MONASH University

MONASH DATA FUTURES INSTITUTE

AI FOR SOCIAL GOOD?

Australian public attitudes toward AI and society

Neil Selwyn, Beatriz Gallo Cordoba, Mark Andrejevic & Liz Campbell

August 2020

Copies of this report may be obtained by contacting:

Monash Data Futures Institute
E: datafutures@monash.edu
W: monash.edu/ai-data-science

13 Rainforest Walk
Level 2, Room 236
Monash University
Clayton VIC 3800

CRICOS provider: Monash University 00008C

Attribution:
Professor Neil Selwyn – Faculty of Education
Dr Beatriz Gallo Cordoba – Faculty of Education
Professor Mark Andrejevic – Faculty of Arts
Professor Liz Campbell – Faculty of Law
AI for social good? Australian public attitudes toward AI and society.
Melbourne, Monash University

MONASH University

MONASH DATA FUTURES INSTITUTE

The Monash Data Futures Institute represents the largest cohort of AI and Data Science researchers in the Asia-Pacific. Together with interdisciplinary specialists from law, education, medicine, sociology and anthropology, we partner with national funding bodies, government and industry to research and develop essential and transformative AI as a force for social good.

Contents

Contents

Tables

Figures

> **AI is humanity's new frontier. Once this boundary is crossed, AI will lead to a new form of human civilization...**
>
> **Audrey Azoulay (Director-General of UNESCO)**

Introduction

Artificial Intelligence is now influencing almost all aspects of society: employment; manufacturing; telecommunications; banking and finance; health services and even our national security. The abundance of research into the legal, ethical and societal implications of AI across industry and government reflects its potential for an enduring, transformative impact. However, the views of the public remain underrepresented.

The Monash Data Futures Institute (in partnership with the Faculties of Law, Education and Arts), is proud to present this first major study of Australian's attitudes to AI. We believe it to be vital to paving the way for more frequent and systematic contributions of public opinion, to the formation of public policy, shaping the future of AI technology in Australia and educating the public about AI's benefits and risks

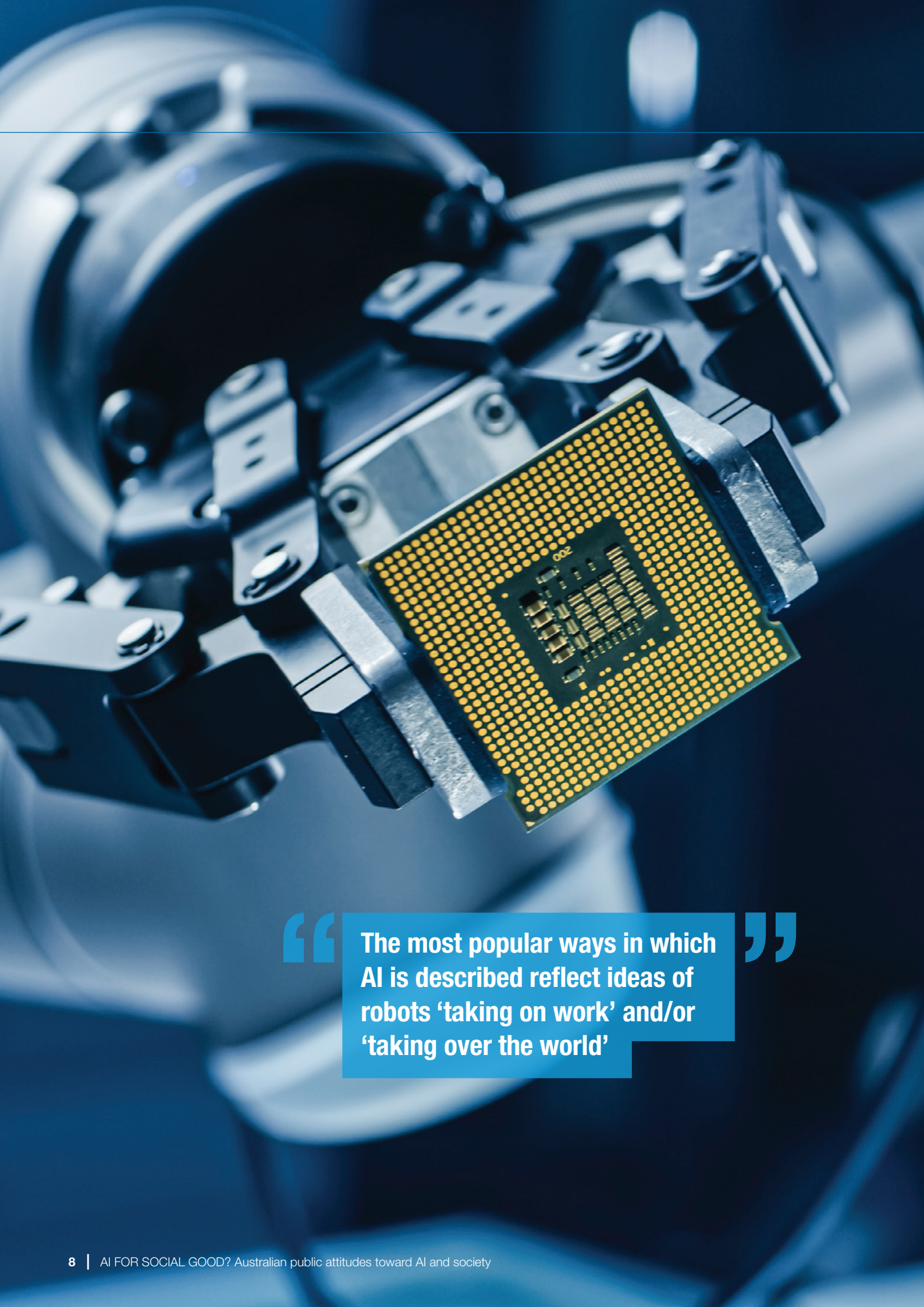
The report's findings show very strong public support for the establishment of a new regulatory body to address AI development, as well as support for the increased governance and oversight of AI development through legislation and industry guidelines. The most immediate priority arising from our findings is the development of public education efforts to enhance what might be termed 'public understandings of AI'.

