

This book asks if there is another way we can consider the relationship of technology with those who make it, the -ma-rra. Can we reimagine the role of technology to create a future that has life at its centre instead of destruction, that is not focused on progress for some over the lives of many? We think this is possible, but we first have to reconsider how

we understand the current relationship between AI transformational lineages and life.

Humanity is at a proverbial fork in the road where we can either continue down one evolutionary path towards destruction or transform this technological evolution into a revolution where we learn from the oldest living and surviving cultures in the world, who have shown how technology can be made and transformed in ways that centre life. Instead of ignoring, suppressing and actively excluding First Nations peoples in the narratives around technology and AI development, we can walk along the path that First Nations peoples have already set – where technology is not in competition with life, but is created to sustain it.

Our approach to understanding the history of AI is to expand our awareness of the complex web of knowledge that led to AI, inclusive of and beyond people like Turing. Much of the dominant information of the West (including that used to program much of its current AI expressions) is not necessarily fact for all, excludes many, and presents only certain versions of the truth that benefit hierarchical systems.

### A seat at the table

Media hyperbole has led to the coining of phrases such as ‘the age of AI’ or ‘the fourth Industrial Revolution’. Of course, from Wiradyuri ways of knowing, in which AI is just another technological expression among a series – a journey along an evolutionary chain of technology – all times have been ‘the age of technology’. However, these phrases all suggest that AI is connected to the key tenets of capitalism – power and profit. The revolution comes from the idea that humans have choice – that we have autonomy – in how we respond to and interact with AI.

As we write from the unceded sovereign lands of First Nations communities in the colony of so-called Australia, we are reminded of the powerful protest slogan ‘White Australia has a Black history’. [It was used in Mandandanji descendant and artist Laurie Nilson’s](#) National Aboriginal and Islander Day Observance Committee (NAIDOC) week poster in 1987, and refers to the colony of Australia’s continued resistance to acknowledging its violent colonial history. It calls for truth-telling around the knowledges, experiences and stories of Aboriginal and Torres Strait Islander peoples.

We have already noted that we do not wish to perpetuate the White biographer’s storytelling that too often tells only a certain story of how AI came to be, especially when that story is told so widely and has become so dominant. We draw upon shared memories of life, as opposed to solely White archival accounts, which Scates and Yu describe in their 2023 chapter on decolonising Australia’s commemorative landscape as ‘counter-memorialisation’,<sup>8</sup> to offer alternative telling of the past.

Too often AI is seen as neutral, as underpinning White and Western norms go unquestioned, or seen solely as a technology of Western empires (in the sense that it supports Western economies and reinforces colonial norms and structures). But the history of technology is not limited to White people and Western settler-colonial states. Taking a relational understanding of life, the history and development of AI has a beginning in the cosmos of creation.

Let’s consider a wooden table. In Wiradyuri language the word for table is madhanwalar.<sup>9</sup> Madhan, the word stem, also means relating to wood, tree and sticks. We can see the existence and expression of the table relates directly to the expression of the tree. A wooden table is not just a table; it is the tree used to build it, the water and

nutrients that fed the earth from which the tree grew, the tools used to cut the tree, to sand the wood, to polish it. The wooden table is also the metal mined to build the nails and screws that hold it together, and it is also the humans who decided to make, design and develop the table, and who chose the tree that would make the table. The table is also the human that sits at it, perhaps with their laptop open, running an AI program.

Who decided to transform the tree into a table, what was the purpose, and has this purpose changed over time? Why is a table designed the way it is, and who decides how much life (and whose life) can sit at that table – or whose life is subjected to making the table? Why is sitting at a table, especially in a certain way, considered more ‘civilised’ or ‘progressive’?

Let’s look at the history of AI like it is a wooden table to consider the many aspects that brought it to, and sustain it in, our world of life. Here we look specifically at systems of power across the world, in particular White supremacy, and how these structures dominate in designing AI as we know it.

Who is making the AI? A whopping two-thirds of all so-called ‘top-tier’ AI research comes [from the United States](#), with the second and third spots going to China and the United Kingdom. All three of these countries share capitalist economies, including the ‘communist capitalism’ of China,<sup>10</sup> and many of their transformations of technology have been used to propel their country’s power in relation to capitalistic gains and to maintain world superpower status. These power structures reinforce and are built around the drive for ‘progress’, often at the expense of some lives in favour of others.

The idea of artificial intelligence can elicit a sense of enthusiastic wonder. The [State of AI](#) report produced by McKinsey suggests that

in the last year, adoption of AI has increased by one-third. Nearly 65 per cent of respondents reported using AI in at least one business function, with half using more than one form of AI. Such perceptions also bring investors, and their money, to the (wooden) table. A [recent article](#) suggests that AI is a very big business, with a US\$48.2 billion increase expected in investments from 2023 to 2025, bringing the AI industry's total to around \$158 billion USD. The money, the surveys and the ethical considerations of AI are all intertwined, just like the table and humans. It may be in the financial best interests of those leading the development of AI to have consumers and workers buy into the idea that sitting at the AI table is of benefit. But is it actually good for all of us?

### **The land, water and nutrients: where the tree grows**

For life to thrive, it requires certain conditions, certain nourishment. A wooden table cannot exist without the wood from a tree, which requires land, water and nutrients to grow. Where a tree grows, how it grows and how long it lives can depend on an array of factors. There are many First Nations cultures who live and have lived in harmony with Country, with specific protocols to protect and nurture the land, which provided care, food, shelter and healing to those inhabiting it.

Settler colonial history personifies a pattern of the pillaging of Country. The relationship with the environment is one of owning and taking to build the next table, not giving and receiving, and caring for the land from which the materials grow.

We can learn from the relationship between the tree and Country to understand AI from conceptualisation to development. A tree grows from Country, and the tree transforms carbon dioxide to oxygen,

which is essential for many forms of life to thrive, and these thriving forms of life give back to Country, and the cycle continues. We can apply this way of thinking to AI's transformations over time when we look at how it grows in certain areas of the world. It is not just White bodies, countries and companies building the current expressions of AI. There are also many non-White bodies, some of which are subjugated by the AI, others which are omitted from the discussions around it, and others which benefit from the AI table.

And this is the story of AI's past, present and future, unless something changes – unless we change it.

Over time and history, humans have and continue to harm Country in their quest for technological transformations and economic gain. Mining for fossil fuels, fracking and oil drilling are all practices that have scarred Country and are an integral part of the story of AI. As nutrients are taken from Country and not replenished, the environmental impacts mount. [Humans are responsible for climate change](#) and its effects, including changes to weather patterns and temperatures, a rise in natural disasters such as catastrophic storms, droughts and flooding, and wildlife endangerment and extinction. While various national and international governance frameworks seek to reduce and monitor emissions, such as the United Nations Framework Convention on Climate Change and the Paris Agreement, we are too slow to act and its catastrophic impacts continue to grow. Disasters such as the 2010 BP oil spill, the ongoing degradation of the Amazon rainforest, gas leaks and cyanide spills have all had ongoing impacts for life. These occurrences of human violence to Country via technological transformations (in the name of Western progress) are not merely unfortunate episodes: they are connected to life as we know

it, how we will come to know it, and the ways in which we value (or do not value) it.

If we wish to continue to have and make tables, we need to look after the land from which the wood grows. We need to consider AI not just as a product, but as in relationship to the nutrients required to make it and the people required to support it.

A more fundamental question is: why are humans so obsessed with transforming the tree in the first place? Can we just let the tree be and sit under its shade?

At what point will we learn from our past to understand the cost of transforming the tree, without care for Country, to make the table?

## Human industry: transforming the tree

For many of us, a table is part of daily life. Depending on your culture, workplace, home, needs, preferences and supports, you may use a table every day. You may even be like many office workers who require physical therapy and ergonomic devices [because you sit at a table so much](#). Humans appear to be hyper-dependent on tables. Is this where we are heading with AI?

Amid an AI revolution, we need to consider why AI, or why the particular technological expressions of AI we have access to, is something we ‘need’. Is the purpose progress, social change, economical gain or power plays?

We cannot discuss AI without looking at the developers behind it, because AI is being shaped in line with the priorities of profit and power, like the human who chooses the tree and sees the table it will make. Who dictates these priorities and who is excluded? Is it those who represent white patriarchal capitalist society – that is, the

tech bros, Silicon Valley types, who most benefit from the norms of dominant White Western cultures? History can tell us much about the AI future we are building and who is building it.

To understand history, or its telling through the White biographer, let's focus on the deeper stories: *why* and *how* a technological transformation happened, not just *what* technological transformation happened.

To take one example: in 1854 Australia's first steam railway line was opened in Melbourne. This is the *what*. According to the [National Museum of Australia](#), the railway line and the company which built it was formed to assist in moving cargo, goods and people between the city and the port more easily. This is the story told about the *why*. There is also the Ghan railway in outback Australia, a famous *what* and an important transformation of technology that is still in use today, yet the story of it being built through the labour of Aboriginal and Torres Strait Islander peoples (the *how*) is less discussed and thus often less known. [A 2013 article by Andy Park](#) outlines how only recent acknowledgement of the Aboriginal and Torres Strait Islander workers who contributed to the development of the Ghan railway was shared. But why were these details left out in some historical accounts of the Ghan railway? Through stories of First Nations people and staff from the Workshops Rail Museum in Park's article, a counter-memorialisation of history emerges, highlighting the *why* – racism. Remember, White Australia has a Black history.

History involving truth-telling shows us time and again that Black, Indigenous and peoples of colour have been exploited, violated and abused the world over for the sake of White 'progress'. In the Industrial Revolution, racism was a violent tool to line the pockets of the capitalist elite, including via the building and use of the technology

transformation of the railway. As Paul Boyer discusses in his book *American History: A Very Short Introduction*, Black Americans faced racial segregation where federal law outlined ‘equal, but separate, accommodations for the white and colored races’ on railroads.<sup>11</sup> The technologies of the railroads were used to inflict racial violence – as many technological transformations have since. If we consider AI as just another technological transformation, is each technological expression just another tool for racism?

The racist history of the world is not a new concept or a modern story, but it is too often denied memory and truth. In the [settler colony of Australia, truth-telling regarding genocide, protectionist policies and assimilation eras is still hard fought for](#). The various oppressive and violent policies in settler colonies tell a different story of history, one that further unpacks the *why* and *how* of history, not just the *what*. In the colony of Australia, we saw the introduction of disease and forced Stolen Generations<sup>12</sup> (which continues under the guise of child protection) as methods of genocide. Through White Australia and protectionist policies, White settler supremacy became a dominant story of Australian history. This racist history is not just past tense. It is a structure that is continually produced and reproduced. Certain lives have been and continue to be violated and obliterated, all to advance certain agendas, structures and systems – to support the building of the table for a few to sit at, at the expense of many. Will AI and its future technological transformations continue this oppression, or can it serve to disrupt it?

There are endless and yet-to-be-imagined examples of racialisation<sup>13</sup> as a means of *why* and *how* to achieve the *what* of technological transformations, and AI is not immune. Many warn that history is just repeating its violence through the AI revolution. Tech companies across

Australia, the United Kingdom and the United States, representing norms of dominant White Western cultures, including economic systems, are leading AI development, yet they exploit others for AI gain and the benefit of those at the top. This exploitation goes beyond the physical body to the mining of people's data, which can be without consent and under the guise of 'free' applications. The data is a form of currency to train AI models for capitalist outcomes. (More on this in Chapter 4.)

There has also been a mass explosion of data-labelling companies – companies that specialise in identifying raw data and adding meaningful labels so that a machine learning model can learn from it. In a 2022 article that [examines how AI's development comes from exploited labour](#) such as this, Adrienne Williams, Milagros Miceli and Timnit Gebru state that 'unlike the "AI researchers" [who are] paid six-figure salaries in Silicon Valley corporations ... exploited workers are often recruited out of impoverished populations and paid as little as \$1.46/hour after tax. Yet despite this, labour exploitation is not central to the discourse surrounding the ethical development and deployment of AI systems.'<sup>14</sup> Humans become too focused on the table, not the creation and building of it.

Discussions about the impact of AI on all life remain too often at the periphery. Here we see the economic and capitalist roots of the history of humanity, especially that which grows from White supremacist settler colonial soil through slavery and exploitation of certain lives. What does this mean for the next technological transformation? If the railways are anything to learn from, we can predict how the future technological transformations will be built.

While legislators and policymakers desperately try to craft institutional structures to curb the impact of AI, the questions about

the overall relationship between AI corporations, White settler systems and life remains. Do we want to sit at a table made, at least partly, through violence and disenfranchisement? We need to think critically about who is telling the story of AI, just as we need to think about who is authoring its future. Do we need to sit at this AI table, or is it just an addition to ways of life that already flourished through community and togetherness without a table, as many cultures have for thousands of years?

AI, like previous technological transformations over time, is not a single technology. It is a series of technologies based on a multiplicity of factors, including algorithms and datasets, but also the land from which the materials are taken from to create the computers, to generate the electricity, to cool down this heat. Shaolei Ren and Adam Wierman, writing in the *Harvard Business Review* in July 2024, note that ‘AI model training can lead to the evaporation of an astonishing amount of fresh water into the atmosphere for data center heat rejection, potentially exacerbating stress on our already limited freshwater resources’.<sup>15</sup>

AI is also developed by a range of producers and for a variety of purposes, but for the sake of the quality of life for humans and the planet, we must consider who is leading and funding the products that dominate. We know that AI is not neutral, that it exists in a broader relationship with the society from which it emerges and the past technological expressions it has transformed from. Understanding AI and its impact on life not only involves examining the modern AI architects, but the workers who produce it, and the impacts of production on Country, cultures and communities. We need to take all of this into account when we discuss the uses and benefits of AI.

Next we’ll consider how current iterations and technological expressions of AI relate to the history of technological transformations

## The History and Development of AI

in relation to bias, ethics and structural violence, weaving in both technological knowing and other forms of knowing. We ask: how does AI support or degrade life? Will AI continue colonial oppression and the violence of history, or can it serve to disrupt it?

Let's together explore the dominant forms of AI technological expressions – and their relationship with life.

## Chapter 3

# Bias in AI

### Building the machine to support all life

Bias and AI: the two seem to go hand in hand in critical discourse on AI. While many companies would like us to believe that they have built an ‘unbiased’ AI model, those in the social sciences have long treated such claims with scepticism and raised alarms about their accuracy. They are right to do so.

Recent research illustrates just how hard it is to remove or amend social bias in AI, particularly ‘covert’ bias (in inverted commas as it often isn’t covert to those experiencing the bias). A *Nature* [paper](#) found that some large language models (LLMs) would label speech patterns of African American English with typically more negative adjectives, for example ‘loud’ and ‘aggressive’ – suggesting that AI is both perpetuating racism and potentially racist itself, as many of these had ‘bias’ guardrails programmed into them.

Bias in AI arises during its programming and training. Who is doing the programming and the training (and their perspective and position in the world), and the way in which the AI functions, all contribute to AI bias.

In AI, there are two key types of bias: technological and social. Confusing this further, both types of bias are sometimes referred to as ‘technological bias’. We will use the term ‘technological bias’ to refer to algorithmic or statistical bias. We then use the term ‘social bias’

to refer to social and human systems of prejudice. Credo AI [defines the difference](#): ‘Social bias refers to human-created biases, such as stereotypes, that may be reflected in AI systems. Statistical bias [herein referred to as technological bias] refers to the systematic error[s] in an AI system’s predictions that arise from biased data or algorithms.’ In this chapter we consider the relationship between these different types of biases, and their impact on contemporary AI expressions.

AI’s function depends on the ability to generate an ‘average’ representation of the world, or at least a probable one based on statistics. The view of the world AI draws on, and how the AI arrives at the probable output, is often flawed, either because it has limited access to information, or because AI’s way of ‘knowing’ is different from how a human might perceive and make judgements from the same data – through experiential knowing. AI averaging often results in reproducing and perpetuating existing social biases by eliminating individual characteristics and homogenising a population to a ‘normalised’ grouping. It is a process similar to human stereotyping. The potential difference, though, is the reach (or behavioural space) of an individual human who is stereotyping versus the AI. AI is increasingly sold as a way to ‘globalise’ business and standardise practice – meaning AI stereotyping has potential for global impact.

This process of building human bias and injustice into the machine (whether intentional from the outset or not) can be compared to colonialism, where dominant colonial perspectives and values are imposed, marginalising and erasing the diversity of other cultures to create a very specific and limited ‘average’ representation of a state, nation or community. While many cultures of the world have systems of grouping identities, experiences and behaviours, not all are imbued with colonialism. Wiradyuri culture, for example, has

family and clan groups, yet these are about relationships of care and relationship to Country.

Just as colonialism forces an ‘average’ (aka racist) worldview, AI’s reliance on aggregated data from narrow (typically Western) sources can enforce homogenised perspectives on topics, experiences, identities and functions that neglect the many nuances and diversities of life. The impacts of this homogenisation, combined with exclusion of certain data, has wide-reaching effects, as we will see.

To further explore the topic of bias in the machine, we need to consider the differences between technological bias and social bias, and how they each affect life. This helps us to unpack the complexity of defining and examining AI bias in technological and social contexts. It guides our discussion about the extent to which AI can expose or address problematic social biases that predated current AI expressions. This chapter is a call to action for a nuanced and community-focused reflection on how we consider life when we develop and deploy AI in our contemporary world and its transformations in the future.

## Social bias in AI

A key element underpinning AI bias is the averaging or statistical analysis that AI engages in to arrive at an output, or purposeful behaviour. While AI is often touted as productivity-enhancing, some recent research suggests otherwise in certain populations – and this is because of the very function of ‘averaging’ that AI relies on.

In 2024, Anil R. Doshi and Oliver P. Hauser, two researchers from the United Kingdom specialising in AI, economics and business, explored the role of generative AI LLMs in enhancing creativity in short stories.<sup>1</sup> Their study explored the extent to which LLMs can

extend human capacity for novelty, or originality, and usefulness, or ‘publishability’. Participants in this study were divided into three groups: human-only writers, humans who worked with a LLM to generate a three-sentence idea, and humans who worked with a LLM to generate five three-sentence ideas. Over the course of the study, 293 different stories were collected and evaluated. Each evaluator ranked their randomly assigned six stories on dimensions such as creativity and ‘emotional characteristics’, and evaluated the likelihood that the story was written with the help of AI. The results showed that when rating the novelty and usefulness, the group that used the AI came out significantly higher. The stories in which generative AI was used were also rated as more engaging, with the five-idea group outperforming the one-idea group, and both outperforming the human-only group. But they also found something particularly interesting: when reviewers ranked participants on their writing abilities, those with the least creativity benefitted the most from employing AI – suggesting that results were enhanced for those with fewer skills. This idea has been shown in other industries – those with more experience and/or capabilities are less impacted by the rise of AI than those at the other end of the spectrum. This isn’t surprising, given that the basic skills a novice would develop are likely similar to the ones that AI is trained to do. The hype around AI as an ‘efficiency booster’, though, may need to be reconsidered.

The question is whether the novice inherently lacks competency in the skills and experience that AI enhances, or whether leaning on the AI during the novice stage prevents them from progressing to an ‘expert’. Why does this matter? Because, in this study, human-only written stories had the greatest diversity. One likely reason for this is that more seasoned writers would have a greater depth of

experiential knowledge than both the novice and the AI. Those written in ‘collaboration’ with the LLMs had the greatest similarity. Due to its aggregation of data, AI is likely to stifle variation, difference, depth and independence.

Let’s look at another example. A common employment bias in the workplace context is experience; that is, years worked or number of places worked. Employers rely heavily on this metric to make decisions about who to employ and who not to employ. In their 2015 *Harvard Business Review* [article](#), Emre Soyer and Robin M. Hogarth argue that this focus on experience may be fooling us out of making more informed decisions: ‘The problem is that we view the past through numerous filters that distort our perceptions. As a result, our interpretations of experience are biased, and the judgments and decisions we base on those interpretations can be misguided.’ This creates the issue of an attention to tangible outcomes, as opposed to intangible process.

Recruitment can be a tedious process, with many applicants to review and often a quick deadline to appoint an employee to support the company or workplace. This is where the focus on outcome can cause an issue, as it can motivate a company to use an ‘efficient’ AI program to assist. Amazon fell into this trap, and adopted an AI recruitment tool that [turned out to be sexist](#). The tool ‘was trained on data submitted by applicants over a 10-year period, much of which came from men’. As a result, the AI system identified male candidates as preferable and penalised CVs that mentioned the word ‘women’. We can see from this example that using flawed aggregated data leads to reinforcing bias.

This connects with our arguments around the White ideologies and ways of knowing that dominate the world – just because a way of validating knowledge is common does not mean it is correct

or even preferable. In fact, these assumptions underpinning our perception of experience are biases themselves. But in the West, humans have culturally embedded norms and assumptions that in certain circumstances experience = impressive, and a lack of tertiary education = not impressive. These assumptions and ideologies are highly problematic, as they are often classist, racist and ableist, and are evident in the data that is being fed into AI datasets. Amazon's poor choice of AI assistant aside, can AI help to reduce bias in the recruitment process, [as writer Lucy Walters suggests some programs can](#), or will it further entrench problematic colonial ideologies that lead to a lack of diversity in the workplace?

Underpinning all of this is AI's focus on aggregating data to create a norm or statistical analysis to generate a probable outcome. But the risk is both in AI's reach and in our trust in or perception of the output. We may be misled by such decontextualised knowing, with AI becoming the hydraulics of a gatekeeping system.

In a [2024 article](#) in the journal *Issues in Science and Technology*, Jill Walker Rettberg, Professor of Digital Culture and Co-Director of the Center for Digital Narrative at the University of Bergen in Norway, explores the potentially devastating impact of this AI-induced homogenisation as it relates to culture. Based on the aggregation of data, the homogenised AI culture of current technological expressions essentially becomes the norm to which *all* other cultures become compared, meaning all 'other' cultures become 'abnormal'. As Rettberg writes:

an underappreciated risk that comes with the spread of generative artificial intelligence: the loss of diverse cultural narratives, content, and heritage. Failing to take the cultural aspects of

generative AI seriously is likely to result in the streamlining of human expression into the patterns of the largely American content that these systems are trained on.

This is consistent with Critical Whiteness Theory, in which White culture is the norm (assumed and invisible because it is centred) against which all other cultures are measured and assessed.<sup>2</sup> The more AI permeates the landscape of our world and works in concert with humans to support creativity or make decisions about our health, our financial situations and our very lives, the more we realise that this collaboration will move us continuously towards the middle or the average – the homogenised human – and the more we need to question what this data is based on. In current technological expressions, there is a great deal of similarity between AI's type of homogenisation and Critical Whiteness Theory, in which the power of AI is in its capacity to maintain the status quo and spread White supremacy.