Similarly, our findings point to high levels of public trust in CSIRO, university researchers and bodies such as the Office of the Chief Scientist of Australia and Australian Human Rights Commissioner – therefore suggesting the benefits of these organisations playing prominent public-facing roles in any efforts to manage and oversee AI development in the future.

Perhaps the most immediate priority arising from our findings is the development of public education efforts to enhance what might be termed 'public understandings of AI'. Indeed, our survey suggests that many people will change their initial opinions and preconceptions about AI when provided with further information, examples and questions.

This survey was administered as a scoping study rather than an exercise in providing specific recommendations or agendas for further action. Nevertheless, our findings suggest that further dialog amongst policymakers and the broader Australian 'AI community' would be of value.

Professor Joanna Batstone
Director, Monash Data Futures Institute
August 2020



> **The most popular ways in which AI is described reflect ideas of robots 'taking on work' and/or 'taking over the world'**

Executive Summary

The past ten years or so have seen artificial intelligence (AI) technologies become a prominent topic of discussion across Australian society. Yet, the ongoing implementation of these technologies remains a highly contested topic. This Monash Data Futures Institute report presents one of the first comprehensive pictures of Australian public understandings, attitudes and opinions relating to AI and society.

Based on a nationally-representative public opinion survey of over 2000 Australian adults, the report examines key areas of public understanding, optimism and concern regarding the societal application of AI technologies. As industry and policy-makers continue to develop, implement and manage AI across most areas of Australian society, this report explores the often-overlooked views of the general public – in many ways, the ultimate 'end users' of these powerful technologies.

Key findings from the survey include:

While nearly nine-in-ten people are aware of the term, the majority of the Australian public consider themselves to have little knowledge of 'artificial intelligence'. Just over one-quarter of respondents described themselves as knowing 'a fair bit' or 'a lot' about AI. These are most likely to be people with computer science or programming expertise, aged between 18–24 years, and/or from households where a language other than English is spoken.

People's immediate understandings of AI are varied. The most popular ways in which AI is described reflect ideas of robots 'taking on work' and/or 'taking over the world'. That said, many respondents have more sophisticated understandings – for example, making immediate associations with computers being programmed to perform tasks, learning from data, and displaying human-like thinking.