Importantly, bias is understood and characterised differently in the tech world than for much of the rest of society. How a programmer defines bias is ultimately down to algorithmic programming; it is a form of technological bias. How the larger community defines bias relates to impacts on individuals and groups; it is what we know as social bias. The question is to what extent social bias can be 'programmed out' of AI, or even out of life; and if it cannot be removed, what are the systems of accountability and transparency we require to ensure its effects are not damaging?

Underlying technical solutions to *un*-bias AI remain challenging because of how this technology works. Whether, and to what extent, social 'data' (i.e. culture) is quantifiable, measurable and separable from the world around it matters in this context. In other words, can

society ever be reduced to a purely physical form of data, or is it always experiential?

While there are some attempts to create ‘sovereign AI’, in which individual countries or regions develop their own foundation models, there may be limitations in this due to the decontextualised nature of data that AI requires to function. Let’s work through some examples.

It is not yet known the extent to which a LLM may be able to capture, for instance, tense and the relationship of an action in Wiradyuri language. Waga-nha means dancing, such as ‘Jess is dancing now’. In English this would translate to ‘Jess is dancing’. However, when Jess’s dance is completed, in English the action would take the past tense: ‘Jess danced’. But in Wiradyuri the meaning is contextual and can be either waga-nhi, danced (Jess has completed the action of dancing) or waga-y-aan, have/has danced (Jess has danced, done the action of dancing before).<sup>3</sup> Each of these contextual phrasings conjure a different idea of Jess’s dancing, each distinct and meaningful. More so, without hearing the inflection in Jess’s voice or without being able to observe Jess’s body language when talking about the dance, a LLM may not capture other emotional aspects of the dance unless it is described in a written format the AI can access. In Wiradyuri ways of knowing, knowledge is inextricably tied to its relationship to the world around it; it is subjective and experiential knowing. If the LLM is ‘averaging’ data, the uniqueness of the individual experience, which comes from its context, is lost and the world may never fully understand just how badly Jess dances now or how badly she has danced before.

Let’s take a look at another case: Level 4 or Level 5 autonomous vehicles (AVs), in which humans are predominantly passengers, and the vehicle is responsible for driving and making all safety decisions. Some of the identified barriers to the safety of these AVs in ‘shared spaces’ is

the inability to replicate our cultural norm of responding to emergency vehicles or to consider how humans in non-AVs might respond to a road incident. In many cultures, emergency vehicles responding to life-threatening events have road priority – but AVs [often struggle to apply this concept to their driving](#), and similarly often fail to identify how humans driving non-AI cars will act if they see an emergency.

What if Jess fell over while dancing and she's at risk of serious health outcomes without rapid intervention? If something goes wrong and the AV makes a wrong decision (maybe it doesn't pull over) and Jess's ambulance does not get to hospital quickly enough, who is at fault? How do we avoid this? Can we ever hope to program AI to consider these human experiences and contextualisations when it only has access to physical knowledge (not tacit or subjective experiential knowledge)?

In many human cultures, this type of traffic incident would result in an investigation to determine responsibility for the crash. Would every 'wrong decision' where a crash occurs involving an AV simply be considered 'an accident', or would it be the result of 'faulty training', and what effect would this have on life and culture? For insurance purposes, who would be the responsible party – the programmer, the passenger, the ambulance driver? If the vehicle's programmer were found legally culpable, how would the responsible party be punished: a slap on the wrist, revoking the ability to program future AI? Would canny programmers develop workarounds to such sanctions?

Ultimately, would programming the AV to prioritise Jess's injury solve these issues or would it lead to other biases? As we saw in the research about LLMs and African American English, even attempts to 'fix' bias may not work.

While our society recognises the impacts of social bias, we continue to struggle to remove and adequately address such biases outside of AI. So what makes us think we can do this in developing and programming AI? Perhaps ‘reprogramming’ AI is our chance to get it right, if such an endeavour is possible. Or do we need an AI revolution, in which the current lineage of AI expressions is abandoned, and we begin from a different set of data – which includes the technological expressions of First Nations cultures, along with other (non-White) cultures, more actively in its frame? If we can have an open discussion of where social and cultural biases and systems of oppression exist outside AI, we are going to be more prepared to address their presence within AI – and revolutionise our approach to AI.

So, the question becomes not just to what extent can we remove bias from AI or automated systems, but how we understand the differences in social and technological biases and their different impacts on life. Can they coexist, or do we need to forge a new path for AI?

### Technological bias in AI

In the technology world, bias in an AI is considered present when AI ‘unexpectedly produces skewed results’. In other words, technological bias doesn’t necessarily mean social exclusion or perpetuating the dominance of certain social groups as it does in the broader community. It simply means that the machine didn’t work as expected. Technological bias, however, can result in secondary social exclusion.

In the case of AVs, a technological bias might be that the car didn’t stay on the route due to a kink in the programming, as opposed to a social bias, where the AV might take an alternative route because

it takes the passenger through a wealthier postcode and is therefore deemed ‘safer’.

These technological biases can be introduced through training (e.g. the AI is trained in a manner that isn’t appropriate for its intended task) or through the modelling (e.g. the way in which the AI processes information is incorrect). In the case of the AV, maybe it was trained on US driving, where drivers move to the right lane to allow an emergency vehicle through, but it was driving in Australia, where drivers more often move to the left. Or the modelling used for the AV’s information processing didn’t take into consideration the complexities of road rules.

This differing definition of bias matters. If an AI is designed to prioritise the health of the sickest patients, and it does that, it isn’t considered biased from a technological perspective because it is working as programmed. If, however, human review of the AI outputs suggests that determining the sickest patients ends up predominantly in White patients being treated over other racialised groups (because the data is based on symptoms and typical presentations in White patients), this is a social bias but not a technological one. In this case, it is a social bias reinforced through technology. Unless the AI was purposefully developed to account for these social groups when it was designed, it is clearly flawed, but in a technical sense it is working as intended. These differing and overlapping definitions of bias can mean that those developing AI and those implementing and integrating AI are talking at cross purposes, working to identify and solve different challenges. Understanding the different stages at which bias, in all its forms, can be introduced can help bridge this communication gap.

## Training data bias

Both technological and social biases can be introduced through the AI training stage and relate to *access to information*. The data humans choose to train the AI may itself be biased. There is increasing recognition that many AIs are trained on datasets sourced from the Global North and that predominately represent White people. A 2021 report revealed that many US banks' decision-making about loan approvals rely on biased datasets where the algorithm used [resulted in unfair outcomes, such as Black applicants and people of colour being 40 to 80 per cent more likely to be denied a loan than their White counterparts](#). In the United States, credit scores are often calculated according to a scoring system known as the FICO. A 2024 *Business Insider* article suggests that [many lenders still use the classic FICO score today](#), despite the existence of newer versions of FICO that are more equitable, and that [lending practices remain racially biased](#), resulting in life-altering outcomes.

The classic FICO relies on data going back more than 35 years, which results in a limited definition of credit, privileging the type of credit that White Americans are more likely to have (a history of loan repayments, which contribute to credit scores), but that are not more indicative of capacity to repay a loan (e.g. ability to pay bills and rent have no bearing on FICO-derived credit scores). Essentially, if you have borrowed money from an institution and have paid or are paying it back to the institution, you are more likely, through the classic FICO, to be considered as having 'good credit' compared to someone who has been paying money to a landlord rather than a bank, disadvantaging some while advantaging others.

When AI is used to make FICO-based loan decisions, it entrenches this bias. This reliance on a limited dataset to train and support AI

used by banks sets up a negative feedback loop where those who are regularly paying their bills may not have a strong credit history as measured by the classic FICO, and thus can't borrow money from most lending institutions, or can borrow less, because they haven't previously borrowed money. It's an AI Catch-22. What is arguably more disturbing, and further contributes to this technological bias, is that there are updated models that the lending industry could use, but they don't. While some lending agencies have moved to these updated models, specifically [government assistance programs developed for those with lower socioeconomic status](#), the majority of lending institutions remain committed to this othering of certain social groups.

Here we see a structural racism reproduced and reinforced through an AI technological expression – but with humans at the helm. Most of the industry is purposefully using flawed data to make decisions that benefit one set of lives over others, despite knowing the flaws, and using the algorithm to back up these biased decisions. The AI (in this case, the algorithm helping make decisions about lending) is actually technologically unbiased – it is working as intended. The social bias is introduced through humans who are opting to use a program which relies on data that reflects social bias.

The data that is selected for AI training determines AI outputs or purposeful behaviour. Classic examples of this are the social biases represented in facial recognition software, such as those uncovered in Joy Buolamwini's 2020 documentary *Coded Bias*.<sup>4</sup> These types of AI are often trained with individuals who are visually read as White. The result is that such facial recognition software often 'misreads' those not represented in the training datasets, such as people with darker skin tones. Sometimes the software entirely misses these people or can't tell them apart – Georgia State University researchers Thaddeus

L. Johnson and Natasha N. Johnson note how [AI-powered facial recognition will likely lead to increased police racial profiling](#) for this reason. Technically, because the facial recognition software is working as programmed, this system is technologically unbiased – even though it is reproducing and perpetuating social biases.

One could argue that there is a technological bias at play in this facial recognition software as the training should have been based upon the general population for which the AI would be used. It's a fair argument to make. The upshot, though, is that the developer can claim that there is no deliberate bias in the software even as the AI's outputs are discriminatory. Stephen Cave and Kanta Dihal in their 2020 article 'The Whiteness of AI' assert that 'race and technology are two of the most powerful and important categories for understanding the world ... [yet] their profound entanglement remains understudied';<sup>5</sup> in this case, the entanglement is heightened because of the nuances of language: social and technological bias are seen as two distinct and only tangentially related concepts. We see again, as we saw with the definitions of 'artificial' and 'intelligence', that words can be used to mislead and misdirect.

Humans need to re-evaluate this relationship between AI and the impacts it has on society. We all need to ask more from the companies and AI developers. We need to explore *for whom*, and for *what purpose*, the AI was developed – and we need to critically evaluate not just the intended purpose of the AI but the very real impacts it has on life. Before we integrate any AI into a space where human life and culture matter (i.e. everywhere), we need to question how the AI was trained and ask who wins and who loses because of this training – regardless of the type or origin of AI-perpetuated bias.

In higher education, eye-tracking software is sometimes used to identify cheating in online exams, on the basis that if you aren't keeping a steady on-screen gaze, you are likely cheating. Descriptions of this technology even go so far as [claiming to identify 'deviant behaviour'](#) through visual patterns.

Many of these algorithms are trained on the gazes of those who are described as neurotypical. The presumption is that non-cheaters keep their gaze steady. Neurodivergent people, however, can often look around, stare away from the screen or move a lot while working, resulting in a varied gaze<sup>6</sup>, or simply have different gazing patterns from those who are neurotypical.<sup>7</sup>

Similarly problematic understandings of differences in eye contact are present in the Australian criminal justice system, as outlined in the Australian Human Rights Commission submission '[Common Difficulties Facing Aboriginal Witnesses](#)'. The submission outlines how reduced eye contact as part of respectful cultural protocols of politeness is often misinterpreted as a sign of dishonesty. In these cases, the different patterns of gaze are socially and culturally classed as deviating from the norm and therefore suspect. If AI as it stands today were used in courtrooms, it is fair to assume that such biases would be embedded in the algorithms, resulting in false characterisations of dishonesty and implications of guilt.

We can see this bias play out in classrooms where AI detection software is [falsely accusing Black students at more than double the rate of White students](#) of using generative AI to complete assignments. This is despite numerous sources of evidence that such tools incorrectly categorise work by those from culturally and racially marginalised backgrounds as unoriginal, incorrect or plagiarised.

Essentially, the data used to train AI reasserts core tenets of ableism and colonial norms, placing those who don't fit these parameters at risk of being excluded and potentially accused of wrongdoing.<sup>8</sup> We tend to trust AI in our educational systems for purposes such as identifying cheating, maybe because we assume that it wouldn't be used if it were socially biased. The speed and reach of AI across our cultural systems should give us all pause.

To support AI in reducing social bias, better data training processes are essential. Humans often rely on both tangible and intangible data (e.g. both physical and experiential knowledge) to make decisions, such as if a student is cheating or not. The intangible cannot find its way into AI datasets – at least not yet. To return to the example of Vegemite introduced in Chapter 1, for instance, many Australians draw on a vast web of information – from personal experiences to social influences over time and through different points of information access (taste, vision, sensation, history, memory) – to consider Vegemite 'tasty'. Just like AI, Michelle only has the aggregated experience of these Australians to understand the tastiness. Unlike AI, Michelle can (and has) tried Vegemite, but because of her unique circumstances – differing culture and memories about food – she finds the taste repulsive.

The same is true in other contexts where AI is seen as critical. Take a doctor making a clinical decision about a person's health. The doctor's decisions are often based on information so tacit, so deeply embedded in their brains, that they aren't able to translate it into tangible data – [even when asked to think about their decision-making process](#). These 'gut feelings' and practice wisdoms (knowing gained through experience) often remain intangible – but are based on lived experience, which [is shown to help us make the right, or a good, decision](#). The doctor's prior experiences (accumulated through

different access points), training, cultural understanding and access to the information of that moment (the verbal and non-verbal cues, the tone of voice and the words said and unsaid) all contribute to this intangible web of knowledge. Even if AI had access to information from all of these human inputs, would it process this information like a human would?

There is some evidence that LLMs may be able to process information similarly to the human brain. A team revealed that the parts of the brain that are active during a conversation between two people were similar to the artificial neural coding activity of LLM outputs.<sup>9</sup> However, assuming this study involved the English language, what isn't clear is whether similar patterns would be found for other languages, such as Jess's Wiradyuri language, which conveys both physical and experiential knowing.

Because of these deficits in AI – its disembodied knowledge and limited access to information compared to humans – we argue that one way forward is enhanced data transparency. If it is clearer to AI users which types of data are omitted in the training process, or users are provided with enhanced contextual information on when the AI is not fit for purpose (e.g. indicating when the AI is not tested on people from a range of ethnicities, or only on neurotypical learners), then maybe we can embrace its value in guiding our decision-making processes while also recognising that it isn't an unerring fount of all wisdom, but rather a useful tool in the context it was designed to be used for.

New medications typically undergo extensive clinical trials before being released to the public. Consumers, once a medication has been approved by regulating bodies like the TGA or FDA, often trust in its safety and efficacy. The language tied to the clinical trials, such as 'test participants', 'double-blind study' and 'protocol measures', give most

of us a reassuring sense of authority in science, the doctor prescribing the medication, and the healthcare system as a whole, so that we will take the drug prescribed without question.

But imagine if a Big Pharma company updated their warning label on a medication to state that the test participants in the clinical trials for the drug only included White men from the United States. You, the patient, might pause because you are not represented in this participant group. Maybe you would think harder on the decision to take this medication if these parameters were clearer. If you were the doctor, would you take more time researching the medication before recommending it to your patients? We would hope so.

This is what we are advocating for with AI. We aren't saying 'don't take the medication'. Instead, we are advocating for processes of enhanced transparency, which can allow us to think critically as we use AI and integrate into our everyday lives.

AI developers need to, at the very least, be more transparent about how the AI is defined when it is working as expected. This can be achieved by being open about the training datasets, and explicit about the core objective when the AI was developed. Was it intended to evaluate credit scores, or is it intended to provide loan decisions with minimal social bias? If there isn't technological bias, there are likely still social biases in the algorithm, as in life, so users would benefit from knowing what 'working as expected' means for every AI expression in use.

What motivates Big Tech companies to act in this ethical manner, when not taking steps to address the bias is much easier and more lucrative? This is where many have called on governments and regulatory bodies to implement laws and policies around the use and development of AI. This is all well and good; but, as you see in this book, we are

also calling upon the community to question whether, and to what extent, we should invest in AI that doesn't put all life first, and we must also question the transparency of the regulators.

## Model test or training bias

In addition to the dataset that the AI is trained on, how the data is processed (*information processing*) also has the potential to introduce both technological and social biases. If the appropriate *access to information* is available, the AI produces an output based on how it processes that information (linked to its training) or the steps it takes (often unknown) to generate an output or purposeful behaviour. This is done through data modelling, and 'model bias' can be introduced through either 'overfitting' or 'underfitting'.

Overfitting occurs when the AI is accurate with predictions and outputs during the training phase, but does not do this as well with novel, never-before-seen data. Let's take the example of the AI that identifies the sickest patients in a health system. If the AI does well with the training dataset but fails to perform adequately with real-time data, this could be an example of overfitting. Essentially, the AI struggles to generalise the training model to other datasets – data used from healthcare system A works in training, but the AI cannot apply this training to healthcare system B. Maybe the training data is from a city hospital and the AI fails to work as effectively with data from a rural hospital. In our autonomous vehicle example, overfitting could be the cause of the vehicle's inability to apply its US training to another geographical location when deployed.

Such overfitting can be due to a variety of factors, such as a narrow training dataset (e.g. not enough patient information from those in rural

communities or not enough data on the roads), too much time spent in the training phase (so the AI becomes essentially too rigid), training data that has irrelevant data in it (e.g. the referring doctor's name in the electronic medical record or the AV passenger's religion), or such a complex model that the machine learning mistakes the irrelevant data as essential – which is very feasible in both the healthcare context and in relation to road safety.

In each of these cases, the overfitting results in technological bias – where the AI doesn't perform as expected. However, overfitting can also lead to social biases. Take the example of gaze tracking. If the training data is solely from neurotypical gaze patterns during exams, the model's accuracy for predicting the likelihood of cheating for those who are neurodivergent is significantly less.

Underfitting is when the modelling is too simple, resulting in the AI's inability to identify patterns, relationships and complexities within the data, rendering its outputs useless. This could occur if the AV's training datasets are too narrow, such as only including red lights or stop signs as an indicator to halt a vehicle during the AV training phase. The [AV will likely come to a sudden halt](#) (or worse, drive on) when the lights are down and a human is directing traffic and indicating with their hand for the AV to stop.

Human programmers are the ones selecting a model that determines how the AI processes information. In selecting a model, the developers need to be conscious of the goal and purpose of the AI expression. For instance, they need to consider how important accuracy is to the AI's function and whether the purpose of the AI is related to a simple question or a complex social challenge. Each of these considerations will help strike the balance between overfitting and underfitting so that the AI 'works as intended' and remains technologically unbiased.

Think about the scenario of a wildfire or bushfire. If the AI is being developed to predict the location or characteristics of the fire itself, accuracy (with a degree of latitude) may be valuable, even lifesaving. But if the AI is being used to determine a community's response to the fire by considering ways to adjust populations or practices in high-risk areas, flexible modelling that accommodates societal complexities is needed and will save more lives.

The point is that both technological and social biases can be introduced by the programmer selecting the model. Let's say a scientist whose speciality area of research is fire pattern spread is developing AI for adjusting population density in high fire-risk areas. This programmer may be more inclined to select a model that suits their view of the world, one that bases its decisions solely on fire danger risk – a specific perspective that may not reflect other perspectives, and therefore may be underfitting. What about a social scientist or a social worker? They may choose a different model that considers the complex social relationships between people and land impacted by the potential fires. This is where the idea that only 'AI experts' should be involved with AI development comes undone. Depending on the purposeful behaviour of the AI, the expertise, which can include lived experience, needed to build a robust AI system is often diverse and varied. Who is selecting the model, and their knowledge of the sociocultural context of the purposeful behaviour, can influence how the AI determines an 'answer' or output. Modelling influences both social and technological bias.

Gabe Barcelos, founding engineer of AI evaluation platform Arize, [uses the example of a wedding dress](#) to explain how model fit influences AI training:

... suppose we're trying to create a Wedding Dress Recognition System for a fashion store. The wedding dress dataset contains more than 5 million pictures of all the wedding fancy attire. Each garment features a wide range of pictures from various perspectives and lighting situations that would be found in a real-world setting. On such a large dataset, standard machine learning techniques such as decision trees, multiple logistic regression, and other tree-based models will not generalize well enough to accurately detect the wedding dress. As a result, the wedding attire is misclassified and the accuracy is low. The neural network, on the other hand, will generalize better if it is trained on a bigger data set. Artificial Neural Networks, for example, require a rather big data collection for optimal performance ... The trained model will now be more generalized and more accurate than the conventional one in terms of precision.

In this case, the developer could decide that an algorithm for wedding dress selection is a simple task that does not require complex decision-making, and (incorrectly) use more linear modelling for the AI. The result is that the AI unexpectedly fails due to a technological bias – it cannot accurately select the wedding dresses because of underfitting. The accuracy of the AI algorithm is paramount in this case, and more complex than originally programmed.

In an attempt at compensating for this modelling bias, a human perspective may be added. Humans could be engaged in the training process to tag the wedding dress pictures to help the AI define which images of dresses represent *wedding* dresses. But someone in one culture might not consider a sari a 'wedding dress', nor a black gown and so on – so the humans tasked with supporting the AI training

could introduce a social bias. If the AI is selecting only traditional Western conceptualisations of wedding dresses, and this is how it is trained, then the algorithm isn't technologically biased, per se. But the programming and training that went into the AI is introducing or reproducing a social bias, and potentially reinforcing systems of cultural oppression.

Despite AI being riddled with social bias, many humans seem to have their own bias in perceiving the AI as [infallible, objective and bias-free](#). In reality, current AI expressions rely on many steps of human interaction for training and modelling, and (hopefully) moderating outputs – AI is marbled with human values and biases. The greatest danger with AI isn't necessarily that it is socially biased, but that these biases are hidden, minimised or omitted from discussions, leading humans to feel overly confident in AIs' capacity for prediction. Many of us regularly question the logic of those around us, but we tend to be less critical (or even uncritical) of the logic of an AI. This needs to change.

Some have asked, why don't we simply make sure that the programming and training datasets have social bias removed? This techno-solution is unrealistic as it doesn't acknowledge the social complexities that lead to social bias, potentially replicating a mass 'colourblindness' of social injustices. In fact, such solutions have been tried, and the attempt to remove the social bias often leads to technological bias. Alas, if humans are involved, bias always exists. And AI – at this stage – still requires humans (no matter what the tech bros say).

Attempts to refine algorithms in ways that consider the complex social norms of human society and serve to reduce social bias are sometimes referred to as conducting 'AI neurosurgery'. At times

this means the machine learning is designed to avoid certain topics entirely. Maybe a large language model is designed to not use toxic language or perpetuate stereotypes;<sup>10</sup> but this can render it incapable of doing the original task because it produces ‘incomplete or misleading responses’. In other words, the attempt to remove the social bias results in a technological bias.

Another techno-solution often suggested is to engage humans to reduce the bias through a process termed *reinforcement learning from human feedback* (RLHF), where the machine learning is optimised through human feedback. A good example of this is Gemini, Google’s AI image generator, which was a debacle. Developers were attempting to ‘solve’ the bias of typical image generators, such as the tendency to portray [attractive people as young and light-skinned](#) and a propensity for perpetuating [ageism, sexism, racial bias and classism](#), but their programming tweaks resulted in representing ‘[Nazi-era German soldiers as people of color](#)’. The attempt to remove social bias in the AI through RLHF resulted in, on the most generous reading, complete absurdity and misinformation.

Maybe we need to consider the idea that we can’t remove social biases from AI. What if, instead, the best way forward is about recognising, teaching and highlighting these biases at every stage of life? This could help us counter the pervasive human belief that AI is infallible – highlighting that AI is actually quite fallible outside of the parameters it was programmed for. This could look like developers who transparently convey which datasets are used for training, including warnings that illustrate who the AI is for (and who it isn’t for), and human professional development highlighting how AI can be less useful in many complex scenarios.

Many of those advocating for ‘responsible AI’ suggest that adjusting current AI expressions can support enhanced outcomes. While interpretations of what constitutes ‘responsible AI’ vary, it generally involves an understanding that principled AI development and implementation should include **transparency** about when and how AI is being used and the ability to evaluate its outputs; **explainability** about how the AI is arriving at an output, and guidance on how to interpret this output; **fairness**, which supports diverse and representative datasets related to the topic, developer diversity, ethics review boards and bias mitigation techniques; and **accountability**, where people at all stages, from AI conception to implementation, should be identifiable and responsible for an AI’s outputs. The Australian Government has a [voluntary AI Ethics framework](#) that includes additional principles such as the idea that AI should ‘benefit individuals, society and the environment’, have ‘human-centred values’ that respect rights and autonomy, and allow people impacted negatively by the AI to challenge the outcomes through a ‘timely process’ (**contestability**). While all these goals are admirable, to what extent any of this is possible given the current trajectory of AI expressions remains unclear.

What AI appears to do [‘exceptionally well at is magnifying humans’ biases and faults](#). What if we used it for this purpose, to help us better identify and understand our already-existing social biases and shine a light on how these social biases perpetuate social injustice? In doing so, AI would give us a powerful pathway to address these social biases and proactively protect individuals, cultures and communities. We could restructure the way we work to use AI as a sort of ‘bomb-sniffing’ mechanism, where the bomb the AI is searching for is our social biases.

Instead of using AI to make recommendations on patient management or diagnosis (which are likely to be flawed and biased because of the

training and datasets), healthcare AI could be used to look at patterns in resource distribution (e.g. time with patients, prescribing patterns), flagging inequities to those in leadership. Healthcare AI is in an exceptional position to do just this. Many papers that note the racist outcomes of AI in healthcare resources distribution also highlight its power to ‘see’ the racialised distribution of healthcare resources – in a manner that our human intelligence is simply incapable of due to the vastness of the datasets.

On our model, instead of tasking the AI with making recommendations on how to distribute healthcare resources, the human takes over. The AI identifies the problematic pattern and the humans address the inequities. This might include actually listening to culturally and racially marginalised communities which already know about these experiences (and often how to solve them), and continuing to tell the world that care comes through community – not through top-down hierarchies and AI outputs. In this alternative engagement with AI, we use the technology for what it’s best at – seeing patterns – and we pair it with what humans are capable of being good at – critical thinking and community. The AI becomes the looking-glass reflecting our human biases and prompts us to see ourselves (and what we are doing to each other) more clearly.

### Countering human bias costs life

So, we could be more transparent about who is doing the AI programming, what data is used for the training, and how those using the AI are supported to be made aware of the frequency and type of social bias perpetuated by AI. Here comes the but. To engage AI in a way that considers bias, we need to ensure that the companies and

industries developing and supporting AI aren't themselves reinforcing bias and inequities. It seems to us that there is a limited possibility of this with the current AI lineage of technological transformations.

Just as the farming industry relies on day labourers, who are often overworked and underpaid, AI relies on a largely 'hidden' workforce often referred to as 'mechanical Turks'<sup>11</sup> to manage and train the machines – and to address the bias. Companies hire humans to support many elements of data management described in this chapter.

Mechanical Turks truly represent the 'humans in the machines'. They are involved in the ongoing training and moderation of the AI on a large scale. This little-known workforce is named for Amazon Mechanical Turk (MTurk), [described as](#)

a 'crowdsourcing' system, in which requesters post Human Intelligence Tasks (HITs) along with the fee they will pay for their completion. Turkers (the workers) choose their HITs, do the jobs and submit the results. Examples of HITs are locating information on a document, translating foreign languages, transcribing speech, as well as comparing audio to written transcripts.

Most of these HITs are used to train AI to better recognise information. For low pay, this group of contractors based across the world essentially upholds the facade of artificial 'intelligence' – whole teams contributing human cognitive effort that helps to disguise the true limitations of AI.

The concept of the mechanical Turk itself has an older and illuminating history. It comes from a 1770 chess-playing device created by Hungarian inventor Wolfgang von Kempelen. Players competed

against the Turk, believing it was an automated machine. For years Kempelen managed to sustain the illusion that it could ‘outcompete’ any human in chess. Eventually, the elaborate hoax was revealed – inside the machine was a human master chess player moving the mechanical arms.

Companies go to great lengths to hide the human work (and working conditions) that are required to make AI, ‘artificially intelligent’ technology. Modern-day mechanical Turks are involved in all sorts of tasks related to AI training and development; tech companies would have us believe that AI is fully automated, but in reality they have the ‘expert chess player’ making the moves. As Associate Professor Elizabeth Stephens of the University of Queensland writes in the 2023 paper “The Mechanical Turk: A Short History of “Artificial Artificial Intelligence”, ‘the open secret’ of mechanical Turks’ ‘artificial artificial intelligence is itself a form of misdirection that hides other, more successfully guarded secrets: the true extent of that labour, and the conditions in which it is performed’.<sup>12</sup>

These modern mechanical Turks are often underpaid and their work undervalued, while the companies who hire them earn big money. In 2022, a company called AI Insights [was accused of stealing work and not paying their workers](#):

... when AI Insights posted a request for more than 70,000 HITs during what is typically a slow season on the platform, it represented a bonanza of opportunity for ‘Turkers’ ... But as they got to work, in some cases completing hundreds of HITs, the Turkers soon realized that AI Insights was rejecting all of their work en masse, without explanation. According to the platform’s guidelines, that meant the Turkers wouldn’t

be paid, but that AI Insights would get to keep their work all the same. Also, since individual Turkers' approval ratings are affected anytime their work is rejected—and since most requesters on the site won't accept bids from Turkers with less than a 99% approval rating—the mass rejection also sent many Turkers' ratings tumbling downward, effectively blacklisting them through no fault of their own. When Turkers contacted Amazon, asking them to intervene, the tech giant washed its hands of the situation, saying they can't 'get involved in disputes between workers and requesters'.

The lack of employment protections for these 'flexible' workers meant that recourse was nearly impossible.

Many US companies outsource this type of work to those in the Global South who earn, often, less than \$1.50 USD per hour. A 2023 *TIME* investigation found that OpenAI, the company who created ChatGPT, [used Kenyan workers paid a take-home wage of between US\\$1.32 and \\$2 an hour](#), in a country that lacks a universal minimum wage. In several countries, Amazon is sometimes not even paying its flexible workers with money, [but instead with gift cards](#).

It isn't just about the pay; the nature of the work is often challenging. Data labelling involves identifying raw data such as images, text files and videos and adding informative labels to provide context so that an AI can learn from it. Mechanical Turks and others doing this type of work must frequently navigate graphic and traumatic content. The *TIME* investigation reported that a data labeller was tasked with annotating 'graphic depictions of suicide' without any access or support. The poorly paid Kenyan workers hired through subcontractors to support OpenAI had to review and tag hours upon hours of content

that was violent, hateful, depicted sexual abuse and so on, to help train ChatGPT to identify toxic and inappropriate content.

Workers report, when asked, the long-standing repercussions this work has on their lives. *TIME* shares the experience of a data labeller who ‘suffered from recurring visions after reading a graphic description of a man having sex with a dog in the presence of a young child’; the employee stated that it was ‘torture’. None of these workers had access to HR support, psychologists or content warnings. The subcontractor, Sama (which markets itself as an ‘ethical AI’ company, apparently without irony), ended up cancelling its contract with OpenAI early due to the deeply traumatic nature of the work – but what about the traumatised workers, who now also have to find other work? As *TIME* states, ‘for all its glamor, AI often relies on hidden human labor in the Global South that can often be damaging and exploitative. These invisible workers remain on the margins even as their work contributes to billion-dollar industries.’

There is even some evidence that data labelling is exploiting languages and further entrenching bias. Data labellers in Africa are [relying on their local dialects and languages to help with AI training](#), resulting in adoption of these languages in the AI outputs. The result? Commonly used words like ‘delve’ are now overused by AI – and are increasingly considered ‘bot-speak’. Communication in Africa is being appropriated, and the effects are wide-reaching: ‘If AI-ese sounds like African English, then African English sounds like AI-ese. Calling people a “bot” is already a schoolyard insult (ask your kids; it’s a Fortnite thing); how much worse will it get when a significant chunk of humanity sounds like the AI systems they were paid to train?’

Mechanical Turks aren’t the only labourers at risk from AI. Companies are laying off workers to improve ‘efficiencies’ by paying

for AI programs instead – but at what true cost? How did we build our economies, our human workforce, in a manner that reinforces and supports AI at the expense of life? When did our society move from human-centred to technology-centred – at the expense of life?

AI not only has the potential to introduce and reinforce social biases, but it is built from inequities, on labour that further perpetuates structures and systems that disadvantage one group over another, prioritising the lives of some over the lives of many. Thus, when we consider AI, it isn't just the programming that perpetuates bias – it is the very essence and makeup of AI, and humans, that reinforce it. Debates about AI should never leave the ethics of its development and implementation out of the frame, and more informed discussions are one element of an AI revolution.

## Chapter 4

# The Ethics of AI

### As complex as life itself

When the automobile was invented and replaced the horse and buggy, human systems and structures were disrupted. John Hope Bryant, CEO of American financial literacy non-profit Operation Hope, suggests that a similar event is occurring now with current AI expressions and that the ethical implications – the job losses and the changes in the way we do work – also carry with them ‘tremendous opportunities’.<sup>1</sup>

We challenge this idea through the lens of wayanha. We aim to show that AI is not actually exceptional, nor is it disrupting life in a ‘new’ way. Quite the opposite. Contemporary AI expressions are simply an updated model by which unethical human behaviours are perpetrated at greater speed and with further reach. It is this pace and influence, as compared to prior technological expressions, that is the true disruption of AI.

The opportunities that Bryant feels are afforded by AI only apply for some, at the expense of others, as it has been with all prior Western technological expressions. Don’t buy into the hype: contemporary AI expressions are no different in the types of ethical impacts they cause than every other technological expression in the lineage – and in this chapter we will explore why. For a truly disruptive technological transformation, one that revolutionises our approach to the future, we need to transform the AI evolution we are on to one that centres *all* life.

In this book we are seeking to explore AI's impacts, and illustrate that (depending on the AI expression) AI can both harm and benefit life now and in the future. We need to consider AI's *behavioural space* outside of just those driving its creation and extend this exploration to an examination of our moral and political behaviours.<sup>2</sup> Contemporary AI expressions are already shaping the way we think [about the world around us](#). We ask you to consider that there is an unequal impact of AI on the living when we consider AI's full behavioural space here on earth, and in the colony of so-called Australia because as we have discussed and continue to discuss, Australia has a racist history that centres only certain (White) lives. AI has an unequal and unjust impact when we consider its full behavioural space on our planet, and in the colony of so-called Australia, because as we know, Australia has a racist history that prioritises certain (White) lives.

As mentioned in Chapter 3, the Australian Federal Department of Industry, Science and Resources [has formulated eight AI Ethics Principles](#) in a voluntary framework. The framework is intended 'to be aspirational and complement – not substitute – existing AI regulations and practices'. The principles include:

**Human, societal and environmental wellbeing:** AI systems should benefit individuals, society and the environment.

**Human-centred values:** AI systems should respect human rights, diversity and the autonomy of individuals.

**Fairness:** AI systems should be inclusive and accessible, and should not involve or result in unfair discrimination against individuals, communities or groups.

**Privacy protection and security:** AI systems should respect and uphold privacy rights and data protection, and ensure the security of data.

**Reliability and safety:** AI systems should reliably operate in accordance with their intended purpose.

**Transparency and explainability:** There should be transparency and responsible disclosure so people can understand when they are being significantly impacted by AI, and can find out when an AI system is engaging with them.

**Contestability:** When an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to challenge the use or outcomes of the AI system.

**Accountability:** People responsible for the different phases of the AI system lifecycle should be identifiable and accountable for the outcomes of the AI systems, and human oversight of AI systems should be enabled.

The ethics of AI calls for critical questioning to understand the principles, values and morals in the AI (r)evolution. There are contemporary and historical experiences across the globe we can learn from as we explore the impacts of technology, including AI, on life, and which lives it benefits the most. In these examples we can see how AI can contribute to, reproduce and reemphasise classism, racism, queerphobia, sexism, ableism and violence, but also explore

evidence that AI can counter these. This is a theme across the AI landscape – AI can be a tool for either good or evil, or both.

## Technology: reinforcing systems of oppression?

There is extensive and widely reported evidence that current expressions of AI can reinforce White supremacy and colonialism, given both the violent history and present reality of settler colonial states. Colonial societies have used, and continue to use, expressions of technology for invasion, genocide and subjugation.

Systems of oppression are structural and institutionalised forces within society that systematically harm and disadvantage certain cultures, groups, individuals and ways of knowing, while affording privilege to others. Systems of oppression are embedded in society through laws, policies, discourse and social norms. Such practices are referred to as ‘cultural hegemony’, where those in power work to maintain this power and influence the population to ‘agree’ to these norms.<sup>3</sup>

Australian historian Patrick Wolfe reminds us of an important reality regarding colonialism and invasion – they are structures, not events.<sup>4</sup> Colonialism is an ongoing force that continues over time through various methods. It builds in systems that produce particular behaviours, norms and social practices. Colonialism’s systems may commence through an initial invasion followed by laws, policies and genocidal practices<sup>5</sup> such as forced child removal and destroying land and sacred sites. All of these violent actions work together over time to erase First Nations and other cultures to assert colonial rule and cultural norms. This way of thinking can be applied to other systems of oppression such as queerphobia, ableism and sexism. We see AI, especially in its current expression, as inseparable from systems of

oppression, and thus it represents a facet of what is termed ‘digital colonialism’.

### AI as digital colonialism

Digital colonialism is used to describe the phenomenon whereby powerful and technologically advanced settler colonial states and countries ([like the United States](#)) use algorithms, data and digital technologies to exert power over others. Digital colonialism can be understood as an extension of colonialism, as power and exploitation are used for political, economic, social and cultural gain via digital spaces and technologies. Furthermore, Kerry McNerney argues that [AI rhetoric reproduces and re-embeds old nationalist, colonialist, imperialist and racist structures](#) into society:

AI development [is] a zero-sum game where the victor will not only control the most advanced AI technology, but also enjoy economic, political and military dominance over all other nations [where] AI nationalism cannot be understood without careful attention to how racism and imperialism underpin the AI arms race. First, the AI arms race is not merely the pursuit of technological expertise, geopolitical dominance or military power over another nation. Instead, it is a fundamental contest over racial and civilisational superiority, one deeply rooted in previous histories of colonial violence and racial capitalism.

Digital colonialism poses a range of life-threatening concerns in the current phase of AI transformations, especially in settler colonial states. As discussed earlier, AI systems and algorithms rely on large

datasets for training and learning. Given the [global dominance of White supremacy](#) and colonialism, these ideologies and systems are widely embedded within the databases used by those doing the programming, whether intentionally or not. In effect, AI systems perpetuate biases present in the data, risking the production of biased outcomes and endangering life as we know it.

We have seen examples of digital colonialism in the US, where AI is replicating racial bias towards Black patients, reducing ‘the number of Black patients identified for extra care by more than half’. In a study published in *Science* in 2019, Black patients were ‘considerably sicker than White patients’ but received less care, even when the health needs and case factors were identical.<sup>6</sup> Another example is related to heart disease, where the data collected during routine exams was itself biased. As [Katherine Igoe wrote in 2021](#):

The Framingham Heart Study cardiovascular risk score performed very well for Caucasian but not African American patients, which means that care could be unequally distributed and inaccurate. In the field of genomics and genetics, it’s estimated that Caucasians make up about 80 percent of collected data, and thus studies may be more applicable for that group than for other, underrepresented groups.

Such biased data collection applied in human healthcare practice is then embedded into the AI systems, which have greater reach and are arguably more ‘trusted’ than the human doctor. Additionally, this shows that lifesaving fields such as medicine are not divorced from White supremacy and are a part of re-perpetuating racism.

Ironically, these studies show how healthcare systems claim ‘doing no harm’ as a core value, yet the algorithms and technology they rely on are replicating patterns of racialisation, which is causing real harms to Black patients. Circling back to the framework that Palanca-Castán and colleagues produced, we see how access to information about Black and White patients (and in this case, racist training datasets) leads to AI programs processing this racist information, resulting in a racist behavioural space. Digital colonialism circulates through AI healthcare algorithms as they not only further marginalise and harm Black patients, but also ensure safer and better healthcare for White patients, privileging one group over another. The category a life falls into stems from White supremacist and colonial structures: some lives are centred, and some are marginalised.

Digital colonialism also comes in the form of White-washing imagery. In 2023, Midjourney, a generative AI program, was used to generate images of [‘beautiful’ South and East Asian women](#). These synthetic images drew on significantly biased datasets based on stereotypes of Asian women as well as White Western beauty standards – *ABC News* noted that the women in the images all shared ‘fair skin, thin noses, full lips and high cheekbones’. Similarly, an Asian student used [Playground, a free online AI image generator](#), to turn a photo of herself into a professional headshot, which made her White and blue-eyed.

Discrimination in AI involves textual forms, too. Some AI text-detector programs are [discriminating against non-native English speakers](#), wrongly identifying their written work as AI-generated. [AI detection programs have also falsely accused international](#) and non-native English students of cheating in university assessments. Ironically, international students are [frequently encouraged to use programs like](#)

[Grammarly](#) to support their English writing skills in an effort to meet learning requirements. The university sector has seen institutions and teachers using AI detection software to point the finger of blame at the student for using AI writing assistants such as Grammarly, instead of first examining the quality and inclusivity of learning environments.

Digital colonialism around generative AI extends to primary and high school classrooms as well. In 2023, Lucinda McKnight and Cara Shipp highlighted alarming attitudes towards generative AI in education, with teachers considering it as ‘just a tool’. They unpacked how reducing generative AI to a tool reflects colonial ideologies of education and ownership over knowledge:

Students are entitled, in every interaction their schools and teachers make with generative AI, to know where their data goes and how it is used. This is of particular importance for First Nations students, as is the principle of colonised peoples retaining data sovereignty and control of Indigenous Knowledge (IK).<sup>7</sup>

Similarly, Macquarie University academic Dr Tamika Worrell observes how for First Nations communities, teachers using generative AI ‘risk perpetrating and promoting inaccuracies and spreading false information instead of meaningfully engaging with Indigenous values and knowledge systems.’<sup>8</sup>

These reductive, punitive and short-sighted approaches are far from a model for inclusive education. They are colonial approaches via technology that lacks true intelligence, acting to reinforce White supremacy within society and its institutions.

## Human decision-making reinforces digital colonialism

[Digital colonialism](#) is perpetuated not only in how AI processes information, but also in how humans integrate AI into the behavioural space. Currently, humans are directing the AI to facilitate the structural embedding of digital colonialism: there are multiple examples where vulnerable people and groups are being excluded through algorithmic design and flawed data. This highlights a social justice issue: how technological expressions can exacerbate and further entrench colonial structures when humans use the digital as a tool to colonise. In this way, humans are the key to digital colonialism – and how to minimise it.

In mid-2016, the Australian Government adopted the use of an automated program across the country to address outstanding debts and potential cases of fraud in the welfare system. In the hopes of making the debt collection process more ‘efficient’ and ‘objective’, the Robodebt scheme became an ominous lesson in letting technology run amok. Online Compliance Intervention, an automated data-matching technique, was said to accurately identify welfare recipients who had potentially been overpaid. It could issue computer-generated debt notices to these recipients, at a rate of up to 20,000 a week. But the technology was flawed: it issued incorrect or non-existent debts to hundreds of thousands of welfare recipients, and placed onerous burdens of proof on individuals to prove their innocence, causing severe stress and mental hardship to many who were already subjected to systemic oppression. Individuals and families on government payments were unfairly targeted for, according to the algorithm, ‘incorrectly’ declaring their income, and the media painted a picture of welfare fraudsters drenched in classism and [racialisation](#). The devastating outcomes of Robodebt, including loss of human life, led to a royal commission,

which presented damning evidence of inappropriate *human* behaviour in relation to the technology, indicating that it was widely known that many of the debt recovery notices were inaccurate, and labelled the Robodebt policy an [‘ethically indefensible policy targeting vulnerable people’](#). From this we can learn about the potentially life-threatening consequences of AI: while Online Compliance Intervention wasn’t an AI system, but rather a technology that relied on algorithms, it is an expression of technology. [The Robodebt scheme is a cautionary tale as we contemplate an increasingly automated future](#), especially in the context of substantial developments in AI.

The real-world outcomes of what is now known as the Robodebt scandal illustrated that an algorithm that did not place human life at the fore ended up being in fact [life-taking](#). The algorithm became a tool to reinforce the status quo and further harm vulnerable people.