There are relatively high levels of support from the Australian public for the development of AI. Having completed the survey, around two-thirds of respondents (63.5%) stated that they 'somewhat support' or 'strongly support' the development of AI. In comparison, only 12% describe themselves as either 'somewhat' or 'strongly' opposed to AI.

Many people will change their opinions on the basis of receiving more information about AI, and being asked to think through issues relating to AI and society. Specifically, we found 43% of the respondents who initially considered themselves 'opposed' to the development of AI to shift subsequently to either a 'neutral' or 'supportive' stance once having engaged with all the survey questions.

We found consistently high levels of support for the use of AI to address social, humanitarian and environmental challenges. Particularly strong support was expressed for the use of AI to address challenges in areas such as health and medicine, alongside environmental challenges and crisis response. The only area with notably lower levels of support is the use of AI to generate culture for popular consumption (such as films, books, music or art).

The majority of the public agrees with the need for industry guidelines and/or new legislation to address all possible harms associated with AI. We find high levels of support for legislation and industry governance related to banning the use of lethal autonomous weapons, ensuring the safety of autonomous vehicles, and protecting data privacy.

The Australian public has high levels of confidence in CSIRO and university researchers to develop AI in the best interests of the public. The lowest levels of confidence to develop AI in the best interests of the public are expressed for Amazon and Facebook.

In terms of managing the development of AI, high levels of confidence are expressed for CSIRO, as well as independent government bodies such as the Office of The Chief Scientist, and the Australian Human Rights Commissioner.

There is very strong support (87.3% of respondents) for establishing a new regulatory body in Australia to govern and promote responsible innovation in the area of AI.

The Australian public is especially hopeful about the potential benefits of AI in healthcare and medicine for advanced diagnosis, development of medicine and disease treatments. Conversely, the prospect of the increased use of AI in the workplace is seen in mixed terms. The most prevalent fears expressed by our survey respondents relate to AI-based surveillance and loss of privacy, alongside the misuse of AI technology by governments and companies acting with malintent.

Notwithstanding specific concerns, our survey finds the Australian public to be generally optimistic about the impact of AI on their lives and society. The majority of the public (69.4%) agrees that AI will do more 'social good' than 'social harm'.

While not necessarily convinced that such technology will ever exist, most people imagine that AI capable of exceeding human intelligence would have an overall 'good' or 'more or less neutral' impact on humanity. Only 5.1% of respondents see the possible future development of high-level machine intelligence as 'extremely bad' and/or maybe even 'leading to human extinction'.

“The Australian public is especially hopeful about the potential benefits of AI in healthcare and medicine for advanced diagnosis, development of medicine and disease treatments”

> **69.4% of the public agree that AI will do more 'social good' than 'social harm'**

Background

This report examines Australian public attitudes toward the increased societal use of artificial intelligence (AI) technologies. At present, the growing use of AI technologies remains a highly contested topic. IT industry and policymakers continue to make claims about the likely transformative outcomes of 'smart cities', 'industrial revolution 4.0' and 'precision medicine'. At the same time, there are growing public and regulatory concerns regarding the over-reach of technological advances in areas such as automated transport, dataveillance and biometric monitoring. In the light of recent controversies over the rollout of COVID-19 tracing apps, racially-biased image recognition, and various other perceived injustices, the ongoing automation and datafication of Australian society is subject to widespread enthusiasms *and* concerns.

On one hand, AI technology has developed rapidly over the past decade – with advances in data storage, the generation of 'big data' sets, and sophisticated computational techniques prompting heightened expectations for AI to be a defining innovation of our time. These are sophisticated technological advances that are capable of powerful outcomes. In tangible terms, various aspects of everyday life are already seeing the benefit of computational systems capable of specialist expert behaviour – from global weather forecasting to traffic flow management. Alongside these tangible examples, are more speculative enthusiasms over the possible development of autonomous agents with high-level 'general intelligence' comparable to human intelligence. In all these cases, then, AI is understandably anticipated to be driving societal development over the next few decades and beyond. As Audrey Azoulay (Director-General of UNESCO) reasons:

> ***"AI is humanity's new frontier. Once this boundary is crossed, AI will lead to a new form of human civilization. … We are faced with a crucial question: what kind of society do we want for tomorrow? The AI revolution opens up exciting new prospects, but the anthropological and social upheaval it brings in its wake warrants careful consideration".*** Audrey Azoulay (Director-General of UNESCO)

These heightened expectations have been accompanied by growing scrutiny (and some scepticism) of the actual implementation of AI technologies to date across different areas of society. Key here are questions over the capacity of these sophisticated computational technologies to address equally sophisticated social, cultural, political and historical issues. These tensions are reflected in suspicions of "a Silicon Valley worldview where advances in software and AI are the keys to figuring out almost any issue" (New York Times 2020). Even experts within the field of AI acknowledge that the technology runs the risk of inhibiting progress in key areas of societal development (such as employment, environment and democracy), leading to growing calls for robust regulatory oversight and ethical standards (Vinuesa *et al.* 2020).

With these thoughts in mind, this report considers how such issues and debates are playing out within general public opinion. Traditionally, AI development has been led predominantly by expert voices in industry, research and policy. However, there is now growing acknowledgement of the need for AI and data science to pay more attention to public understandings, sentiments and options. As Zhang and Dafoe (2020, p.187) contend:

> ***"While technologists and policymakers have begun to discuss the societal implications of machine learning and AI more frequently, public opinion has not shaped much of these conversations. Given AI's broad impact … the public, particularly those underrepresented in the tech industry, should have a role in shaping the technology".*** Zhang and Dafoe (2020, p.187)

As such, it is important for AI and data science communities in Australia to pay attention to public opinions and sentiments toward AI – both as an influence on future societal implementations of the technology, and to identify areas where public understandings of AI might be developed. Public acceptance will be a key factor in determining the success of the implementation of AI in society. Indeed, the Australian public plays a number of pivotal roles – most obviously as active 'end-users' of some AI applications, and as unknowing subjects of many others. In addition, the public also constitutes the electorates that shape government policymaking, as well as the shareholders and customers of the corporations that produce the technology. All told, AI innovation over the next few decades will be shaped increasingly by public opinions and understandings.

The significance of public understandings and opinions has been highlighted in recent controversy and push-back against AI-based technologies such as facial recognition and automated systems such as the 'Robo Debt' online compliance intervention system. These recent experiences suggest that when it comes to the development of technologies with wide-ranging social consequences, public understandings and opinions need to be anticipated and considered in advance. Otherwise policy makers and the public more generally find themselves in a reactive mode – attempting to address, after the fact, the unanticipated consequences of the rapid deployment of new technologies. In response, policymakers, developers and researchers have an obligation to ensure that the technology is implemented in accordance with civic commitments and the public interest.

Against this background, the present report joins a burgeoning literature looking at public understandings, perceptions and attitudes toward AI (e.g. Fast and Horvitz 2017, Ipsos/MORI 2016, Royal Society 2017, Cave *et al.* 2019). Previous studies point to a range of responses – from enthusiasms over speculative applications, to fears over job losses, and lack of trust in key actors. Often, these studies highlight what Aphra Kerr *et al.* (2020, n.p) describe as the "disconnect between societal expectations and practice around the design and deployment of AI". One of the main aims of the present report, therefore, is to explore the extent to which issues from these recent studies elsewhere in the world correspond to the Australian context.

Of particular significance to our present study is a large-scale survey of 2000 US adults conducted during June 2018 detailing attitudes toward AI governance issues (Zhang and Dafoe 2019). This survey went beyond broad-level concerns and focused on a number of areas relating to ethical issues surrounding AI and society, views over the regulation and governance of these technologies, and expectations of future capacity of AI technology. This survey highlighted a range of issues and trends – some of which are pursued in our own investigation of the Australian context. Notably, Zhang and Dafoe (2019) identified substantial variations in support/opposition to AI development in terms of income, educational background, gender and technology background.

So, in what ways are these issues evident in the ongoing integration of AI technology into Australian society? Drawing on our own nationally-representative survey of just over 2000 adults across Australia, this report provides insights into the following areas of questioning:

- Public understandings of AI (i.e. what the public consider to constitute 'artificial intelligence', and how knowledgeable people consider themselves to be);
- Public support for the development of AI in different societal domains (e.g. health, workplace, education, transportation, crime & justice, communities);
- Public opinions regarding AI governance challenges in the next 10 years (e.g. attitudes relating to fairness, accountability & transparency issues in different societal domains, surveillance, social effects);
- Levels of trust in different public and private actors to develop and manage AI in the best interests of the public;
- Future expectations for AI applications – notably speculative forms of 'high-level machine intelligence' where technology would have the capacity to learn or understand tasks better than the average human.

As this report now goes on to detail, while there is some notable patterning in Australian public responses to the prospect of the societal implementation of AI, there are also surprising similarities and points of consensus between groups that might be assumed to hold conflicting views. All told, these findings have significant implications for the development of AI-related policymaking and practice over the next 10 years or so.

We conclude that these findings demonstrate the need for AI developers, researchers, industry actors, policymakers, and regulators in Australia to pay closer attention to public opinion. Above all, the roll-out of AI-driven products and processes needs to become a topic of on-going consultation, dialogue and consensus-building amongst the Australian public – especially if ideals of 'AI for social good' are to be realised.

> **This report presents an analysis of data generated by a nationally-representative survey of 2019 Australian adults aged 18 years and over**

Methodology

This report presents an analysis of data generated by a nationally-representative survey of Australian adults aged 18 years and over (n=2019). A 73-item questionnaire was developed to gauge public opinions on the areas of questioning outlined in the previous section. The survey comprised seven main sections (sections marked with an asterisk include items adapted from Zhang and Dafoe 2019):

- Respondent background and demographics;
- Support for the development of AI*;
- Opinions regarding AI for social good – i.e. the application of AI to social, humanitarian and environmental challenges;
- Opinions regarding societal challenges raised by AI – e.g. issues of privacy, fairness, equality and other human rights*;
- Confidence in organizations to develop and manage AI in the best interests of the public*;
- Expectations for the future development of AI*;
- Hopes and futures regarding AI and society.

The study was conducted by *WhereTo* Research using participants from the Online Research Unit (ORU) online panel cohort. The survey was administered to members of the ORU panel, and responses collected between April 1st and April 24th 2020. This resulted in a sample of n=2019 adult residents eligible to vote Australia.

The final sample (see Table 1) was broadly representative of Australian population figures in terms of gender, region and socio-economic status (see Appendix for further details). Data were then weighted by age group to correct for online panel deviation from the Australian population.

Table 1. Survey sample (n=2019) and weighted sample by main demographic characteristics.

	n	Percentage	Weighted n	Weighted percentage
Gender				
Female	1004	49.9	1022.1	50.8
Male	1008	50.1	990.7	49.2
Age range (years)				
18–24	287	14.2	245.1	12.1
25–29	211	10.5	196.3	9.7
30–34	267	13.2	194.6	9.6
35–39	234	11.6	175.9	8.7
40–44	176	8.7	169.8	8.4
45–49	239	11.8	174.5	8.6
50–54	61	3	162.3	8
55–59	82	4.1	159.5	7.9
60–64	98	4.9	141.1	7
65+	364	18	400.2	19.8
State				
ACT	70	3.5	64.1	3.2
NSW	640	31.7	610.9	30.3
NT	32	1.6	30.6	1.5
QLD	401	19.9	421.9	20.9
SA	141	7	148.4	7.4
TAS	42	2.1	52.8	2.6
VIC	521	25.8	523.4	25.9
WA	172	8.5	167.1	8.3