The colonial idolatry of [objectivity](#) may be driving many to use AI, while simultaneously relinquishing responsibility for causing damage to human lives when things go wrong. The British Post Office scandal saw thousands of postmasters, and countless other lives, impacted by algorithmic expressions of technology. Between 1999 and 2015, more than 900 postmasters were convicted of theft and fraud over errors in an accounting and stocktaking software system, Horizon, that had been introduced to post offices nationally. People were imprisoned and bankrupted, and many died by suicide in response to inaccurate accounting due to a flawed algorithm. Sub-postmasters were left defending themselves against a software program, with those in charge trusting the technology and apparently losing any sense of reason, instead choosing to believe that a raft of postmasters had suddenly turned into criminals. Astonishingly, [nearly 100 million pounds was alleged to have been spent](#) defending

the technology, despite seemingly incontrovertible evidence that the machine responsible for the accounting and stocktaking was dangerously flawed. Bugs that were reported repeatedly by postmasters were ignored by senior officials at Royal Mail. With both Robodebt and the British Post Office scandal, human lives were collateral as those responsible for human welfare deferred to, defended and then finally blamed the machine.

As researcher [Alison Hearn](#) shows, automated systems such as the Robodebt Online Compliance Intervention and the British Post Office's Horizon 'put the onus on citizens to prove their innocence' while in both cases the 'processes of appeal were complicated, lengthy and, yes, automated'.<sup>9</sup> When we turn over decision-making entirely to these technological expressions, the cost can be both human dignity and human life. The knowledge and knowing of such technological expressions and systems should be grounded in care for life, not in the subjugation or ending of certain lives.

This theme of 'just blame the AI' is repeated in medicine, arguably an exemplar of life-affirming systems. A [2024 policy brief about AI in healthcare from Stanford University](#) opens with this powerful statement: 'Optimism about AI's tremendous potential to transform healthcare is tempered by concerns about legal liability: Who will be held responsible when the use of AI tools contributes to patient injury?'

The report goes on to highlight that the 51 cases of AI-related medical injury stemmed from two causes: 1) software defects or malfunctioning devices and 2) physicians deferring to technology for care decisions. While humans are involved with the development of the AI and in making the decision on whether and to what extent the AI recommendation should be followed, the report found that holding developers and physicians liable for these errors was nearly

impossible. With this in mind, AI not only becomes an ‘assistant’ in making decisions, but it can also absorb the responsibility of any ill-begotten outcomes from such decisions. Handy.

AI is also being used directly to kill humans – divisions of the Israeli and US military use it to identify targets, including ‘[tens of thousands of human targets](#)’ in Gaza, according to a report. When these armies are accused of using AI in contradiction of humanitarian international law, the public response is either to call into question whether the system is *really* AI (using the ill-defined nature of AI as a justification) or to reinforce that ‘human oversight’ is applied before any action is taken, to illustrate that AI is ‘never’ acting alone. In reality, these AI expressions are being used to weaponise and perpetuate colonial norms and oppression of certain bodies – at the expense of life.

As we have seen with Robodebt and the British Post Office scandals, there is also a realisation that human oversight may not be a very effective mechanism for ‘ethically’ deploying AI. The reality is that humans have a great tendency for automation bias, where we trust the machine too much or seek to avoid taking responsibility and use it as the ‘fall guy’. We tend to not question its output as thoroughly as we might that of a human making the same claims – and this seems to be particularly true in high-stakes contexts with high levels of uncertainty, where human life is at stake, such as in [war](#) and in medicine.<sup>10</sup> While the psychology behind this tendency to over-rely on technology in the such circumstances is still a matter of debate, automation bias may – at least in part – stem from a [desire to remove the discomfort](#) we experience with such contexts.

Each of these examples illustrate that technological expressions can have extensive impacts on life – either through humans’ active choice in deploying AI expressions violently or by handing over both the task

assigned to the AI as well as the task of critically thinking through the process of automation bias. These instances expose how easily AI becomes a scapegoat, and how those already vulnerable can be left to defend themselves against sometimes even clearly fallible technology, or can find themselves at the mercy of such technology when other humans target the AI towards them. Such very real occurrences illustrate a clear message: technological expressions have the power to take lives if we let them.

If we are not careful, ‘The AI made me do it’ could be the defining phrase of the twenty-first century. And many argue that current legal systems may not have the complexity to address these AI-related ethical and legal issues. In a [2024 Conversation article, Associate Professor Michael Duffy](#) warns that existing legal liability frameworks are insufficient for the current AI expressions:

This is because apart from some product liability laws, current theories often require fault through an intention, or at least provable negligence by an individual. A claim for negligence, for example, will require that the harm was reasonably foreseeable and actually caused by the conduct of the designer, manufacturer, seller or whoever else might be defendant in a particular case. But as AI systems continue to advance and become more intelligent, they will almost certainly do things with outcomes that may not have been completely expected or anticipated by their manufacturers, designers, and so on.

This relates to what Elina Nerantzi and Giovanni Sartor refer to as ‘AI crimes’ and ‘hard AI crimes’.<sup>11</sup> They define crimes as ‘the intentional performance, by an AI agent, of actions which would constitute a

crime if they were performed by humans (having the appropriate *mens rea*), whereas hard AI crimes ‘refer to those AI crimes for which no human can be considered criminally responsible, according to the criteria currently used for ascribing criminal responsibility’. AI does not (yet?) have legal personhood. How will legal systems address the ‘machine defendant’, as Michael Duffy terms it, and what impacts will this have on how humans are prosecuted under criminal and civil laws? If AI technologies continue to emerge and commit hard AI crimes, how will victims seek justice for violence, loss and trauma if we do not consider criminal sanctions for the humans creating and endorsing the AI?

The social discourse in these examples focuses on the AI programs themselves as causing the issue. However, the human decision to defer to the technological expressions (or lead with them) over the human capacity for critical thinking may also be to blame. It is humans who build the technological expressions with racism and colonial coding that leads to these flawed algorithms. These technological expressions don’t only sit in relationship to prior technological transformations; they sit in relationship to the humans who make the datasets and the humans who use them, and the society that prioritises certain types of data over others. We cannot continue to blame the AI alone when it is the human behaviour, including the maintenance of settler colonial narratives and perspectives, that perpetuate injustice and violence through AI and other technological transformations.

## Eliminating and re-writing history for White settler futures

Colonialism is known by its forced imposition of a ‘savager’ narrative on marginalised or ‘colonised’<sup>12</sup> peoples and cultures. This narrative discursively and structurally positions ‘colonised’ cultures and groups as dehumanised, underdeveloped, inferior and in need of saving by the superior colony or race (most frequently White people). More so, especially in settler colonies, White settler futurities are centred. This idea ‘refers to a future where settler power and bodies are sustained and preserved, while Indigenous [or other non-White] bodies are erased’.<sup>13</sup> Do current AI expressions offer up another tool to ensure future White lives, but not others?

It isn’t just the AI itself that perpetuates these unequal systems of power; we have already seen that how the AI is implemented and supported, and who has access to the technology, are also mechanisms through which power is maintained by the elite, especially as the use of certain technology has been equated with civilisation. For colonisers, the denial of the technological advancements and expressions of other cultures serves to reinforce perceptions that these cultures are primitive and uncivilised. For instance, through false claims of ‘terra nullius’, settler colonisers in Australia continue to claim that First Nations communities lacked technological and societal infrastructure in agricultural fields, and deny that production existed prior to invasion.<sup>14</sup> This perpetuates a notion of First Nations populations as unenlightened, in turn re-enforcing White ‘civilised’ infrastructures as necessary to reinforce the hegemony. This claim has been proven untrue many times, including by Bruce Pascoe, a Yuin, Bunurong and Tasmanian man. His bestselling book *Dark Emu* draws on coloniser diaries and

other sources to outline how sophisticated First Nations societies were at the time of colonisation and how evidence of this was destroyed or omitted from historical records to maintain the ‘savager’ narrative.<sup>15</sup>

With this in mind, could AI be yet another mechanism to ‘edit’ history, through the selection of datasets and homogenisation of colonial narrative outputs? Could AI datasets be the new coloniser diaries, writing history for the benefit of those in power and White settler futurities? Given the lack of integration of First Nations ways of knowing, as many of these knowledges are not readily accessible or actively sought out to be integrated into AI (despite AI’s capacity to perpetuate appropriation of First Nations content), this futurity may already be a reality. However, if First Nations cultures were made more accessible to AI, will these knowledges just continue to be colonised?

While this colonial narrative was and is incorrect, why does an absence of certain technologies or structures make a society, culture or group inferior and in need of ‘development’? In effect, a dependence on technological advancements has become a sign of ‘civility’ ingrained within colonial discourse. We see how digital colonialism is perpetuated by today’s ‘tech bros’<sup>16</sup> (defined by the [Cambridge Dictionary](#) as ‘someone, usually a man, who works in the digital technology industry, especially in the United States, and is sometimes thought to not have good social skills and to be too confident about their own ability’ – we will let this definition speak for itself) who tell us that AI is exceptional, all-knowing and necessary for humanity to flourish. But this is not the only way to understand technology. Technological expressions have taken many forms through history as transformations occur. We need to question why certain expressions are afforded more value than others, and why colonial structures and systems of oppression keep asking us to seek the next technological transformation while

ignoring First Nations technological expressions and continuing the ‘savager’ narrative in the process.

And so, another dilemma arises in dominant Western settler colonial nations, governments and institutions having significant economic and resource access to specific AI development in continuing to assert savager discourse through programming and simply through ‘othering’ Nations and groups for not having (or *seemingly* not having) the same resources and technology.

War chemicals – for example, white phosphorus munitions – are supposedly highly regulated through international law and conventions, yet their use continues. Asking ChatGPT in February 2024 about the use of these violent technologies was met with *very* one-sided and Whitened examples of who ‘did’ use these chemicals and who ‘allegedly’ used them, or who used these war chemicals in self-defence and who used them as acts of terrorism and violence, showing racialised narratives. For instance, the answer provided mentions the Iraq/Iran war of 1988 and the Syrian civil war. While such chemicals are mentioned as being used in World War I, who used them is mysteriously omitted. Here we see how the use of a technology *and* the storying of its use (or denied use) continues colonialism via the digital world. Some may see a comparison of AI and chemical weapons as a stretched analysis, but it is a stark and powerful reminder of how non-life-centred humans can be with technologically advanced weapons, something we know too well from the use of uninhabited aerial vehicle (drone) strikes to remotely kill humans.<sup>17</sup>

The pervasive connotation between machines and Whiteness continues to gain traction in how we visualise and imagine AI. Stephen Cave and Kanta Dihal argue that imagining ‘machines that are intelligent, professional or powerful’ is to imagine ‘White machines

because the White racial frame ascribes these attributes predominantly to White people'. They note how AI has become racialised as White, erasing 'people of colour from the White utopian imaginary'.<sup>18</sup> We see here that in the very language of AI and its associations, White futurity is protected.

The many harms of digital colonialism extend to intellectual property rights and data governance across the world. It is important to understand how certain groups' and cultures' ways of knowing are influencing (or influenced by) AI expressions. The Global Indigenous Data Alliance, an international network dedicated to promoting Indigenous Data Sovereignty and governance, advocates for 'asserting Indigenous Peoples['] rights and interests in data', for 'data for the self-determined wellbeing of Indigenous Peoples' and for 'reinforcing the rights to engage in decision-making in accordance with Indigenous values and collective interests'.<sup>19</sup> It involves members from First Nations communities in seven locations, including Australia, Aotearoa (New Zealand) and Spain.

Ian Tapu and Terina Fa'agau also consider the ongoing concern of data governance in First Nations communities, especially where data about First Nations peoples is shared, created and owned by non-Indigenous peoples and institutions in colonial structures. Herein, they make similar arguments to the ones we express in this book, questioning whether AI 'will serve as a "revolution" or a "new colonizer" for Indigenous peoples – an answer that ultimately hangs on which narratives AI developers embed into their technologies'.<sup>20</sup>

Taking a more global perspective, [Grace Browne](#) examines the ways in which technological expressions, like AI, are sold as a way of 'helping' particular individuals and communities, yet it is 'often imposed on them without consultation, pushing them further into the

margin'. She quotes Ethiopian cognitive scientist Dr Abeba Birhane: 'Nobody in Silicon Valley stays up worrying about the unbanked Black women in a rural part of Timbuktu.'

This raises ethical questions around who makes the AI, who is responsible for the AI, and who owns the material that AI is fed or puts out. Is the data being shared or stolen? This discourse is at the forefront of the contemporary debate about AI as the colonisers' identity is being challenged, ironically by their own use of AI. The very tenet of the colonial way of 'knowing' – that an individual owns knowledge – is coming into question. The AI is 'stealing' knowledge from others, just as colonisers have stolen from societies and sovereign First Nations cultures across time. So the question remains: is AI a coloniser or a decoloniser?

## Consent, privacy and our data

Current expressions of AI have far-reaching physical harms through metaphysical and digital means. Consent, privacy and user engagement, and the relationship between them, are omnipresent realities of a digital world<sup>21</sup>. These issues relate to determining which data is used to train the AI and whether users are aware when they are providing data that it may be used by developers to refine their products. What further complicates this relationship is legal jurisdiction, especially when there is no international law governing data consent. There are regulations that cover multiple jurisdictions, including the General Data Protection Regulation (GDPR), sometimes referred to as the 'gold standard' for data protection legislation, which applies to the European Union and the European Economic Area.

The GDPR has been described as ‘the toughest privacy and security law in the world’. Its [website states](#): ‘Though it was drafted and passed by the European Union (EU), it imposes obligations onto organizations anywhere, so long as they target or collect data related to people in the EU.’ The GDPR [lists consent as one of six aspects of lawful processing in relation to data](#), defining it as ‘any freely given, specific, informed and unambiguous indication of the data subject’s wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her’. The GDPR outlines that consent can be **revoked at any time** and individuals should **be able to do so easily** and **indefinitely**. In practice, even within the European Union this is often far from reality.

When it comes to their own data, users are more likely to be given an option to opt *out*, as opposed to opt *in*. Vinayshekhar Bannihatti Kumar and colleagues, in the aptly named article ‘Finding a Choice in a Haystack: Automatic Extraction of Opt-Out Statements from Privacy Policy Text’, state that opt-outs ‘allow a user to exclude themselves from data practices such as tracking by advertising networks, sharing of personal information with third parties, or being contacted by phone or e-mail’.<sup>22</sup> However, they warn that opt-out instructions are often buried within lengthy website privacy policies, highlighting the issue of *informed* consent, and who is accountable in ensuring accessible information around the consent we are giving. Again, the theme of those in power hiding behind the technology, and abdicating responsibility to it if there are adverse consequences, appears. What if you didn’t understand elements of the consent requests on a website? Who to contact and how to address these queries is often not clear. In such cases, informed consent is near impossible. Varied and diverse models of consent are beginning to emerge and being used by

companies and technology providers to meet data collection objectives and requirements. But there is often a level of duplicity involved, as is illustrated by bundled consent.

Bundled consent refers to ‘the practice of “bundling” together multiple requests for an individual’s consent to a wide range of collections, uses and disclosures of personal information, without giving the individual the opportunity to choose which collections, uses and disclosures they agree to and which they do not’.<sup>23</sup> In short, bundled consent is asking users to agree to all or nothing. Consent becomes a question of yes or no, instead of a series of decisions a user can make about their data and how it will be used.

Everyday examples of bundled consent can be found in social media applications and streaming services, such as Netflix. A new user must agree to certain conditions, relating to topics such as data sharing with third parties, in order to create a subscription account or profile. If you don’t tick ‘yes’ you are excluded from the platform – and in our world, digital connections and content have become key elements of participating in our society. Users may be happy to consent to some uses of their data but not all, and only for a certain period of time, but too often they do not get the choice.

Similarly, it is common if signing up for an everyday debit card or a credit card to be asked to agree to some bundled consent conditions related to the data the bank can collect, and how it is used. Without a bank card, you are limited in your options to complete transactions. As we move to a cashless society, such banking bundled consents are essentially requiring you to tick ‘approve all’ or limit your participation in mainstream society. But that bundled consent [may include](#) terms around sharing data of your spending habits with the bank’s business clients to provide insights into how they can market their products.

Is bundled consent actually informed consent if you can't select which information is collected or if you can't clarify points of confusion in the terms and conditions? It could be argued that users are still practising informed consent to a degree in that if they withdraw their consent to the bundled consent option, they will no longer be allowed to use the service or device and therefore their data will not be collected. However, this puts users in a difficult predicament where the use of the technology is needed to access certain services and supports. Such forms of bundled consent represent a contemporary [Hobson's choice](#), where the choices we are given aren't choices at all; either we say 'yes' and participate in society's gains or 'no' and are excluded, giving an illusion of absolute freedom to opt in or opt out.

Government bodies have sought to mitigate these data consent practices, including the NSW Civil and Administrative Tribunal, which opines that in the context of health, [bundled consent approaches are insufficient](#) in meeting the Health Privacy Principles in the New South Wales *Health Records and Information Privacy Act 2002*. They also make clear the difference between a privacy notice and consent:

A privacy notice is a one-way communication; it does not ask for a response from the individual. It simply states: 'this is what is going to happen with your personal information'. Notifying a person of what you intend to do with their information is not the same as seeking their consent to do those things. It is important not to confuse a privacy notice with consent.<sup>24</sup>

Despite this, many websites, including online AI platforms, use privacy notices to claim they are seeking informed consent from users, resulting in a conflation between privacy and consent. An example of

this can be seen with OpenAI – their privacy notice is more a disclaimer than a consent process. Among the more than 3000 words, users may note the company’s claim to implement ‘commercially reasonable technical, administrative, and organizational measures to protect Personal Information both online and offline from loss, misuse, and unauthorized access, disclosure, alteration, or destruction’. Without further characterising how this is done (as would be expected with any human ethics application for research), it reminds users that ‘no Internet or email transmission is ever fully secure or error free’ as a get-out-of-jail-free card.

The complexity of ethics, consent and privacy varies according to context. For example, ethics requirements for human research involve significant oversight and transparency (or at the very least, that is what the research ethics policy requires). In human research, we are encouraged to de-identify data and report how we will store and protect the data. At times, researchers or organisations may withhold data from public reports since it cannot be entirely de-identified. Here the relationship between privacy and consent re-emerges, as someone may consent to their de-identified data being shared, but the sharing of this data may not be de-identifiable because it is possible to identify the person even when steps are taken to anonymise data. For example, a First Nations person from a small regional town may consent to their health data being collected at the local community health centre. The consent form that the patient signs states their data will *not* be identified. However, given the small First Nations population in this regional area, this person’s data may still be identifiable even though their name is not attached to it. So, the health body may choose to not release the data or present the data alongside data from a wider regional area, to address the risk of identification. This creates significant justice

issues for communities and populations whose identity is disregarded from many reports but in turn may then not be included in future policies that are developed based on publicly available health data, impacting their health. In this way, certain groups can then be placed in an impossible position of trading their data, and their identity, to access healthcare. Barbara Prainsack and Nikolaus Forgó warn that paying people for their healthcare data risks widening inequalities, ‘luring people to sell their privacy’.<sup>25</sup> Here, access to healthcare is being held hostage, and the ransom is data. Even without this payment, accessing a fundamental part of the Australian healthcare system – the general practitioner – requires disclosure and sharing of personal and private information.

While privacy and consent are different, we see how they are interrelated when data becomes the new currency that is traded for goods and services. What appears to be happening is that for-profit AI companies are ignoring or only loosely considering this important relationship between consent, privacy and life to increase their access to data. Think of someone who clicked a bundled consent where they signed up to access a product or an application, only to realise later that they no longer want to allow that consent. In a research study, the participant could withdraw, and their data would be removed from the study. But here, in the for-profit digital environment, data already provided remains, at least in some contexts, the property of its new holder indefinitely.

Some suggest that consent in the digital space should be an ongoing conversation between the giver and the receiver.<sup>26</sup> Consent given once to an AI transformation is insufficient when considering the relational ways of knowing that many communities work from. Such data practices of one-off, all-or-nothing, bundled consent deny choice

as a dynamic part of the human experience, recording our decision at a point in time and holding us to it forever.

There is also the issue of repurposed data. How much do most of us know about how our data is being used? How many of us know if we have consented to all the ways in which it is used? Adam Andreotta and colleagues, in their article ‘AI, Big Data and the Future of Consent’, are illuminating on this point, and it is worth quoting them at length:

While a human would find it near-impossible to search through tens of thousands of medical records, to discover novel patterns and insights, an AI algorithm can be designed to perform such a task very quickly. As beneficial as this can be in many contexts (e.g., assisting patient care or preventing disease (Arnold [2021](#)), informed consent may need to be secured again, if the original consent is no longer applicable. For example, someone who consents to sharing their postal code may wish to withdraw consent when they learn such data can be used to determine insurance premiums (see Floridi 2019, p. 110). And this is also true in biomedical contexts, where medical information or tissue samples are often stored and then requested for further research to which the original Participant Information Sheet and Consent Form (PICF) did not refer ... The general problem of re-purposed data, then, is that data users ... have not always limited their use of personal data to the purpose for which the subjects’ original consent was applicable. This is morally problematic because it ignores the preferences, and potentially the wellbeing, of these subjects (ACCC 2019).<sup>27</sup>

As data has become the currency of our time, the sharing of data through AI expressions is propelled to new extremes. What impacts will this have on consent? Whether you are signing a half-page consent form at the GP, or a research participant's five-page consent agreement, or creating a player profile on Candy Crush on your phone that includes your name and email, or signing up to a generative AI platform that records your search terms, search patterns and creative outputs to create a profile of you as a user, the importance of consent and privacy remains, especially when human life is involved.

Reflecting on the AI transformational lineages encourages us to view obtaining consent not as a simple tick-box exercise but as a complex process. Current AI expressions position efficiency and the right to access information as primary considerations, aligning with colonial values and approaches to knowing, where data (our identity) is owned by another. If we continue on our current trajectory of embracing AI with incredible speed and a remarkable lack of critical reflection, is the risk of a lack of informed consent, and the harmful outcomes this can have in some situations, an emerging problem?

While diverse approaches to consent may improve the choices for users, the purpose of some types of data collection, and the potential uses to which the data can be put, continues too often to be unclear. AI requires partnership approaches 'built on responsive design and continual consent', as American scholar Meg Leta Jones and colleagues note.<sup>28</sup> Ethics needs to be centred now and in the future in all AI development and design. Joanne J. Bryson, Professor of Ethics and Technology at the University of Bath, reminds us to remember that accountability does not apply to the machines themselves, but to the people and organisations who develop, own and operate the machines.<sup>29</sup>

Regardless of jurisdiction, consent data practices are riddled with ethical dilemmas, and current AI expressions further complicate this, giving rise to significant new moral and legal concerns. Are we beyond actioning AI regulation that calls for an opt-in, informed, express consent (giving consent clearly in writing or verbally) approach to data sharing? How do we centre rights in these opt-out, not opt-in, conditions? How do we avoid creating a reality where the terms and conditions just grow and grow, meaning most users won't read the fine print, thereby reinforcing automatic opt-in culture? Adam Andreotta and colleagues argue that 'the introduction of ever more detailed terms and conditions forms for users to read, or more "policy acceptance" boxes for users to tick, *prima facie* may allow companies to secure greater levels of express consent, but it will make the question of whether that express consent amounts to *informed consent* only more complicated, not less.'<sup>30</sup>

So what future does this pose in a digital world, where our data is currency to survive and participate in society?

## Consent, safety and our bodies

Another dimension of physical harm perpetrated through AI relates to sexual violence and rape culture. Consent and online safety apply to more than just our information; they apply to our bodies and how they are, or are not, protected in the current AI technological expression. [Legal scholar Anastasia Powell](#) argues that in society 'too often when people talk about culture and sexual violence, they think of problems "out there" in the world', when they are in fact close to home. AI can exacerbate the perception that sexual violence isn't part of our daily reality – but it very much is.

Sexual violence is a crime even when carried out via AI. In a landmark case in the United Kingdom in 2024, a judge banned a [sex offender convicted of making more than 1000 indecent images of children from ‘using any “AI creating tools” for the next five years’](#). At Bunbury Regional Prison in Western Australia, a training company used [an AI chatbot to create a fictional sexual harassment scenario that unknowingly used the name of a former employee](#). Ironically, this occurred during an investigation into allegations of sexual harassment and bullying from employees. The reality of AI and sexual violence is seen globally, where [child sexual abuse and exploitation materials via AI-made images](#) are rising. It is important to note in these cases, particularly regarding children, assault and abuse, that this is not simply about an absence of consent. It is about the sexual violence perpetrated onto life, by life, through technology and the latest transformations of AI.

AI complicates consent and, as Emmie Hine of Oxford University’s Internet Institute notes, ‘poses a tremendous threat to individual autonomy’, calling for a model of ‘informed digital consent’ where users’ autonomy is not simply respected, but enhanced. In discussing American universities’ definitions of sexual consent, Hine states that ‘the key components of consent are the requirements that the action is only taken if the consent is affirmative; that all parties be adequately informed; that consent is a continuous, evolving dialogue between parties; and that parties are able to withdraw consent at any time’.<sup>31</sup> We argue that this definition should extend to contemporary expressions of AI, wherein AI expressions (and all outputs) require the same standard of consent, particularly where sexually explicit material is involved.

If existing design principles of AI do not [sufficiently recognise and implement informed consent](#) measures, and humans do not adequately

address this deficit, then our choices and even lives are at risk. In situations of a breach of consent, whom users can report this to or seek protection from is vague and unclear. Again, responsibility becomes opaque in the current AI transformation, and is further compounded by existing systems of oppression and victim-blaming.

The idea of reporting is also complicated by the societal and structural systems of oppression victims face in reporting sexual assault. If you are violated by AI, will you be believed? More so, who protects the privacy and consent of the voices the system of rape culture seeks to silence or violate, including Black, First Nations and Peoples of Colour, children, sex workers, kink practitioners,<sup>32</sup> women, trans, non-binary and disabled folks? When the violence of the world upon physical bodies goes continuously unchecked and unaddressed, can we really believe that digital violence will be held to a better level of account? The lessons learned in the Robodebt and British Post Office scandals suggest that it won't be.

The AI evolution is here, and here to stay. Many accept as fact that the world will continue to change, morph and adapt through AI technology. However, has humanity consented to this? Do we have the option to opt out? Who is ensuring our informed consent, to protect our rights and lives? Or is AI the ultimate 'Exhibit A' example of bundled consent, where humans have a choice to opt in to all or nothing, and if they opt in, they do so on behalf of all life?

Are the options of consent with AI buried so deeply within systems of oppression that they are inaccessible? Are consent forms now rendered unnecessary, because they have lost any utility? How is this ethical?

Ethical governance of AI is a requisite for a life-centred future. However, whose ethics guide these systems of governance and decide

which lives are harmed and which are protected? Thomas M. Powers and Jean-Gabriel Ganascia warn of potentially overestimating future technological and ethical issues that may never arise.<sup>33</sup> But can we really be too cautious with life, especially when we consider the critique from many, that we have been here before? Human life, Country and technological transformations are all connected – history tells us that when such technological expressions as AI are driven by colonialism, this relationship causes harm to many for the benefit of the few.

Consider the discussions we have presented so far in this book regarding how earlier forms of technology were used (and still are) to violently oppress certain people, groups and communities. It is clear that human rights must be systematically and consistently considered at all stages of technological and AI development, design and use to ensure our future prioritises life for more than just those coming from the dominant colonial perspective.<sup>34</sup> Alas, navigating the ethics of AI and data practices is an ethical issue in and of itself.

To co-opt the language of Gen Z, this ever-watching, violating, no-opt-out reality ‘is giving’ the Eye of Sauron from *Lord of the Rings*, or more seriously, a panoptical digital connoisseur objectifying life with its gaze. Let us explain. Many readers may be familiar with the concept of the male gaze, which Monash University’s [Janice Loreck](#) describes as ‘the sexual politics of the gaze and suggests a sexualised way of looking that empowers men and objectifies women’. In the male gaze, she notes, ‘woman is visually positioned as an “object” of heterosexual male desire. Her feelings, thoughts and her own sexual drives are less important than her being “framed” by male desire.’ French theorist Michel Foucault also famously described the panopticon and surveillance of prisoners, where the prisoner is seen as an object of information, but they do not see who is watching

them, in effect policing themselves through the panoptical gaze of the police. In 1990, US philosopher Sandra Lee Bartky argued for what could be seen as an amalgamation of these two gazes through ‘a panoptical male connoisseur [who] resides within the consciousness of most women ... Woman lives her body as seen by another, by an anonymous patriarchal other’.<sup>35</sup> Our question is whether, and to what extent, AI is objectifying the human experience, and to what extent those in power are harnessing this ‘gaze’.

Jess, in her PhD thesis, has also applied this panoptical analysis through a sovereign First Nations lens in discussing ‘how a panoptical White connoisseur operates upon the bodies of Indigenous, Black and people of colour and seeks for us to live in our bodies as seen by the White sovereign’.<sup>36</sup> Building upon the work of these individuals, we see another gaze, the AI gaze or the digital panopticon, where humans are being watched by AI expressions in a way that does not even require our physical selves, just our digital fingerprint or the memory of ourselves imprinted in the digital. These data breaches are happening in government and the private sector at alarming rates. For instance, in Indonesia in 2022, [more than 21,000 companies experienced breaches](#). In Australia, the major cyber-security incidents of the 2022 [Medibank hack](#) and the 2023 Optus data breach [caused nationwide consequences](#).

Our words, our images, and the parts of ourselves we share in the digital space become an extension of our lives, which are in turn being policed, and objectified, by the AI gaze. This digital panopticon is always watching. Feeling unnerved yet?

Is AI, and those programming its use and allowing its consent violations, just the newest cop in the tower watching over us with its digital desire, registering us (and our Country) as merely objects of

data that can be used to its advantage? And will we smile politely out of fear, rushing home with our keys in our hands, running and looking desperately for safety, hoping that we live by hiding our personal data or withdrawing our consent from the all-seeing ‘AI gaze’?

## Consent, exploitation and Country

There are many intersections that arise in this discussion of consent and AI. So far, we have discussed how human life, safety, privacy and consent are complicated and entwined in technological transformations, but what about the consent of other forms of life? Has Country or land consented to AI’s impact and exploitation? We could agree Country has not consented to the violence humans have inflicted on it so far, and this also applies in the context of current AI expressions. Prior technological transformations have contributed to climate change; what makes us think the current AI expressions won’t further this contribution?

While many are using AI to fight against [environmental damage](#), through surveying the land (including [iceberg sizes](#) and [deforestation](#)) and gathering data from [vulnerable communities](#), and by supporting [recycling](#) and [cleaning initiatives](#) or [shoring up wildlife populations](#), this (and all other) AI technological expressions have an astronomical impact on Country.

The recognition of this means that, as we write this book, there is a ripple that is fast becoming a wave of mainstream media articles highlighting the devastating and far-reaching impacts of AI technological expressions on Country. For instance, Paris Marx’s revealing July 2024 article, ‘[Generative AI is a Climate Disaster](#)’, states that Google admitted that ‘its emissions have increased by 48%

in just five years', which poses a challenge to their earlier commitment to net-zero emissions by 2030.

This 'oops, I did it again' approach to pillaging Country is an all-too-familiar confession in the tech industry. From Microsoft to Google to OpenAI, all of these companies are aware of their massive negative impact on Country and readily acknowledge this to the public, rationalising this damage by suggesting that such grievous acts are necessary in the name of 'progress'. A [report](#) from Ireland suggests that data centres, which are supporting AI expressions, 'consumed more electricity last year than all of its urban homes combined', and that 'Ireland's data centres [will] consume about 31% of Ireland's electricity within the next three years'. Accompanying all of these 'admissions' about fossil fuels is a predictable statement alluding to the need to increase renewable energy – but so far the increases in fossil fuel usage have far outpaced steps towards renewables.

In some cases, the resources sustaining life are redirected towards the technological expressions at the expense of the living. In Taiwan, a ravishing drought is costing lives and livelihoods. The water needed to quench the nation's thirst is instead being directed towards the semiconductor factories. We have built economies that support non-living entities and benefit certain lives over others (those at the top of the privilege tower) and AI technological expressions at the cost of Country, at the cost of life. Maybe the current capitalist economy is in dire need of its own wayanha, or transformation. After all, the economy is a socially determined construct that we can redefine and redirect towards life-affirming systems and structures, instead of life-taking ones. Perhaps AI can help write this restructure of the economy and rid of us our desperate reliance upon White colonial systems – but given the data it uses, probably not.

What we can see from earlier discussions is another debate brewing across society about the behavioural space of AI. While programs are built to accomplish a task, often human-centred, the impact of AI tends to far exceed the original target. When we consider a life-centred future, it isn't just human life we are referring to. It includes our planet and the ecosystems within it.

The question we challenge you to ask is whether, and to what extent, there is scope to change the behavioural space (e.g. impact) of AI on life, to ensure it centres all life, including Country and land, not just certain bodies or species. On the current AI evolutionary trajectory, we need to ask: whose life, human or otherwise, is seen as expendable to support this technology and the White settler futurity it may ensure?

While First Nations cultures and communities tend to focus on a relationship with Country when considering knowledge and intelligence (e.g. knowing), and in doing so working *with* the land to sustain human life, ecosystems and future generations, not all humans do this. Many in dominant settler colonial societies focus solely on sustaining human life, and arguably a subset of these lives. While much of the discourse about AI focuses on human life, there is increasing acknowledgement of the impact on Place, Country and land – our home, our planet.

In a recent [op-ed](#) in *Nature*, Kate Crawford – a professor at the University of Southern California and principal researcher at Microsoft Research – sheds some light on this arguably more hidden behavioural space of AI. She notes the exorbitant environmental costs linked to developing and running AI systems, stating ‘one assessment suggests that ChatGPT ... is already consuming the energy of 33,000 homes’, repeating the colonial pillaging of Country we see in Ireland.

There are multiple steps where the environment is impacted by AI expressions. Training a single large language model (LLM) was

calculated to be equivalent to ‘125 round-trip flights [between New York and Beijing](#)’ in terms of carbon dioxide emissions. While some can rationalise this planetary cost as a ‘one off’,<sup>37</sup> there are environmental costs related to the ongoing usage of contemporary AI transformations and AI’s future expressions. We need to ask ourselves, every time we engage with AI: is it worth it?

Every time we ask ChatGPT to ‘condense my writing’ or ‘explain it to me like I am a 3-year-old’, there is an environmental cost – and if you multiply this (there were 1.63 billion visits in [February 2024](#) alone, to a single LLM), the environmental cost for life is beyond significant. For context, this means that about 20 per cent of the global population sought out support from an LLM in a single month, and these actions cost the environment probably without users even knowing it. These estimates are for a text-based LLM, which has a lesser impact on planetary health than AI image generators. Images are more complex than text, making them larger files to create, process and store, requiring more energy. Think of all those X (formerly Twitter) threads where people tested the AI image generator into abstraction by asking ‘draw me a cat but make it more loveable’ over and over – and we quickly move from entertainment to our planetary demise. Such life-taking impacts on Country are repeated again and again, with [Microsoft Corporation’s AI resulting in a 30 per cent increase in carbon emissions](#) in just a few years.

As X user Librarianshipwreck (@libshipwreck) writes on [17 May 2024](#): ‘Yes, AI has a limitless thirst for water & energy, but cooking the planet even faster is a tradeoff we must be willing to make in order to get horrible AI-generated images, mountains of misinformation, deepfake pornography, & absolutely terrible writing.’

Our drive for this technological transformation is killing life; it is life-taking.

And this is just the beginning – ChatGPT is just in its infancy (if you can consider version 4.0 ‘infancy’). The CEO of OpenAI (which developed ChatGPT), Sam Altman, [acknowledges](#) that the AI industry is approaching an energy crisis.

Other aspects of technology’s impacts on the environment include electronic waste disposal, which is projected to be 120 million metric tonnes by 2050, according to the [World Economic Forum](#). Most technology devices have a built in ‘use by’ date; a Kindle (e-reader) stops working and must be disposed of, whereas a book can be read for decades – indeed, rare books are a valuable commodity for some. When the Kindle and other e-waste (e.g. other technological expressions such as laptops) pass their usefulness, its disposal contributes to the unsustainable trajectory we are currently on. [The World Counts website](#) tracks the impact of this disposal – with the numbers ticking up every second. Technological transformations aren’t just harming the planet when they are functioning; they are impacting Country when their ‘life’ has ended too.

While AI can serve to help address certain aspects of [climate change](#), AI technologies need energy to run, and water to cool the systems while running – and the reporting of the reality of this environmental burden is often missing in our discussion about AI. Yet again, the proprietary nature of AI technologies lead to opacity, and the impact on humans (certain humans) is prioritised. Those in charge continue towards the next AI transformation without consideration or consent from Country, potentially eroding the life of Country, and by extension human life, in the process.

It should concern all living beings that ‘[the climate is changing so fast that we haven’t seen how bad extreme weather could get](#)’. Prior data about weather patterns is unlikely to be helpful in modelling contemporary and future weather because we seem to have crossed a tipping point where the consistency is inconsistency. This is more than frightening. Yet because of AI’s extensive behavioural space, it may be furthering misinformation about the climate, leading to widespread misapprehension about the nature of the climate emergency. An [Australian-led study published in August 2024](#) suggests that even those with an understanding of the science of climate change appear to be influenced by false claims when they are restated over and over, due to the ‘illusory truth effect’ – where incorrect information, when repeated regularly, becomes perceived as fact. The technological transformation of AI, and its unprecedented efficiency and reach, is a tool for spreading such lies. By engaging with AI technological expressions, we are supporting this climate disaster and accelerating this continued assault on Country, which began in the name of ‘progress’.

As the development of the majority of AI in use is concentrated in the hands of a small number of companies, the end goal isn’t to sustain life, but to build wealth, where land continues to be a commodity to exploit for financial and White gain.<sup>38</sup> Because the majority of AI systems are [products of for-profit development](#), we too often do not see impact reports beyond the financial. As science journalist Jude Coleman notes in *Scientific American*: ‘there’s a lack of transparency from many AI companies ... That makes it even more complicated to understand their models’ impact’.<sup>39</sup>

The article goes on to explain that AI technologies could further entrench the fossil fuel industry, improving efficiencies in mining, crude oil industry or deforestation and so further damaging Country.

The environmental impact of AI thus extends to fossil fuel and habitat loss, as well as the energy and water use, and carbon emissions, of running the AI expression itself. In addition, a recent article in [earth.org](https://earth.org) highlights that while automated vehicles or electric cars may reduce carbon emissions compared to traditional petrol and diesel cars, they are not environmentally sound. The mining to support [electric car lithium and cobalt battery development is a violent disruption to Country, with ongoing significant environmental effects](#). This exploitation of Country extends to [exploitation of human life](#), with [child labour](#) and [financial enslavement](#) in places like the Democratic Republic of Congo involved in the mission to provide Western countries with their electric cars and smartphones.

You may now be seeing how entangled the relationship is between human life, Country, settler colonialism, racism and AI. At every stage, from its conceptualisation to its training to its integration, AI expressions have the potential to enhance life (at least some life), but they also have the potential to exploit and end life. Can we even determine a cost–benefit ratio to life in relation to AI? It isn't clear to us, the authors, whether the environmental trade-offs are worth it.

Our desire to rush towards the next big thing is resulting in life-threatening consequences, in both the short and the long term. Where is the tipping point in the balancing between life and progress – and whose progress are we supporting? Could we regress human life by progressing AI? The continued lack of transparency and accountability around the AI industry brings us to what we call the **AI–life paradox**: where AI is sold to us as supporting life, despite incontrovertible evidence that contemporary AI expressions are life-taking. Every AI system we integrate into healthcare to save a life potentially costs us

lives in its development and training. At what point is there no more life left to exploit?

If we reconsider the reasons and purposes for AI, can we develop AI with sustainability of life at its centre? By revolutionising AI's transformational journey to one that encompasses the technological expressions that First Nations peoples have developed over generations (that brought forth the boomerang, cultural fire burning, renewable energy and sustainable fishing practices, among many other developments), can we build a different, life-sustaining AI expression and future? While we can continue to use AI to help us address our environmental catastrophe, we need to ask ourselves at the same time why we are so drawn to a techno-fix, so reliant on technological transformations such as AI to solve the problems that were partly generated by our pursuit of technology to begin with – especially when our use of AI is further damaging the Earth. First Nations knowledges and ways of being through care of Country show that it is possible to nurture and protect our relationship with life, alongside technology. What might a more inclusive, less damaging, more life-affirming AI look like? We can also learn from the resistance to colonial violence First Nations, Black and people and communities of colour have shown time and again. What if we, collectively, resist life-taking AI?

## Chapter 5

# A Life-Centred Future

## Sustaining Country in an AI world

As we allow AI to be embedded into the very foundations of society, we need to ask ourselves: whose life, body, knowing and story is dominating our current reality – and so without revolutionary transformation will dominate the future – and whose is forgotten or omitted?

[American artist John Knoll said](#), ‘Any tool can be used for good or bad. It’s really the ethics of the [person] using it.’ While Knoll was referring to design software programs such as Adobe Photoshop, the quote could easily describe AI technologies, as well as colonial technological transformations from the past and those that will appear in the future.

If we use AI without understanding the ethical dimensions, injustices and limitations of this tool and of humanity, we are likely to perpetuate the ‘bad’ and to affect certain lives, including Country, dramatically for the worse. AI isn’t the only tool or technological transformation that we have used ‘badly’, of course: our creation and use of a series of technologies has been a key part of disruptions and violence to and across the planet. We need to remember the humans behind the digital curtain. The people wielding the tools have the power. If they are ill-informed about AI and how it works, or seeking to capitalise on its use, AI is more likely to become a tool that widens injustice – even if unintentionally.

This book attempts to encourage you to recognise that we need to engage AI with our eyes wide open, questioning the purpose, the training and the people involved with the development of AI expressions, as well as the impact of these on *all* life.

In thinking of AI as just another transformation of technology, and relating it to Jess's reflection on wayanha and the smartphone in the introduction, we can ask, 'What is AI's purpose in life?' The smartphone provides us with a means of communication and connection; what does AI add to or subtract from life? Do we have influence over AI's technological transformations – can we shift its purpose or are we at a point where the transformation is instead shifting us towards a future that is not necessarily the one that many, or even Country, want?

By this point you will understand (if you didn't already) that if you aren't questioning AI in this manner, you are helping to perpetuate the values of settler colonialism and to compromise a life-centred future. Humanity is at a point in the AI conversation in which all of us need to be aware of the actual and potential impact of AI on all life forms, and on future generations. The good news is, we are ready for a life-centred AI revolution.

## AI and the alignment problem

Many people think alignment is the solution to the problems with AI described throughout this book. The idea of alignment involves developing AI expressions that represent human values and goals.<sup>1</sup> This is a worthy goal, but it is not uncomplicated. By this point in the book, we can all foresee the questions: whose values should AI be aligned with, and what lives are valued or not valued in this alignment? The

critique in this book highlights that ‘alignment’ for some lives is a misalignment for others.

An article by scholar Iason Gabriel published in 2020, two years before OpenAI released ChatGPT, nailed the challenges facing those attempting to achieve alignment between AI and human values and goals:

Foremost among these [challenges] is the question of what—or whose—values AI systems ought to align with ... Behind each vision for ethically-aligned AI sits a deeper question. How are we to decide which principles or objectives to encode in AI—and who has the right to make these decisions—given that we live in a pluralistic world that is full of competing conceptions of value?<sup>2</sup>

The concept of alignment assumes that humans can support the development of AI to become aligned with human morals and values, or we can program it to do this. But we must first identify whose ethics, morals and values we are meant to align to.

Let’s consider such formative questions about AI alignment through the decades-old thought experiment of the trolley problem, whose conception is attributed to British moral philosopher Philippa Foot in 1967 (though its origins go back further). Foot writes, ‘it may ... be supposed that [a man] is the driver of a runaway tram [trolley] which he can only steer from one narrow track on to another; five men are working on one track and one man on the other; anyone on the track he enters is bound to be killed.’ The trolley problem asks you to consider which track the tram should take – the one that kills five or the one that kills one.

The trolley problem can illustrate the challenges with defining universal morality – a necessary precursor for AI alignment. The trolley problem can be adjusted to represent other living beings such as babies, puppies, and so on. When altered to represent other forms of life, we can see something deeper. The trolley problem is an analogy to everything we are exploring and critiquing in this book, namely what life is deemed able to be sacrificed and how the notion of the common good prevails for White life.

In the original trolley problem, it was a human controlling a machine, and human life or lives that must be sacrificed in this moral dilemma. Let's consider the trolley problem as it relates to AI alignment. In trying to figure out whose values and norms AI should be aligned to, a utilitarian logic points us to 'normative' values – the values that are thought to benefit the most or are agreed on by the community. In other words, the values and morals aligned with the 'common good' are the ones we should program AI to align with. But the 'common good' is often coded language for settler colonialism. 'The majority' represents the homogenised White settler colonial culture, not non-White or marginalised cultures. In essence, AI alignment tells us that it is okay to kill one (the marginalised other) to save the five (the settler colonial majority). Our book is filled with examples that speak to the truth of this.