	n	Percentage	Weighted n	Weighted percentage
Urbanity (postcode data)				
Capital city	1324	65.6	1281.5	63.5
Regional	695	34.4	737.9	36.5
Highest education level completed				
Year 12 or below	737	36.7	728	36.2
TAFE/Technical	189	9.4	196.5	9.8
Certificate or Diploma	409	20.4	423.6	21.1
University degree or higher	672	33.5	660.5	32.9
Annual household income before tax				
Up to $20,000	138	7.7	138.4	7.7
$20,000 to $39,999	319	17.7	343.4	19.1
$40,000 to $59,999	321	17.8	327.8	18.2
$60,000 to $79,999	249	13.8	250.6	13.9
$80,000 to $99,999	210	11.7	198.5	11
$100,000 to $119,999	180	10	180.3	10
$120,000 to $149,999	167	9.3	162.9	9
$150,000 or more	218	12.1	200.7	11.1
SES (postcode)				
Low	621	30.8	613.7	30.4
Medium	810	40.1	841.3	41.7
High	588	29.1	564.4	27.9
Language background other than English				
No LBOTE	1563	77.9	1617.3	80.5
LBOTE	443	22.1	390.6	19.5
Identify as Aboriginal or Torres Strait Islander				
Do not identify as Aboriginal or Torres Strait Islander	1925	96.4	1934.5	96.9
Identify as Aboriginal or Torres Strait Islander	72	3.6	62.2	3.1
Identify as having a disability				
Do not identify as having a disability	1786	89.3	1768.1	88.3
Identify as having a disability	213	10.7	233.5	11.7
Vote preference in last Federal elections				
Labor	558	34.9	550.8	33.6
Liberal/National Alliance	694	43.3	738.2	45
The Greens	162	10.1	157.1	9.6
Other	187	11.7	193.5	11.8
Computer science/programming background*				
No relevant background	1476	73.1	1517.5	75.1
Relevant background	543	26.9	501.9	24.9

* Respondents were asked if they had programming experience and/or had taken at least one university-level course in computer science/engineering.

> **While 89.4% of people are aware of the AI, the majority of Australians consider themselves to have little knowledge of artificial intelligence**

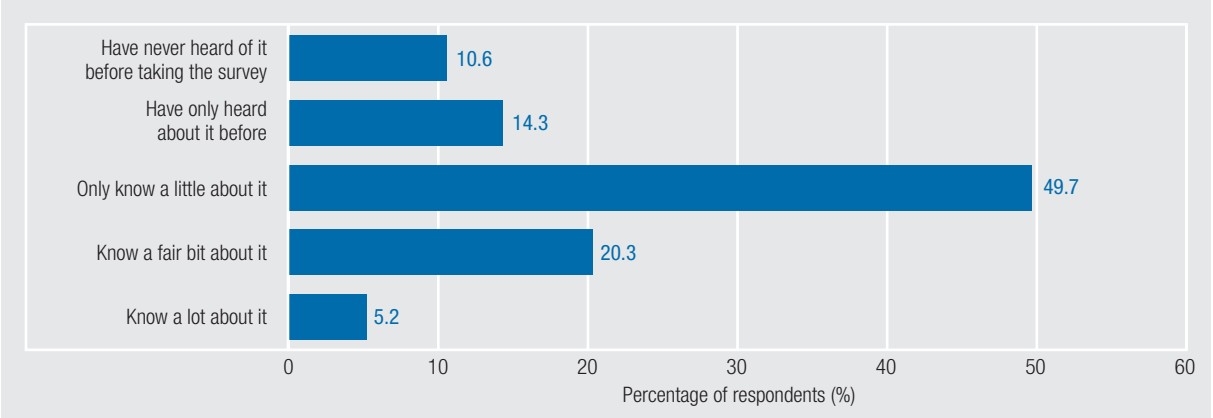
Results

1. Public familiarity and understandings of 'artificial intelligence'

1.1 How much do people know about artificial intelligence?

Prior to the survey giving any working definitions and examples of artificial intelligence technology, respondents were first asked to gauge their own familiarity and knowledge of the topic. Nearly nine-in-ten respondents (89.4%) claimed to have heard of AI prior to taking the survey. However, only 5.2% of respondents claimed to 'know a lot' and 20.3% to 'know a fair bit'. Generally, the majority of respondents claimed to only have passing knowledge and familiarity with the topic of AI (see Figure 1).

Figure 1. Self-reported levels of familiarity with AI technology.