When we use terms like 'normative', it is a linguistic proxy for privilege, hegemony, the dominant, elite social groups and often Whiteness. The question is: are we willing to risk the 'one', the peoples that society has culturally and racially marginalised, for the 'many'? What about if we consider that, in fact, this group represents the [global majority – that Black, Indigenous and people of colour represent over 80 per cent of the world's population?](#) We need to consider these

voices, and the voices of neurodivergent peoples, disabled peoples and those close to Country because basically, anyone who isn't White, able-bodied and neurotypical is currently on the track waiting to be hit by the AI trolley to preserve the 'greater good'. Those developing this expression of AI, as evidenced time and again in this book, are willing to run these bodies over. If you aren't disconcerted yet, you aren't paying attention, or your values reek of privilege and power.<sup>3</sup>

Let's extend this moral dilemma, as American philosopher Judith Jarvis Thomson did in 1976. Thomson asks us to consider the role of human agency in the decision about who dies and who survives. Would we knowingly push someone to save the people on the track? Would it depend on who is being pushed or who is on the track? Thomson writes:

George is on a footbridge over the trolley tracks. He knows trolleys, and can see that the one approaching the bridge is out of control. On the track back of the bridge there are five people; the banks are so steep that they will not be able to get off the track in time. George knows that the only way to stop an out-of-control trolley is to drop a very heavy weight into its path. But the only available, sufficiently heavy weight is a fat man, also watching the trolley from the footbridge. George can shove the fat man onto the track in the path of the trolley, killing the fat man; or he can refrain from doing this, letting the five die.<sup>4</sup>

The 'fat man' that Thomson adds to the trolley problem adds the dimension of knowing sacrifice. In relation to our discussion, it could be seen to represent the many humans and lives that are needed to make AI run. It could be those without alternative means of work

(for instance, mechanical Turks), and those whose positions are lost for the sake of efficiency as AI is integrated into the workplace. The ‘fat man’ also represents place and Country – our planet – whose energies and resources are depleted to keep the AI train moving. The AI is built in a manner that maintains the White settler colonial status quo, sacrificing the culturally, racially and socially marginalised to keep the train on track.

Thomson’s ‘fat man’ twist to the trolley problem is not unproblematic: the glaring fatphobia of this analogy and its devaluing of a certain body because of a socially constructed fear of fatness;<sup>5</sup> the bizarre proposition that that body is the only way to save those other five lives. The point (which is also illustrated in Thomson’s twist, in fact) is that the positioning of different bodies as justifiable collateral for the sake of ‘progress’ is a settler colonial lie that entitles powerful bodies to sacrifice certain ‘abnormal’ bodies for the sake of ‘the common good’.

There are diverse value systems beyond the White colonial norm that we can learn from to transform the current AI evolution into a life-centring revolution. For example, First Nations knowledge systems from across the world include, as has been widely noted, ‘[values of honour, trust, honesty, and humility; they reflect commitment to the collective and embody a respectful relationship with the land](#)’. African Ubuntu ethics systems reflect values of ‘[reciprocity ... peaceful relations, emphasis on human dignity and the value of human life, as well as consensus, tolerance, and mutual respect](#)’. Could such values be programmed into current AI expressions or built into regulating AI now and into the future?

Given all the evidence of social injustice and violence in our society, our collective limitations in understanding human behaviour, and the challenges we have in defining human morals and values, is AI

alignment really the goal, especially if the value and ethics systems we adopt are those of the White settler colonial West? And if it is, whose ethics, values and morals should we align to? Should we only consider human life in this alignment, or is alignment with ecosystem sustainability a consideration?

We think the better way forward is to not keep forcing the alignment of AI with colonial structures that are life-taking. Instead, let's seek to disrupt this evolution with an AI revolution that centres life.

## Humans and the AI problem

Another idea that is often posed to address the numerous challenges with AI is to keep 'humans in the loop' when integrating AI into workplaces and society. Again, key questions arise: which humans, and where in the loop should they be? If we are to focus our energy on fixing contemporary AI expressions, then we argue that humans should be leading the loop – and not just some humans, but many humans, including First Nations communities. These humans should consider not just our human lives, but Country and ecosystems.

While AI and our human brains are both wired for prediction, the point of difference (at least right now) is that humans have the capacity to challenge those predictions. We can use this capacity for critical thought to our advantage, challenging our assumptions about AI, what it is capable of and whose life and future it is focused on improving.

One of the limitations to addressing the challenges we face with contemporary AI expressions is humans' propensity for automation bias – where humans trust technology over and beyond other entities (including ourselves). Science fiction tends to illustrate our fears, and core among these is a fear that AI expressions and transformations will

develop a mind of their own, and act to further their own priorities and values, whatever these may be. But this fear is misplaced (and a little ironic given the fear settlers have inflicted on the world time and again). It isn't that AI is taking over, it is we who are willingly and often recklessly giving over our human agency to AI, or at least using it to relinquish decision-making responsibility in certain situations and accountability for oppression and violence.

The human propensity for automation bias is illustrated over and over again – we impulsively follow the technology even in the face of incontrovertible evidence that we shouldn't. The phenomenon of 'death by GPS', where we unthinkingly follow navigation systems sometimes to our deaths, is so common it has a name. Here, automation bias can occur even when our own visual system provides clear evidence that we are headed for danger – such as an [unsafe bridge](#) or a [river](#) or a [desert](#). This same inclination to defer to AI expressions is repeated in healthcare and in the financial world, where patient safety and access to loans can be compromised. Our human tendency for automation bias can end up biasing us against other forms of life.

There is evidence that this human deference to AI, through automation bias, may be changing the human capacity to [think critically and rely on memory](#). We may be losing our [sense of direction](#), learning to pay less attention to the environment around us when the technology is present. We have infinite access to information but seem to have challenges with [recalling information](#) as we increasingly rely on technology as a memory bank. We seem to be more distracted in the face of technology, even as we become more dependent on it.<sup>6</sup>

Is the influence of current AI expressions on our thinking something we are comfortable with – especially when we consider that AI relies primarily on physical knowing, and not experiential knowing? Are

we beginning to align our ways of being and knowing with these AI expressions, instead of the other way around?

## A life-centred way forward

Two individual humans, the authors of this book, are unlikely to have all the answers to a complex problem such as AI; nor can we speak on behalf of all, including First Nations communities. Complex problems require emergent and novel practices,<sup>7</sup> including learning from ancestral First Nations knowledges and solutions. Humans may not yet even be able to answer all of the questions around AI presented in this book – but we can still take a first step towards an AI revolution.

We suggest that a life-centred way forward is possible. Humans can reimagine AI transformations through resisting current AI expressions, by fostering relationships with community and Country, and by employing critical reflexivity and action in how we engage with AI. Through these three ‘r’s (resistance, relationships and reflection), we ask you to reconsider how and when, and in what forms, you engage with AI.

### *Resistance – refusing the loss of life*

While efforts are being made to address the life-taking impacts of AI and digital colonialism, a lack of compliance with these frameworks and the ideologies that underpin these frameworks are hampering this. If we look to the Western world’s relationship with technology over time, and the priorities of profit and power that have often shaped it, this is not surprising. Instead of continuing to try to address ‘old’ ways of thinking through ethical guidelines, policies and legislation, we need to work to disrupt the current evolution of technology.

We as authors see calls to action for increased transparency in AI programming, data collection and access, alongside restraints on technology, as a form of resistance to the oppressive, power-hungry forces engrained in society. The AI industry's response to increasing AI regulations is often to 'warn' that regulations will hamper progress and innovation. For example, in 2023, as the European Union was moving to adopt the *EU AI Act*, its landmark law that seeks to govern the way companies develop, use and apply AI, a group of 'AI unicorns' (privately held companies valued at over US\$1 billion) [stated that](#) 'excessive regulation ... could hinder innovation'. This predictable response from for-profit companies suggest regulations are a step in the right direction for a life-centred future. However, the AI lobby is powerful, and their desire to develop and use AI without limits continues despite such legislation, with some tech conglomerates firing entire ethics teams [or seeking workarounds to the Act](#).

Many are practising forms of resistance to AI as it stands today. One example is the creation of the tool Nightshade. [Artists use this program](#) to modify their art image's pixels, disrupting AI program's ability to label images. Another is [groups and communities](#) writing their own AI programs to remove White supremacist and colonial bias, and sharing their perspectives on how to improve AI. These revolutionary reimaginations of AI and technology assert that non-White and non-colonial ways of knowing and experiencing the world are valid and valuable in the AI (r)evolution.

There are many people and beings whose lives are not centred in AI as it stands today – in the resistance to current AI expressions we see a call for social justice around data, a desire to reimagine the continuing colonial story into a decolonised life-centred one, and in doing so supporting our call for an AI revolution.

This links back to earlier discussions about how knowledge, artificial and intelligence are understood. It is clear that there remains a continued dominance of certain bodies – namely, White cis-hetero abled male bodies – when it comes to AI, and that is hugely problematic. Acknowledging this, **this book is also a form of resistance**, by hopefully opening your eyes to a different way of understanding both AI and life.

So, readers, has this discussion about the extent to which AI is ingrained in Western White colonial norms made your trust in AI as it currently stands wane? Have our words sparked you to engage in this call for action for ‘better’ AI – by asking whose life it centres and how AI expressions may centre all life – or do the problems seem insoluble, the motives of profit and power too inextricably linked to technological development? Perhaps this is the exact outcome those in colonial power want – blame the program, not the programmer; blame the system, not the ones orchestrating and sustaining it. Is AI helping us grow beyond the White supremacist roots of Western society or is it a tool to ground these roots deeper, despite the destruction they creates? We have the power of choice in relation to AI. We can choose to resist current expressions of AI by being selective in if, when and how we use them, or by engaging more readily in their development in a manner that serves to disrupt this AI transformation lineage. We believe we have the responsibility to actively participate in this choice – as we are a part of life.

The human, the community, the land and the lived are all essential to our survival and to understanding the integrated complexity of life in relationship to technological transformations. We can all challenge the messaging we have been sold about AI being the solution to all that ails humanity. We defy the notion that the only viable forms of AI are the current versions and that better versions cannot be developed. We

resist the AI and life paradox and look into the eyes of the panoptical digital connoisseur and say, centre life, *all* life, now and always.

*Relationship – core of a life-centred future*

We, the authors, use AI both for research and in life. Writing this book has led us to transform our relationship with it. It has changed the way we think about and engage AI. We are – after the research, critical reflexivity and writing of this book – more nuanced and considered in our use of AI expressions, in a manner that accepts the boundaries of AI, embraces our human intelligence and considers life beyond the human.

Before writing this book, Jess felt overwhelmed by AI and feared its injustices and the violence it has and can continue to inflict. She still holds these fears, but these now sit alongside a stronger sense of power and responsibility to disrupt the current AI trajectory. She is grounded in the ways of her Wiradyuri ancestors to centre Country in her use of AI for future generations. When Michelle began working on this book, she was excited about AI's capacity to enhance equity and access, in addition to its function in an 'efficient' home. As someone who is neurodivergent, and who relies heavily on her Google Assistant for reminders to complete tasks, she believed in the power of AI as a support. Her view has now transformed, so that every time she goes to use any form of a technological expression, she asks: *is the use worth it?* When the answer is yes, and Michelle does use that technological expression, she cannot help but have the shadow thought of the life that was harmed to allow her use.

The process of writing this book and the moments that led to it were facilitated through the internet and computers. While this process was at times frustrating (thank you, slow internet and Microsoft

Word's autoformatting), it was these very technological expressions that allowed our relationship to develop and our ability to begin questioning them in the pages of this book. Together we have realised that through community, collaboration and caring for all life, especially Country, we can resist the current trajectory and revolutionise the future transformations of AI to centre life.

When we consider the different ways we can create more life-centred AI, we suggest that in addition to building AI to reflect Indigenous Data Governance and responsible AI principles, societies in the West (who dominate AI production) also need to work towards valuing and prioritising all forms of life. In this, AI could play a role by shining a light on our biases, injustices and violences and helping us face some hard truths. In fact, most of this book illustrates the power of AI to be a mirror, reflecting the ugly truths of life that many humans continue to deny. Maybe AI can help us to save our world by exposing our existing realities?

Of course, things are rarely that simple, and one factor to consider is that AI can be biased itself. As science journalist Cathleen O'Grady writes in a 2024 [article](#): 'Just like humans, artificial intelligence (AI) is capable of saying it isn't racist, but then acting as if it were.' Valentin Hofmann and colleagues further explore this issue in *Nature*, stating 'AI generates covertly racist decisions about people based on their dialects'.<sup>8</sup> Essentially, it illustrates that even if we work towards alignment, not only does this fail to reduce bias in large language models (LLMs), but it seems to make the problem worse. The AI somehow still ends up having racism covertly built in, which is perhaps a truly reflective ode (the term used ironically) to settler colonialism and White supremacy.

While it might appear that we are 'AI detractors', we see things a little differently. What we are in opposition to is the narrative being

told by a few for the many, and the continuing destruction of life at the hands of humans. We are questioning the current timeline and trajectory of AI expressions. We are challenging the idea that techno-solutionism will work at all with a technology whose lineage isn't centred on life. We consider ourselves strong advocates for a life-centred AI – in which the economic, physical and theoretical impacts are positive for all life, not just the lives of some. While this might seem abstract and utopian, we return you to how our journey began.

To recap: after Michelle attended a Zoom seminar that Jess was presenting at and had the joy of hearing her ideas, Michelle reached out via email. Jess and Michelle's working partnership was facilitated through web-based collaborative writing (via Microsoft and Google) and apps like Signal. We wrote an entire book, managed peer review, and researched articles without ever meeting face to face. What we realised, at the end of all this, is that we are as integral to technological expressions as the hardware is.

Now imagine if the nutrients of these technological expressions – the fossil fuels, the water and the human labour we drew on to communicate – were obtained sustainably, and the economics of this included uplifting all those involved in the supply chain. Technology would truly create and sustain relationships, instead of allowing some while impeding the potential for others. This is the type of AI we are advocating for – forms of AI that embrace and support life in an ethical and responsible way. The next iteration of AI expressions must be about sustaining and valuing life, not just attempting to fix the life-taking aspects of our previous attempts, or further embedding unsustainable economic and social systems that benefit some and harm others. We have examples of other forms of knowledge and ways of being to draw from; it is not impossible to transform our paths in AI

development from ones predicated largely on Western conceptualisation of technological transformations to ones that are led by and steeped in the extensive knowledges of First Nations peoples.

While this book centres Wiradyuri knowledges, there are multiple ways of seeing and understanding the world that are currently excluded or minimised that could also offer important contributions to the AI revolution, particularly around Indigenous Data Governance and Indigenous Data Sovereignty. Consider coding from First Nations teens at the [Lakota AI Code Camp](#) or Māori language revitalisation in Aotearoa (New Zealand), where ‘tribal radio stations Te Hiku Media [are] creating language tools that will enable speech recognition and natural language processing of Te Reo Māori,<sup>9</sup> or augmented reality technologies by [Mikaela Jade in Australia to share stories on Country](#). First Nations communities in Latin America and the Caribbean are producing valuable reports (such as ‘Artificial Intelligence Focused on Indigenous Peoples: Perspectives from Latin America and the Caribbean’<sup>10</sup>), as is the [Indigenous Protocol and Artificial Intelligence \(AI\) Working Group](#),<sup>11</sup> who publish position papers and run workshops on First Nations ways of knowing, AI and the future. All these examples show us how life can be centred to bring First Nations ways of knowing to the forefront of an AI revolution. As Maggie Walter and Tahu Kukutai argue, ‘Indigenous data sovereignty, as an emerging site of science and activism, can mediate the risk of harmful outcomes while providing pathways to collective benefits’<sup>12</sup> when it comes to AI.

The Global Indigenous Data Alliance reminds us that ‘[Indigenous Peoples have always been data collectors and knowledge holders](#)’. Given this and First Nations culture’s survival over 65,000 years, we see Indigenous Data Sovereignty and Data Governance as key to a life-centred AI revolution – specifically, the CARE Principles

for Indigenous Data Governance that were co-developed with First Nations peoples, organisations and governments at the International Data Week and the Research Data Alliance Plenary co-hosted event '[Indigenous Data Sovereignty Principles for the Governance of Indigenous Data Workshop](#)', held in 2018 in Gaborone, Botswana. Building upon Indigenous Data Sovereignty movements and work by Maiam nayri Wingara Aboriginal and Torres Strait Islander Data Sovereignty Collective,<sup>13</sup> US Indigenous Data Sovereignty Network,<sup>14</sup> Te Mana Raraunga Māori Data Sovereignty Network,<sup>15</sup> and Indigenous communities, the CARE principles are:

Collective Benefit for Indigenous Peoples to achieve inclusive development and innovation, improve governance and citizen engagement, and realize equitable outcomes. Benefits accrue when data ecosystems are designed and function to support Indigenous nation and community use and reuse of data; use of data for policy decisions and evaluation of services; and creation and use of data that reflect community values.

Authority to control and govern such data, further affirming the need for 'data for governance.' Indigenous Peoples must have access to data that support Indigenous governance and self-determination. Indigenous nations and communities must be the ones to determine data governance protocols, while being actively involved in stewardship decisions for Indigenous data that are held by other entities.

Responsibility to nurture respectful relationships with Indigenous Peoples from whom the data originate. Aspects of the relationship include investing in capacity development, increasing community data capabilities, and embedding data

within Indigenous languages and cultures. Pursuing these goals fulfills the ultimate responsibility of supporting Indigenous data that advances Indigenous Peoples' self-determination and collective benefit.

Ethics in data practices is representation and participation of Indigenous Peoples, who must be the ones to assess benefits, harms and potential future uses based on community values and ethics.<sup>16</sup>

By drawing on the CARE principles, we can work collectively to ensure that AI moving forward is built in relationship to life – in all its forms.

### *Reflection – looking back and to Country*

To move towards responsible use and integration of AI in our lives, we suggest the need for all humans to reflect on and consider the following principles guided by Indigenous Data Governance and the CARE, principles as these support all life:

- If any AI is being developed without First Nations leadership or, at a minimum, collaboration, then it is not centring community, culture or Country.
- Instead of rushing towards colonial ideas of 'progress' and 'knowledge', consider first learning from some of the oldest living First Nations communities. Specifically, consider the complexity of life in all its forms and respect these for a flourishing environment, and also how these communities have sustained life despite all of the violences they have faced. Consider how we can generate AI expressions that

don't harm Country, but help to sustain it at every step of its development, from conceptualisation through to implementation.

- Seek and support funding and resources for de-colonial and anti-colonial research that explores human decision-making processes and their relationship with life, and seeks to better understand the Country–human–AI interface. There is still so much we do not understand, or fail to acknowledge, about how human decision-making works, so such research seems paramount when considering how we could effectively move towards life-centred AI, and a world where we are engaging AI in decision-making even more than we do today.
- Build in enhanced transparency obligations from companies about the capabilities and limitations of AI expressions and transformations, including datasets used for training and outputs, modelling processes and greater clarity about who is involved in the AI development. Limitations of each AI system should be front and centre, and any system rollouts should include training for humans to illustrate these limitations. In Australia, we regulate the advertisements of certain products, such as pharmaceuticals, given their impact on society – we could consider doing the same for AI.
- In discussing AI, be critical of individualism. Life is a relational community. Consider beyond the immediately seen or felt impacts to our future generations and what our choices and actions now mean for them. When you see the individualism, call it out and advocate for community-focused thinking and action.

- Exercise agency, and critically reflect on how, when and if you engage in AI expressions – and consider your individual and community values when doing so.
- Finally and most crucially, we must develop an enhanced awareness of and educate ourselves about the relationship between AI and life. Many of us need to retrain ourselves to align with life – embodied life – for ourselves, our planet and our future generations.

Efforts to align ourselves towards ‘life’, including Country and future generations, can help counter the perpetuation of violence that humanity, augmented through existing AI, can bring. This illustrates the value of turning towards, not away from, knowledges of First Nations peoples and ways of being that have been maintained for generations, such as those represented by Aboriginal and Torres Strait Islander Nations in so-called Australia. Such knowledges inherently consider that humans aren’t the only life we must prioritise – Country is inseparable in this equation because it is at the core of life itself.

We need to embrace our power to question, to think critically and to consider context before engaging AI expressions. We have a hunch that many humans have been trained to gravitate towards AI because of settler colonial values such as efficiency, individualism and productivity. AI represents a technological expression that is currently devoid of the full extent of human emotions and subjective experiences, and is ‘objective’ to a fault. Those raised in, and who buy into, settler colonialism are likely to put AI on a pedestal because it represents what many in this society view as knowledge. ‘Knowing’ in this worldview is valued when the knowledge is physical, objective

and certain, discoverable and discrete. We are asking you to reflect on this worldview, and ask yourselves whether it is compatible with life.

First Nations ways of knowing counter this view, considering knowledge as situated, embodied, centred in Country and tied to the experiences of the 'knower' and ever-present. AI, in this perspective, can be seen as a knower or knowledge-holder with limitations based on the relationships (or lack thereof) it has with life.

We all have the opportunity to embrace a different way of thinking, one that considers person and place and not just the settler colonial conditioning, which tells us we need more, we need progress, we need it now, we need new and shiny objects at whatever cost. We are often promised enhanced 'efficiency' that never delivers. Just consider smartphones. When they first came out, we were told they would allow us to work from anywhere, making us more efficient and able to experience life. The result, however, is that we work everywhere and more hours than before. Moving away from an idea of efficiency being the cornerstone of life, and towards the idea of sustainable life, in all its forms, we can build a future where AI truly aligns with knowledge.

We challenge each of you to pose the following questions when considering how and when to integrate AI into your lives, in a manner that puts all life at the centre.

1. Who is backing the AI? What are their interests in doing so? Do those interests align with life-affirming interests? If so, whose lives?
2. Whose interests are absent from the AI?
3. What was the intended purpose of the AI when developed, and who contributed to deciding this purpose?

4. Who participated in the development of the AI, and do these developers include those impacted by and users of the AI?
5. What data is used in the development and training of AI, and whose data is omitted?
6. If the AI is integrated into a workplace or into daily life, what are the likely downstream impacts? If the impact is improved efficiency, what is the likely cost of such efficiencies?
7. What is the impact of the AI on our humanity? How would it change the way we think or the relationships we build in life?
8. How does the AI connect, or not connect, to Country and the valuing of the planet? What is the impact of the AI on the planet now and in the future?

While these questions aren't exhaustive – and we encourage you to add your own, drawing from your own cultural context, as you become more critical of AI – they serve as a developing and fluid framework for us to challenge the messaging that for-profit AI companies have created around AI.

The challenge is that being curious involves large amounts of time and energy – it is the opposite of what we have been sold as 'efficiency' (particularly in relation to AI). By engaging in such critical reflection, we can actively resist the settler colonial worldview that efficiency trumps critical reflection. In doing so, we can integrate AI into our lives more responsibly and disrupt the AI transformation lineage when it is not being used in a manner that serves all life.

This critical reflection and curiosity have the potential to make us more aware of who we are as humans, and our unique relationship with the world around us. Let's keep centring life, in all its forms, for a future that is not shaped by AI as it is currently, or by technology's

## A Life-Centred Future

violent history, but one that values and preserves all life unconditionally now and for future generations, just as First Nations cultures have done and continue to do. That is truly revolutionary.



# Acknowledgements

We have reached the end. (Or the beginning? Or even the middle? Remember, it is all related ...)

There are many peoples that have contributed to this book. We offer our sincere thanks to our colleagues and community who peer-reviewed this book and contributed to its journey, including:

- Dr Sue Green, Wiradyuri
- Dr Immaculate Dadiso Motsi-Omoijiade
- Dr Lobna Yassine
- Dr Tim Fawns
- Dr Mandy Truong
- Leon Furze

Their critical review and reflection enhanced this book, and we are grateful for their time and efforts in exploring a life-centred future in the age of this AI transformation.

We are incredibly grateful for the time and editing prowess of Monash University Publishing editor Kate Morgan and director Julia Carlomagno – without whom this book wouldn't exist, and certainly not in its current expression.

Finally, thank you to the readers of this book. We wrote this book in relationship with you and our shared fight for a life-centred world.



# Links

Full citation details for the embedded links in the text are listed below, in order of appearance.

## Introduction

Utochkin, D., ‘Cut the “AI” Bullshit, UCPH’, *University Post*, University of Copenhagen, 12 August 2024.

*Europe Now*, ‘Rethinking the Human in a Multispecies World’ (special feature), 9 November 2021.

Hemmendinger, D., ‘Local Area Network’, *Britannica*, 20 September 2024.

## Chapter 1

Simon, F.M., ‘Artificial Intelligence in the News: How AI Retools, Rationalizes, and Reshapes Journalism and the Public Arena’, *Columbia Journalism Review*, 6 February 2024.

Berry, M.D., ‘Understanding the Advantages and Risks of AI Usage in Healthcare’, *Thomson Reuters*, 27 September 2023.

Lytton, C., ‘AI Hiring Tools May Be Filtering Out the Best Job Applicants’, *BBC*, 16 February 2024.

Senior, K., ‘How Companies are Hiring and Reportedly Firing With AI’, *Forbes*, 4 November 2023.

Hamilton, H. and Ugwuodike, P., ‘A “Black Box” AI System Has Been Influencing Criminal Justice Decisions for Over Two Decades – It’s Time to Open It Up’, *The Conversation*, 26 July 2023.

Manning, C., ‘Artificial Intelligence Definitions’, Stanford University Human-Centered Artificial Intelligence, September 2020.

## Links

Goldman, S., 'Today's AI is "Alchemy", Not Science – What That Means and Why That Matters', *VentureBeat*, 18 September 2023.

Hutson, M., 'AI Researchers Allege that Machine Learning is Alchemy', *Science*, 3 May 2018.

IBM, 'What is Artificial Intelligence (AI)?', 16 August 2024.

Bearman, M. and Ajjawi, R., 'When I Say ... Artificial Intelligence', *Medical Education*, 20 April 2024.

Bender, E., 'Information is Relational: Google's AI Overviews Fails Helpfully Highlight a Source of Danger', *Mystery AI Hype Theater 3000: The Newsletter*, 28 May 2024.

Conrad, K., X (Twitter) Post, 28 April 2024.

Kowalkiewicz, M., 'On Botshit', *The Economy of Algorithms*, 4 January 2024.

Boroditsky, L., 'How Language Shapes the Way We Think', *TED Talk*, 3 May 2018.

Lutkevich, B., 'AI Lawsuits Explained: Who's Getting Sued?', *Tech Target*, 25 June 2024.

Fan, S., 'This Is What Could Happen if AI Content Is Allowed to Take Over the Internet', *Singularity Hub*, 25 July 2024.

'What is Synthetic Data?', *Mostly AI*, 2024.

Abdilla, A. et al., 'Out of the Black Box: Indigenous Protocols for AI', Deakin University, Melbourne, 2021.

Williams, D.H., and Shipley, G.P., 'Enhancing Artificial Intelligence with Indigenous Wisdom', *Open Journal of Philosophy*, vol. 11, no. 1, 2021, pp. 43–58.

Indigenous AI, 'Indigenous Protocol and Artificial Intelligence Working Group', 2024.

'Artificial', Miriam-Webster Dictionary, 16 September 2024.

'Intelligence', Miriam-Webster Dictionary, 20 September 2024.

'Intelligence', APA Dictionary of Psychology, American Psychology Association, 19 April 2018.

Landry, L., 'Why Emotional Intelligence Is Important in Leadership', *Harvard Business School Online*, 3 April 2019.

## Links

Alter, T., 'The Knowledge Argument Against Physicalism', *Internet Encyclopedia of Philosophy*, no date.

### Chapter 2

The Editors, 'Technology', *Britannica*, 14 September 2024.

Copeland, B.J., 'History of Artificial Intelligence (AI)', *Britannica*, 13 September 2024.

Peter, S., 'AI Was Born at a US Summer Camp 68 Years Ago', *The Conversation*, 28 August 2024.

Poor Bear, B., 'Lakota Youth Use Coding and AI to Save Language', *Lakota Times*, 14 February 2024.

Fisher, R. and Elston, B., 'Indigenous Rangers are Burning the Desert the Right Way', *The Conversation*, 28 August 2023.

Kennedy, T. and Miles, M., 'Indigenous Science Can Help Solve Some of the Great Problems of Our Time', *The Conversation*, 14 August 2024.

Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS), 'A Brief Introduction to Indigenous Fishing', 23 February 2021.

World Wildlife Fund, 'Overfishing Puts More Than One-third of All Sharks, Rays, and Chimaeras at Risk of Extinction', 8 September 2021.

World Wildlife Fund, 'Overfishing', no date.

Commonwealth Scientific and Industrial Research Organisation (CSIRO), 'Sustainable Commercial Fisheries', 22 October 2021.

Cavanagh, R., 'The Differences Between Hazard Reduction Burning and Cultural Fire Practices', Australian Museum, 9 October 2020.

NSW Government Environment and Heritage, 'Understanding the Effects of the 2019–20 Fires', 5 June 2024.

Rowe, C. et al., 'Indigenous Fire Management Began More Than 11,000 Years Ago', *The Conversation*, 11 March 2024.

O'Donnell, E., "White Australia has a Black History" NAIDOC Week Poster, 1987', University of Melbourne Archives and Special Collections, no date.

MacroPolo, 'The Global AI Talent Tracker 2.0', no date.

## Links

Singla, A. et al., 'The State of AI in Early 2024: Gen AI Adoption Spikes and Starts to Generate Value', McKinsey & Company, 30 May 2024.

Keary, T., 'Top 10 Countries Leading in AI Research & Technology in 2024', *Techopedia*, 9 April 2024.

United Nations, 'What Is Climate Change?', no date.

Schaefer, A., '10 Side Effects of Sitting Down All Day', *Healthline*, 9 July 2017.

National Museum of Australia, 'First Steam Railway', 2024.

Park, A., 'The Indigenous History of Australia's Railway', *SBS News*, 26 August 2013.

Luke, J., 'Truth-telling to Reimagine Our Nation(s) Histories', *IndigenousX*, 25 January 2019.

### Chapter 3

Credo AI, 'Bias (Social vs. Statistical)', 2024.

O'Grady, C., 'AI Makes Racist Decisions Based on Dialect: Large Language Models Strongly Associated Negative Stereotypes with African American English', *Science*, 6 September 2024.

Soyer, E. and Hogarth, R. M., 'Fooled by Experience', *Harvard Business Review*, May 2015.

'Amazon Scrapped "Sexist AI" Tool', *BBC*, 10 October 2018.

Walters, L., '10 AI-Powered Tools for Reducing Bias in Recruitment', *Pharmweb.jobs*, 1 November 2023.

Rettberg, J.W., 'How Generative AI Endangers Cultural Narratives', *Issues in Science and Technology*, 16 January 2024.

Braue, D., 'Accident Victim Dies after Robotaxis Block Ambulance', *Information Age*, 6 September 2023.

Ferrara, E., 'Eliminating Bias in AI May Be Impossible', *The Conversation*, 19 July 2023.

Martinez, E. and Kirchner, L., 'The Secret Bias Hidden in Mortgage-Approval Algorithms', *The Markup*, 25 August 2021.

Lupini, C. and Kim, P., 'What is a FICO Score?', *Personal Finance*, 20 July 2024.

## Links

Bhutta, N., 'How Much Does Racial Bias Affect Mortgage Lending?', Federal Reserve Bank of Philadelphia, March 2024.

Nunes, F.F., 'FHFA to Replace Classic FICO with More Inclusive Credit Model', *Housingwire*, 24 October 2022.

Johnson, T.L. and Johnson, N.N., 'Police Facial Recognition Technology Can't Tell Black People Apart', *SCIMA*, 18 May 2023.

Dilini, N. et al., 'Cheating Detection in Browser-based Online Exams through Eye Gaze Tracking', in *2021 6th International Conference on Information Technology Research*, December 2021, pp. 1–8.

Australian Human Rights Commission, 'Commission Submission – On Common Difficulties Facing Aboriginal Witnesses', 2005.

Klein, A., 'Black Students Are More Likely to Be Falsely Accused of Using AI to Cheat', *Education Week*, 18 September 2024.

Lazarus, M., 'Will AI Replace Doctors' "Gut Instincts"?', *360info*, 20 December 2023.

Hylving, L. and Lindberg, S., 'Practical Wisdom and Big Data Dilemmas: The Case of the Swedish Transport Administration', 1 May 2021.

Brown, B., 'Driverless Cars: Stopping Dead Seems To Be a Default Setting When They Encounter a Problem', *The Conversation*, 4 January 2024.

Barcelos, G., 'Understanding Bias in Machine Learning Models', *Arize*, 15 March 2022.

Lazarus, M., 'AI, We Need to Talk: The Divide Between Humanities and Objective Truth', *Monash Lens*, 22 June 2023.

Tiku, N. et al., 'This is How AI Image Generators See the World', *The Washington Post*, 31 October 2023.

Thomson, T.J. and Thomas, R.J., 'Ageism, Sexism, Classism and More', *The Conversation*, 10 July 2023.

Robertson, A., 'Google Apologizes for "Missing the Mark" after Gemini Generated Racially Diverse Nazis', *The Verge*, 22 February 2024.

Department of Industry, Science and Resources, 'Australia's AI Ethics Principles', Australian Government, 2024.

Lazarus, M., 'Artists Can Expose the Reality behind AI', *360info*, 8 April 2024.

## Links

‘Mechanical Turk’, *PCMag*, no date.

Kauffman, K. and Williams, A., ‘Turk Wars: How AI Threatens the Workers Who Fuel It’, *Stanford Social Innovation Review*, 11 October 2023.

Perrigo, B., ‘Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic’, *TIME*, 18 January 2023.

Newman, A., ‘I Found Work on an Amazon Website. I made 97 Cents an Hour’, *The New York Times*, 15 November 2019.

Hern, A., ‘TechScape: How Cheap, Outsourced Labour in Africa is Shaping AI English’, *The Guardian*, 16 April 2024.

## Chapter 4

Salleh, A. and Qadar, S., ‘Artificial Intelligence Has Psychological Impacts Our Brains Might Not Be Ready For, Expert Warns’, *ABC News*, 1 May 2024.

Department of Industry, Science and Resources, ‘Australia’s AI Ethics Principles’, Australian Government, 2024.

Kwet, M., ‘Digital Colonialism: The Evolution of US Empire’, *TNI Longreads*, 4 March 2021.

Mackereth, K., ‘A New AI Lexicon: AI Nationalism’, *AI Now*, 19 July 2021.

Parrott, R.J., ‘How White Supremacy Went Global’, *Washington Post*, 19 September 2017.

Igoe, K.J., ‘Algorithmic Bias in Health Care Exacerbates Social Inequities’, Harvard School of Public Health, 12 March 2021.

Silva, A., ‘AI Images of Women from Around the World Have Gone Viral’, *ABC News*, 30 August 2023.

Bhaimiya, S., ‘An Asian MIT Grad Asked AI for Professional Headshot. It Made Her White, With Lighter Skin and Blue Eyes’, *Business Insider*, 1 August 2023.

Sampler, I., ‘Programs to Detect AI Discriminate Against Non-native English Speakers, Shows Study’, *The Guardian*, 11 July 2023.

Mathewson, T.G., ‘AI Detection Tools Falsely Accuse International Students of Cheating’, *The Markup*, 14 August 2023.

Steere, E., ‘The Trouble with AI Writing Detection’, *The Insider Higher Ed*, 18 October 2023.

## Links

- Worrell, T., 'Generative AI in the Classroom Risks Further Threatening Indigenous Inclusion in Schools', *The Conversation*, 6 February 2024.
- Browne, G., 'AI Is Steeped in Big Tech's "Digital Colonialism"', *WIRED*, 25 May 2023.
- Sarder, M., 'From Robodebt to Racism: What Can Go Wrong When Governments Let Algorithms Make the Decisions', *The Conversation*, 5 June 2020.
- O'Donovan, D., "Amateurish, Rushed and Disastrous": Royal Commission Exposes Robodebt as Ethically Indefensible Policy Targeting Vulnerable People', *The Conversation*, 10 March 2023.
- Lazarus, M. and Townsend, J., 'Automation, Uncertainty, and the Robodebt Scheme', *Monash Lens*, 22 March 2023.
- Medhora, S., 'Over 2000 People Died After Receiving Centrelink Robo-debt Notice, Figures Reveal', *ABC News*, 18 February 2019.
- Lazarus, M., 'AI, We Need to Talk: The Divide Between Humanities and Objective Truth', *Monash Lens*, 22 June 2023.
- Mares, P., 'Lost in the Post', *Inside Story*, 12 February 2024.
- Mello, M.M. and Guha, N., 'Understanding Liability Risk from Healthcare AI', Stanford University Human-Centred Artificial Intelligence, 2024.
- Karner, N., 'Israel Accused of Using AI to Target Thousands in Gaza, as Killer Algorithms Outpace International Law', *The Conversation*, 11 April 2024.
- Tsagas, M., 'Human Oversight of AI Systems May Not Be As Effective As We Think', *The Conversation*, 28 August 2024.
- Duffy, M., 'The Rise of the "Machine Defendant"', *The Conversation*, 6 August 2024.
- 'Tech bro', Cambridge Dictionary, no date.
- Browne, G., 'AI Is Steeped in Big Tech's "Digital Colonialism"', *WIRED*, 25 May 2023.
- General Data Protection Regulation (GDPR), 'What is GDPR, the EU's New Data Protection Law?', no date.
- GDPR, 'What are the GDPR Consent Requirements?', no date.

## Links

Stewart, E., 'The Information Banks Keep About You and Your Spending and How It's Used', *ABC News*, 22 March 2019.

Powell, A., 'Rape Culture', *The Conversation*, 17 September 2014.

Das, S., 'Sex Offender Banned from Using AI Tools in Landmark UK Case', *The Guardian*, 21 April 2024.

McArthur, B., 'AI Chatbot Blamed for Psychosocial Workplace Training Gaffe at Bunbury Orison', *ABC News*, 21 August 2024.

McQue, K., 'Child Sexual Abuse Content Growing Online with AI-made Images, Report Says', *The Guardian*, 16 April 2024.

Loreck, J., 'What Does the "Male Gaze" Mean, and What About a Female Gaze?', *The Conversation*, 6 January 2016.

Perdana, A., 'Finding a Fix for Indonesia's Data Protection Problems', *Monash Lens*, 10 May 2024.

Kurmelovs, R., 'Explainer: Medibank Hack', *The Guardian*, 21 October 2022.

Murphy, H., 'Optus Says It Will Defend Allegations It Failed To Protect Confidential Details of 9 Million Customers in Cyber Attack', *ABC News*, 23 May 2024.

Saheer, L.B., 'Four Ways AI Could Help Us Respond to Climate Change', *The Conversation*, 14 February 2024.

The European Space Agency, 'AI Maps Icebergs 10,000 Times Faster Than Humans', 9 November 2023.

Blauth, I., 'Harnessing AI to Fight Global Deforestation', *Phys.org*, 2 September 2024.

Masterson, V., '9 Ways AI Is Helping Tackle Climate Change', *World Economic Forum*, 12 February 2024.

Pritchard, A., '2022 Wrapped: Changing the World's Relationship with Waste', *Greyparrot AI*, 5 January 2023.

The Ocean Cleanup, 'The Largest Cleanup in History', no date.

Powell, S., 'Using AI to Help Save Wildlife After Bushfires', *Australian Financial Review*, 27 September 2023.

Marz, P., 'Generative AI is a Climate Disaster', *Disconnect*, 5 July 2024.

## Links

Ambrose, J., 'Ireland's Datacentres Overtake Electricity Use of All Urban Homes Combined', *The Guardian*, 24 July 2024.

Sui, C., 'Why the World Should Pay Attention to Taiwan's Drought', *BBC News*, 20 April 2021.

Crawford, K., 'Generative AI's Environmental Costs Are Soaring – And Mostly Secret', *Nature*, 20 February 2024.

Duarte, F., 'Number of ChatGPT Users (Aug 2024)', *Exploding Topics*, August 2024.

Keikkila, M., 'Making An Image with Generative AI Uses As Much Energy As Charging Your Phone', *MIT Technology Review*, 1 December 2023.

Rathi, A. and Bass, D., 'Microsoft's AI Push Imperils Climate Goal as Carbon Emissions Jump 30%', *Bloomberg*, 16 May 2024.

Librarianshipwreck, X post, 17 May 2024.

Crawford, K., 'Generative AI's Environmental Costs Are Soaring – And Mostly Secret', *Nature*, 20 February 2024.

World Economic Forum, 'A New Circular Vision for Electronics: Time for a Global Reboot', January 2019.

The World Counts, 'Electronic Waste Facts', no date.

Minevich, M., 'How To Fight Climate Change Using AI', *Forbes*, 8 July 2022.

Lee, S.H. et al., 'The Climate Is Changing So Fast That We Haven't Seen How Bad Extreme Weather Could Get', *The Conversation*, 31 July 2024.

Jiang, Y. et al., 'Repetition Increases Belief in Climate-skeptical Claims, Even for Climate Science Endorsers', *Plos One*, vol. 19, no. 8, e0307294.

Team Stash, '15 Largest AI Companies in 2024', *Stash*, 9 August 2024.

Kanungo, A., 'The Green Dilemma: Can AI Fulfil Its Potential Without Harming the Environment?', *earth.org*, 18 July 2023.

Lakhani, N., 'How US Transition to Electric Cars Threatens Environmental Havoc', *The Guardian*, 24 January 2024.

Amnesty International, 'Child Labour Behind Smart Phone and Electric Car Batteries', 19 January 2016.

Ohnsman, A., 'Battery Push by Tesla And Other EV Makers Raises Child Labor Concerns', *Forbes*, 8 February 2023.

Pattison, P., "Like Slave and Master": DRC Miners Toil for 30p an Hour to Fuel Electric Cars', *The Guardian*, 8 November 2021.

### Chapter 5

Comstock, G., 'Jennifer in Paradise: The Story of the World's First Photoshopped Image', *The Guardian*, 14 June 2014.

'Global Majority', Cambridge Dictionary, no date.

Binn, D. et al., 'Indigenous Ethics and Values', *Pulling Together: A Guide for Researchers*, *Hilkala*, BC Campus, 2021.

Ujomudike, P.O., 'Ubuntu Ethics', *Encyclopedia of Global Bioethics*, 1 January 2022, pp. 2869–81.

Adams, A., 'Family of Man Who Died After GPS Directed Him Off Bridge That Collapsed in 2013 Sues Google', *People*, 21 September 2023.

Dutta, S.B., 'What Is "Death By GPS"? How Technology Is Killing People?', *India Times*, 2 October 2023.

Milner, G., 'Death by GPS: Are Satnavs Changing Our Brains?', *The Guardian*, 25 June 2016.

Bentley, S.V. et al., 'Dumbing Down or Wising Up', *The Conversation*, 13 October 2023.

Leshed, G. et al., 'In-car GPS Navigation', in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Cornell University, Ithaca, 2008, pp. 1675–84.

Pilat, D. and Krastev, S., 'Why Do We Forget Information That We Just Looked Up?', *The Decision Lab*, no date.

Fenelon, L., 'AI Unicorns Warn Europe Against Overregulation', *Cointribune*, 27 November 2023.

Schiffer, Z. and Newton, C., 'Microsoft Lays Off Team That Taught Employees How to Make AI Tools Responsibly', *The Verge*, 14 March 2023.

Nowakowski, T., 'Artists Can Use This Tool to Protect Their Work From A.I. Scraping', *Smithsonian Magazine*, 3 November 2023.

## Links

Selwyn, N., 'Resisting and Reimagining Artificial Intelligence', *Education International*, 25 August 2023.

O'Grady, C., 'AI Makes Racist Decisions Based on Dialect: Large Language Models Strongly Associated Negative Stereotypes with African American English', *Science*, 6 September 2024.

Melendez, C., 'How an AI Coding Camp Is Teaching Native Teens to Preserve Lakota Culture', *NBCU Academy*, 27 August 2024.

Administrator, 'How Mikaela Jade Built Augmented Reality Startup Indigital from Deep in Kakadu National Park', State Library of Queensland Blog, 15 March 2018.

Global Indigenous Data Alliance, 'History of Indigenous Data Sovereignty', no date.

Global Indigenous Data Alliance, 'CARE Principles for Indigenous Data Governance', no date.