NB. Survey item asked: *"How familiar are you with AI technology?"*

As shown in Table 2, these levels of familiarity and knowledge were patterned by a number of key characteristics. For example, while just over one-quarter of all respondents (25.5%) describes themselves as knowing 'a lot' or 'a fair bit' about AI, these levels of knowledge were most prominent in the following groups:

- Experience of programming and/or computer science/engineering at university (50.6%)
- Age range 18 to 24 years old (50.8%)
- Speak a language other than English at home (43.3%)

Knowing 'a lot' or 'a fair bit' about AI was also more prevalent in the following other demographic groups:

- High SES (33.6%)
- Educated to the level of university degree or higher (32.5%)
- Male (32.5%)
- Living in a capital city (30%)

Conversely, are the group of respondents describing themselves as either having never heard of AI before, or only heard the term without knowing anything about it. This group also constitutes around one-quarter of all respondents (24.9%). These low levels of familiarity and knowledge most prominent in the following groups:

- Aged 65 years and above (37.3%)
- Educated to level of TAFE/technical (37%)
- Annual household income $20,000 to $39,999 (32.1%)
- Females (30.5%)

Table 2: Self-reported levels of familiarity with AI technology by demographic groups.

	Know a lot about it	Know a fair bit about it	Only know a little about it	Have only heard about it before	Have never heard of it before taking this survey
Gender* **					
Female	3.5	15.1	50.9	17.7	12.8
Male	6.9	25.6	48.6	10.8	8
Age range (years)* **					
18–24	9.6	41.2	40.4	3.7	5.1
25–29	12.1	24.4	40.4	8.9	14.2
30–34	11.6	24.8	41.6	10.1	11.9
35–39	5.3	23.3	49.4	9.9	12.1
40–44	5.2	18.6	57	10.2	8.9
45–49	4	22.1	50.3	16.2	7.5
50–54	1.7	16.9	55	16.6	9.8
55–59	0	11.5	64.1	16.8	7.7
60–64	0	14.6	50	25	10.4
65+	1.7	8.5	52.5	22.8	14.5
Urbanity (postcode data)* **					
Capital city	6.6	23.4	47.2	12.3	10.5
Regional	2.7	14.8	53.9	17.9	10.7
Highest education level completed* **					
Year 12 or below	4.3	20.7	45.5	15.2	14.4
TAFE/Technical	4.6	12.2	46.2	25.9	11.1
Certificate or Diploma	2.9	16.8	53.5	16.1	10.6
University degree or higher	7.8	24.7	52.7	8.9	5.9
Annual household income before tax* **					
Up to $20,000	8.5	17.8	45.3	11.1	17.3
$20,000 to $39,999	3.9	16.9	47.1	18.9	13.2
$40,000 to $59,999	4.6	16	52	14.1	13.3
$60,000 to $79,999	6.6	21.8	41.7	19	10.9
$80,000 to $99,999	6.4	21.5	47.7	16.6	7.8
$100,000 to $119,999	4.7	31	47.2	8.1	9
$120,000 to $149,999	5.6	21.1	55.8	13.4	4.2
$150,000 or more	5.8	24.1	57.4	7.6	5.1
SES (postcode data)* **					
Low	3.2	18.4	49.7	16.4	12.4
Medium	5.1	17.7	50.2	15.6	11.4
High	7.4	26.2	48.9	10.1	7.4
Language background other than English* **					
No LBOTE	3.4	17.7	52.2	16.2	10.4
LBOTE	12.4	30.9	39.2	6.8	10.7
Identify as Aboriginal or Torres Strait Islander*					
Do not identify as Aboriginal or Torres Strait Islander	4.9	20.5	49.9	14.4	10.3
Identify as Aboriginal or Torres Strait Islander	11.7	16.4	41.2	13.7	17

	Know a lot about it	Know a fair bit about it	Only know a little about it	Have only heard about it before	Have never heard of it before taking this survey
Identify as having a disability					
Do not identify as having a disability	4.8	20.3	49.7	14.6	10.7
Identify as having a disability	7.9	20.3	50.3	12	9.4
Vote preference in last Federal elections**					
Labor	5.7	20.7	52.1	12.1	9.4
Liberal/National Alliance	4.4	18.1	50.8	15	11.6
The Greens	6.5	25.8	54.3	6.5	6.8
Other	5.3	23.2	49.6	18.2	3.8
Computer science/programming background***					
No relevant background	1.5	15.5	53.5	18.1	11.5
Relevant background	16.2	34.4	38.4	3	8