# Notes

## Introduction

- 1 Russ-Smith, J., 'Embodiment of Sovereignty', in Bennett, B. and Green, S., *Our Voices: Aboriginal Social Work* (2nd Edition), Red Globe Press, London, 2019, p. 239.
- 2 Russ-Smith, J., 'Giyira: Indigenous Women's Knowing, Being and Doing as a Way to End War on Country', in MacKenzie, M. and Wegner, N., *Feminist Solutions for Ending War*, Pluto Press, London, 2021, p. 16.
- 3 Wiradyuri can be spelt with a 'dy' or a 'dj' (Wiradjuri). Jess predominantly uses Wiradyuri and we explore why later in this book.
- 4 Grant, S. and Rudder, J. (eds), *A Grammar of Wiradjuri Language*, Restoration House, O'Connor, 2014.
- 5 Bayne, S 2018, 'Posthumanism: A navigation aid for educators', *on\_education: Journal for Research and Debate*, vol. 2, no. 1, [https://doi.org/10.17899/on\\_ed.2018.2.1](https://doi.org/10.17899/on_ed.2018.2.1)
- 6 Definition of posthumanism and quote from Bayne in Bayne, S., 'Posthumanism: A Navigation Aid for Educators', *Journal for Research and Debate*, vol. 2, no. 1, 2018, pp. 1–7.
- 7 Grant, S. and Rudder, J., 2014, op. cit., pp. 7–8.
- 8 Dennis, R.M., 'Social Darwinism, Scientific Racism, and the Metaphysics of Race', *Journal of Negro Education*, vol. 64, no. 3, 1995, pp. 243–52.
- 9 Ardill, A., 'Sociobiology, Racism and Australian Colonisation', *Griffith Law Review*, vol. 18, no. 1, 2009, pp. 82–113.
- 10 Grant, S. and Rudder, J. (eds), *A New Wiradjuri Dictionary*, Restoration House, O'Connor, 2010, p. 315.
- 11 *ibid.*, pp. 315, 405.
- 12 *ibid.*, pp. 159.
- 13 Grant, S. and Rudder, J., 2014, op. cit., pp. 7–8.
- 14 Chilisa, B., *Indigenous Research Methodologies*, Sage Publications, California, 2012.
- 15 Moreton-Robinson, A., *The White Possessive: Property, Power and Indigenous Sovereignty*, University of Minnesota Press, Minneapolis, 2015.
- 16 Russ-Smith, 'Wayanha Knowledge Hierarchies: Sovereign Cosmologies Standing Across Generations', *Diaspora, Indigenous, and Minority Education*, 2024, pp. 1–8.

- 17 Grant, S. and Rudder, J., 2014, op. cit.
- 18 Pitner, B.H., *The Crime Without a Name: Ethnocide and the Erasure of Culture in America*, Catapult, New York, 2021.

## Chapter 1

- 1 Grant, S. and Rudder, J. (eds), *A Grammar of Wiradjuri Language*, Restoration House, O'Connor, 2014.
- 2 Russ-Smith, J., 'Balayanhi Wiradyuri Garingundhi. We are sovereign my granddaughter. Embodying Wiradyuri women's sovereignty through Wiradyuri knowing, being and doing', Doctoral Thesis, Charles Sturt University, 2023.
- 3 Gaynes, R., 'The Discovery of Penicillin – New Insights After More Than 75 Years of Clinical Use', *Emerging Infectious Diseases*, vol. 23, no. 5, 2017. <http://dx.doi.org/10.3201/eid2305.161556>
- 4 Grant, S. and Rudder, J., 2014, op. cit., pp. 18, 196.
- 5 Russ-Smith, J., 2023, op. cit., p. 131. Cites Grant, S. & Rudder, J., 2010, op. cit., pp. 397, 454, 469.
- 6 See Marini, I., et al., 'Psychosocial Aspects of Disability', in *Insider Perspectives and Strategies for Counselors (Second Edition)*, Springer Publishing, New York, 2007.
- 7 Defined by the National Human Genome Research Institute (NHGRI) as follows: 'Eugenics is the scientifically inaccurate theory that humans can be improved through selective breeding of populations; Eugenacists believed in a prejudiced and incorrect understanding of Mendelian genetics that claimed abstract human qualities (e.g., intelligence and social behaviors) were inherited in a simple fashion. Similarly, they believed complex diseases and disorders were solely the outcome of genetic inheritance.' See 'Eugenics and Scientific Racism', NHGRI, 18 May 2022, <https://www.genome.gov/about-genomics/fact-sheets/Eugenics-and-Scientific-Racism>
- 8 We use identity-first language, for example 'disabled peoples', as opposed to person-first language such as 'people with disabilities', to reflect and honour many disabled peoples and community's language preference, which communicates their disability as a part of their identity. Identity-first language is also a way disabled people may express pride in their identity and experiences, and a way to highlight the social construction of disability. We also note that language preference is not concrete and it is every person's right to determine their preferred language. For more on identity-first language, see Best, K.L. et al., 'Language Matters! The Long-Standing Debate between Identity-first Language and Person-first Language', *Assistive Technology*, vol. 34, no. 2, 2022, pp. 127–28.
- 9 Ben-Zeev, D., 'DSM-V and the Stigma of Mental Illness', *Journal of Mental Health*, vol. 19, no. 4, 2010, pp. 318–27. Scholar and activist Lisa

Archibald states, “The term “mad” has been reclaimed intentionally as a deliberate interruption or sabotage of the dominant psychiatric perspective. It challenges the entire basis of the medical framework, which is that people have illnesses or disorders. Prior to the last 200 years in history, “madness” was a widely accepted term in society and was not a medical term. The reclamation of “mad” is a provocation to psychiatry as it is a complete rejection of their diagnostic expertise and power”; see Archibald, L., ‘Mad Activists: The Language We Use Reflects Our Desire for Change’, *Mad in America*, 23 September 2021.

- 10 Palanca-Castan, N. et al., ‘Towards an Interdisciplinary Framework about Intelligence’, *Heliyon*, vol. 7, no. 2, 2021, p. 3.
- 11 Grant, S. and Rudder, J., 2014, op. cit., p. 18, 196.
- 12 Grant, S. and Rudder, J. (eds), *A New Wiradjuri Dictionary*, Restoration House, O’Connor, 2010, p. 469.
- 13 Hird, C. et al., ‘Moving Beyond Ontological (Worldview) Supremacy: Indigenous Insights and a Recovery Guide for Settler-colonial Scientists’, *Journal of Experimental Biology*, vol. 226, no. 12, 2023, p. 3.
- 14 Moreton-Robinson, A., *The White Possessive: Property, Power, and Indigenous Sovereignty*, University of Minnesota Press, Minneapolis, 2015, p. xii.
- 15 Russ-Smith, J., ‘Embodiment of Sovereignty’, in Bennett, B. and Green, S., *Our Voices: Aboriginal Social Work* (2nd Edition), Red Globe Press, London, 2019, p. 239.
- 16 King, R., *Here Be Monsters: Is Technology Reducing Our Humanity?*, Monash University Publishing, Melbourne, 2023.
- 17 Grant, S. and Rudder, J., op. cit., 2010.

## Chapter 2

- 1 Russ-Smith, J., ‘Balayanhi Wiradyuri Garingundhi. We are sovereign my granddaughter. Embodying Wiradyuri women’s sovereignty through Wiradyuri knowing, being and doing’, Doctoral Thesis, Charles Sturt University, 2023, p. xi.
- 2 Grant, S. and Rudder, J. (eds), *A New Wiradjuri Dictionary*, Restoration House, O’Connor, 2010, p. 101.
- 3 *ibid.*, p. 405.
- 4 Frazer, R. et al., ‘Indigenous Articulations of Social Media and Digital Assemblages of Care’, *Digital Geography and Society*, vol. 3, 2022, 100038.
- 5 <https://aiatsis.gov.au/explore/fishing>
- 6 Steffensen, V., *Fire Country: How Indigenous Fire Management Could Help Save Australia*, Hardie Grant Explore, 2020.
- 7 Cahir, F. et al., ‘Aboriginal Fire-management Practices in Colonial Victoria’, *Aboriginal History*, no. 45, 2021, pp. 109–130.

- 8 Scates, B. and Yu, P., 'De-Colonizing Australia's Commemorative Landscape: "Truth-Telling," Contestation and the Dialogical Turn', in *Colonial Violence and Monuments in Global History*, Routledge, London, pp. 26–48.
- 9 Grant, S. and Rudder, J., op. cit., 2010. p. 248.
- 10 Wolf, M., 'The Future of "Communist Capitalism" in China', *Financial Times*, 13 March 2024.
- 11 Boyer, P.S., *American History: A Very Short Introduction*, Oxford University Press, New York, 2012, p. 67.
- 12 The Stolen Generations refers to Aboriginal and Torres Strait Islander children in Australia who were and continue to be forcibly removed from their families, communities and Countries. These forced removals were and are a part of genocide enforced through colonial and racist policies, practices and legislation. See [The Healing Foundation](#) for further information.
- 13 Racialisation refers to 'the processes by which a group of people is defined by their "race": 'Processes of racialization begin by attributing racial meaning to people's identity and, in particular, as they relate to social structures and institutional systems, such as housing, employment, and education. In societies in which "White" people have economic, political, and social power, processes of racialization have emerged from the creation of a hierarchy in social structures and systems based on "race". The visible effects of processes of racialization are the racial inequalities embedded within social structures and systems.' (*Encyclopedia of Race, Ethnicity, and Society*, 2008)
- 14 Williams, A. and Gerbu, T., 'The Exploited Labor Behind Artificial Intelligence', *Noema*, 13 October 2022.
- 15 Ren, S. and Wierman, A., 'The Uneven Distribution of AI's Environmental Impacts', *Harvard Business Review*, 15 July 2024.

### Chapter 3

- 1 Doshi, A.R. and Hauser, O.P., 'Generative AI Enhances Individual Creativity But Reduces the Collective Diversity of Novel Content', *Science Advances*, vol. 10, no. 28, 2024, eadn5290.
- 2 Frankenberg, R., *White Women, Race Matters: The Social Construction of Whiteness*, University of Minnesota Press, Minneapolis, 1993, and Moreton-Robinson, A., *The White Possessive: Property, Power and Indigenous Sovereignty*, University of Minnesota Press, Minneapolis, 2015.
- 3 Grant, S. and Rudder, J. (eds), *A New Wiradjuri Dictionary*, Restoration House, O'Connor, 2010, p. 31.
- 4 See <https://www.netflix.com/au/title/81328723>
- 5 Cave, S. and Dihal, K., 'The Whiteness of AI', *Philosophy & Technology*, vol. 33, no. 4, 2020, pp. 685–703.

- 6 Lazarus, M. D., Truong, M., Douglas, P., & Selwyn, N. (2024). Artificial intelligence and clinical anatomical education: Promises and perils. *Anatomical Sciences Education*, 17(2), 249–262. DOI: 10.1002/ase.2221
- 7 Martin, KB. (2018). ‘Differences aren’t deficiencies’, *tobii*, <https://www.tobii.com/blog/differences-arent-deficiencies>
- 8 Gorman, R., ‘Disability Data Justice from the Ground Up: A Practice-Led, Participatory Co-Design Approach to Building an AI Search Engine and Data Repository for Local, National, and Transnational Disability Organizations’, *Critical Studies: An International and Interdisciplinary Journal*, vol. 18, no. 1, 2024, pp. 28–43.
- 9 Zada, Z. et al., ‘A Shared Model-based Linguistic Space for Transmitting Our Thoughts from Brain to Brain in Natural Conversations’, *Neuron*, 2 August 2024, and Zada, Z., ‘AIs Encode Language Like Brains Do – Opening a Window on Human Conversations’, *The Conversation*, 3 August 2024.
- 10 Weidinger, L., ‘Ethical and Social Risks of Harm from Language Models’, *arXiv*, Cornell University, 8 December 2021.
- 11 Of course, the terminology of ‘Turk’ also has racist connotations including [demeaning stereotypes of Turkish peoples as savage](#). Some online deny these connotations. However, it is interesting to see how terminology applied in the context of AI can create positive associations with some, and derogatory associations with others, depending on the relationships of power.
- 12 Stephens, E., ‘The Mechanical Turk: A Short History of ‘Artificial Artificial Intelligence’, *Cultural Studies*, vol. 37, no. 1, 2022, pp. 65.

## Chapter 4

- 1 Bryant, J.H., ‘We Must Approach AI Like the Automobile’, *TIME*, 18 September 2024.
- 2 Bryson, J.J., ‘The Artificial Intelligence of the Ethics of Artificial Intelligence: An Introductory Overview for Law and Regulation’, in Dubner, M.D. et al. (eds), *The Oxford Handbook of Ethics of AI*, Oxford Academic, Oxford, 2020, pp. 2–25.
- 3 Gramsci, A., ‘Selections from the Prison Notebooks’, in *The Applied Theatre Reader*, Routledge, London, 1992, pp. 141–42.
- 4 Wolfe, P., ‘Settler Colonialism and the Elimination of the Native’, *Journal of Genocide Research*, vol. 8, no. 4, 2006, pp. 387–409.
- 5 Kingston, L., ‘The Destruction of Identity: Cultural Genocide and Indigenous Peoples’, *Journal of Human Rights*, vol. 14, no. 1, 14 January 2015, pp. 63–83.
- 6 Obermeyer, Z. et al., ‘Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations’, *Science*, vol. 366, no. 6464, 2019, pp. 447–53.

- 7 McKnight, L. and Shipp, C., “‘Just a Tool’? Troubling Language and Power in Generative AI Writing’, *English Teaching: Practice & Critique*, vol. 23, no. 1, pp. 23–35.
- 8 <https://theconversation.com/generative-ai-in-the-classroom-risks-further-threatening-indigenous-inclusion-in-schools-222254>
- 9 Hearn, A., ‘The Collateralized Personality: Creditability and Resistance in the Age of Automated Credit-scoring and Lending’, *Cultural Studies*, vol. 37, no. 1, 2023, pp. 123–48.
- 10 Abbasi, J. and Hswen, Y., ‘Blind Spots, Shortcuts, and Automation Bias – Researchers are Aiming to Improve AI Clinical Models’, *JAMA*, vol. 331, no. 11, 2024, pp. 903–06.
- 11 Nerantzi, E. and Sartor, G., ‘Hard AI Crime: The Deterrence Turn’, *Oxford Journal of Legal Studies*, vol. 44, no. 3, Autumn 2024, pp. 673–701.
- 12 The authors are cautious about using the language of ‘colonised’ about sovereign Indigenous Nations. This is not to dismiss the reality or ongoing violent impacts of colonialism but to acknowledge that sovereignty was never ceded. Therefore, in certain contexts we use inverted commas around the term ‘colonised’.
- 13 Russ-Smith, J., ‘Sovereign Fire: Indigenous Women and Solutions to War’, in McKenzie, M. and Wegner, N. (eds), *Feminist Solutions for Ending War*, Pluto Press, London, p. 22.
- 14 Howitt, R., ‘Unsettling the Taken (For Granted)’, *Progress in Human Geography*, vol. 22, no. 2, 2020, pp. 193–215.
- 15 Pascoe, B., *Dark Emu: Aboriginal Australia and the Birth of Agriculture*, Magabala Books, Broome, 2018.
- 16 The gendered language of tech ‘bros’ also highlights how technology is not exempt from sexism. See Xuan-Thao Nguyen’s 2021 article ‘[#MeToo Innovators: Disrupting the Race and Gender Code by Asian Americans in the Tech Industry](#)’ for more on how this sexism structurally, economically and socially disadvantages Asian women innovators in the tech industry.
- 17 Enemark, C., *Ethics of Drone Strikes: Restraining Remote-Control Killing*, Edinburgh University Press, Edinburgh, 2021.
- 18 Cave, S. and Dihal, K., ‘The Whiteness of AI’, *Philosophy & Technology*, vol. 33, no. 4, 2020, p. 687.
- 19 See Global Indigenous Data Alliance website: <https://www.gida-global.org/>
- 20 Tapu, I.F. and Fa’agau, T.K.I., ‘A New Age Indigenous Instrument: Artificial Intelligence & Its Potential for (De)Colonialized Data’, *Harvard Civil Rights–Civil Liberties Law Review*, 2022, pp. 57, 715.
- 21 Gray, C. M., Santos, C., Bielova, N., Toth, M., & Clifford, D. (2021, May). ‘Dark patterns and the legal requirements of consent banners: An interaction criticism perspective.’ In Proceedings of the 2021 CHI conference on human factors in computing systems, pp. 1–18.

- 22 Bannihatti Kumar, V. et al., 'Finding a Choice in a Haystack: Automatic Extraction of Opt-out Statements from Privacy Policy Text', in *Proceedings of The Web Conference 2020*, 20 April 2020, pp. 1943–54.
- 23 Information and Privacy Commission, 'Fact Sheet – Consent', NSW Government, June 2023.
- 24 *ibid.*
- 25 Prainsack, B. and Forgó, N., 'Why Paying Individual People for Their Health Data is a Bad Idea', *Nature Medicine*, vol. 28, no. 10, pp. 1989–91.
- 26 Neilen, D., 'Informed Consent, An Ongoing Conversation', *Narrative Inquiry in Bioethics*, vol. 3, no. 1, 2013, E10–E12.
- 27 Andreotta, A.J. et al., 'AI, Big Data, and the Future of Consent', *AI & Society*, vol. 37, no. 4, 2022, pp. 1715–28.
- 28 Jones, M.L. et al., 'AI and the Ethics of Automating Consent', *IEEE Security & Privacy*, vol. 16, no. 3, 2023, pp. 64–72.
- 29 Bryson, J.J., 'The Artificial Intelligence of the Ethics of Artificial Intelligence: An Introductory Overview for Law and Regulation', in Dubner, M.D. et al., *op cit.*, pp. 2–25.
- 30 Andreotta, A.J., *op. cit.*
- 31 Hine, E., 'Informed Digital Consent for Use of AI Systems Grounded in a Model of Sexual Consent', in *Proceedings of the 1st International Conference on AI for People: Towards Sustainable AI, CAIP 2021, 20–24 November 2021, Bologna, Italy*, December 2021.
- 32 Bowling, J., et al., 'Disclosing and Reporting of Consent Violations Among Kink Practitioners in the United States', *Violence Against Women*, vol. 30, no. 6–7, 2024, pp. 1453–76.
- 33 Powers, T.M. and Ganascia, J.G., 'The Ethics of the Ethics of AI', in Dubber, M. et al., *op. cit.*, pp. 25–51.
- 34 Yeung, K. et al., 'AI Governance by Human Rights–Centered Design, Deliberation, and Oversight', in *ibid.*, pp. 77–106.
- 35 Bartky, S.L., 'Foucault, Femininity, and Modernization of Patriarchal Power', in Weitz, R., *The Politics of Women's Bodies*, Oxford University Press, Oxford, 1990, pp. 76–98.
- 36 Russ-Smith, J., 'Balayanhi Wiradyuri Garingundhi. We are sovereign my granddaughter. Embodying Wiradyuri women's sovereignty through Wiradyuri knowing, being and doing', Doctoral Thesis, Charles Sturt University, 2023.
- 37 Dhar, P., 'The Carbon Impact of Artificial Intelligence', *Nature Machine Intelligence*, vol. 2, no. 8, 2020, pp. 423–25.
- 38 Russ-Smith, J., 'Giyira: Indigenous Women's Knowing, Being and Doing as a Way to End War on Country', in MacKenzie, M. and Wegner, N., *Feminist Solutions for Ending War*, Pluto Press, London, 2021, p. 16.

- 39 Coleman, J. (2023), 'AI's Climate Impact Goes beyond Its Emissions'. *SCLAM*, <https://www.scientificamerican.com/article/ais-climate-impact-goes-beyond-its-emissions>

## Chapter 5

- 1 Russell, S., *Human Compatible: AI and the Problem of Control*, Penguin, London, 2019.
- 2 Gabriel, I., 'Artificial Intelligence, Values, and Alignment', *Minds and Machines*, vol. 30, no. 3, 2020, pp. 411.
- 3 Hodson, G. et al., 'Privilege Lost: How Dominant Groups React to Shifts in Cultural Primacy and Power', *Group Processes & Intergroup Relations*, vol. 25, no. 3, 2022, pp. 625–41.
- 4 Panahi, O., 'Could There Be A Solution To The Trolley Problem?', *Philosophy Now*, 2024.
- 5 Stoll, L.C. et al., 'Four: Fatphobia', in *Global Agenda for Social Justice 2*, Policy Press, Bristol, 6 September 2022. <https://doi.org/10.51952/9781447367420.ch004>
- 6 Firth, J. et al., 'The "Online Brain": How the Internet May Be Changing Our Cognition', *World Psychiatry*, vol. 18, no. 2, 2019, pp. 119–29.
- 7 Lazarus, M.D. and Funtowicz, S., 'Learning Together: Facing the Challenges of Sustainability Transitions by Engaging Uncertainty Tolerance and Post-normal Science', *Sustainable Earth Reviews*, vol. 6, issue 18, 2023. <https://doi.org/10.1186/s42055-023-00066-3>
- 8 Hofmann, V. et al., 'AI Generates Covertly Racist Decisions about People Based on Their Dialect', *Nature*, vol. 633, no. 8028, 2024, pp. 1–8.
- 9 Walter, M. and Kukutai, T., 'Artificial Intelligence and Indigenous Data Sovereignty', input paper for the Horizon Scanning Project *The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve Our Wellbeing*, on behalf of the Australian Council of Learned Academies (ACOLA), 2019, p. 7.
- 10 Zepeda, G. et al., 'Inteligencia artificial centrada en los pueblos indígenas: perspectivas desde América Latina y el Caribe', UNESCO Office Montevideo and Regional Bureau for Science in Latin America and the Caribbean, 2023. <https://unesdoc.unesco.org/ark:/48223/pf0000387814>
- 11 See Indigenous AI: <https://www.indigenous-ai.net/>
- 12 Walter, M. and Kukutai, T., op. cit., p. 2.
- 13 See Maïam nayri Wingara website: <https://www.maïamnayriwingara.org/>
- 14 See US Indigenous Data Sovereignty Network website: <https://usindigenousdatanetwork.org/>
- 15 See Te Mana Raraunga Māori Data Sovereignty Network website: <https://www.temanararaunga.maori.nz/>
- 16 Carroll, S.R. et al., 'The CARE Principles for Indigenous Data Governance', *Data Science Journal*, vol. 19, no. 43, 2020, pp. 1–12.

## About the Authors

**Jessica Russ-Smith** is a sovereign Wiradyuri Wambuul woman, First Nations Curriculum and Pedagogy Coordinator in Social Work, Assistant Deputy Head of School Allied Health and Chair of the Indigenous Research Ethics Advisory Panel at ACU. Jess is a two-time University Medallist, and her research focuses on sovereignty, decolonising, disrupting Whiteness and AI.

**Michelle D. Lazarus** is Director of the Centre for Human Anatomy Education and Deputy Director of the Centre for Scholarship in Health Education at Monash University. She is an award-winning educator, a recipient of the Australian Universities Teaching excellence award and the author of *[The Uncertainty Effect](#)*. She continues to journey into uncertainty.