***p-value<0.001, **p-value<0.05, *p-value<0.1 for a chi-square test of independence between familiarity and the demographic characteristic.
NB. Survey item asked: *"How familiar are you with AI technology?"*

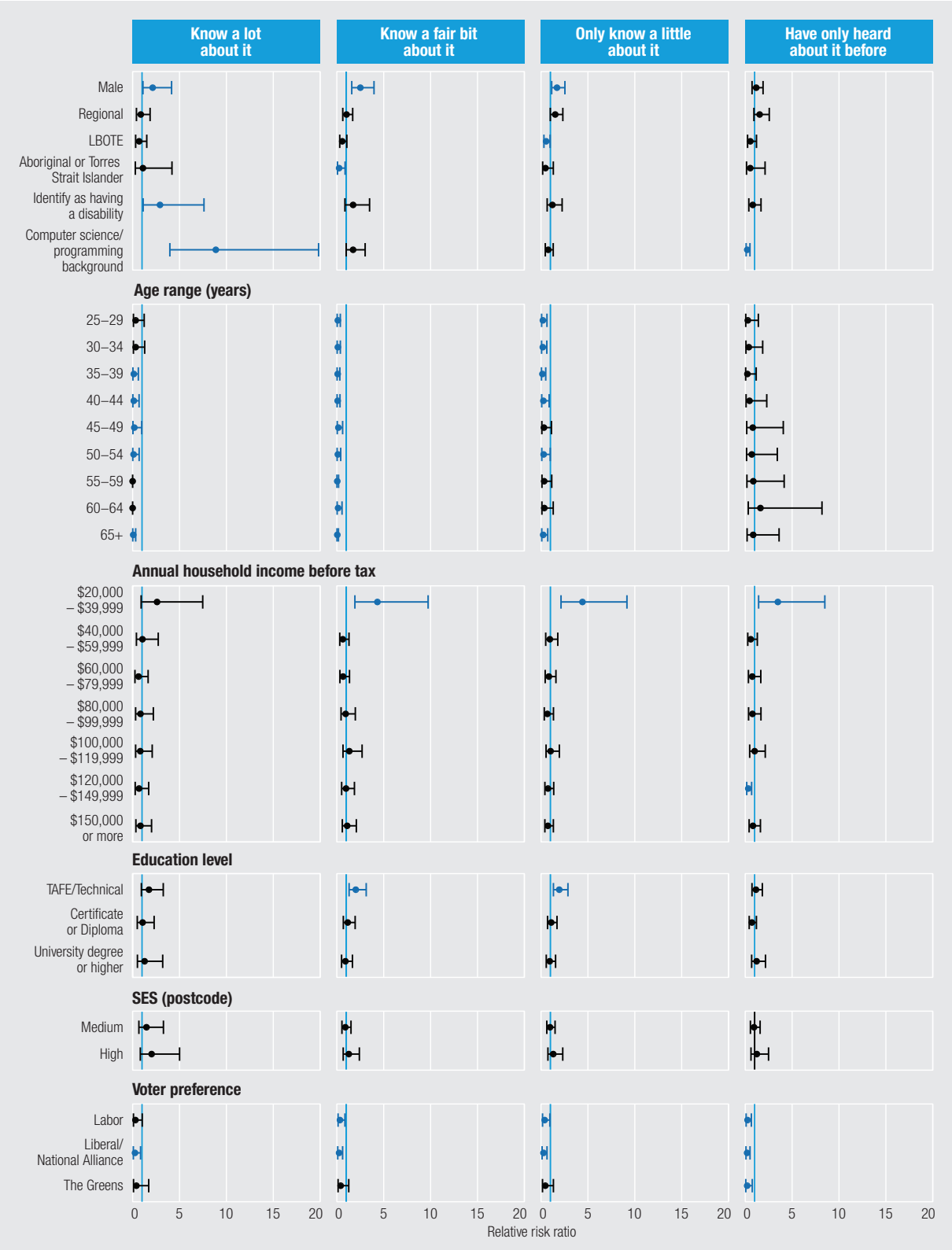
When background characteristics are taken together as a whole, these levels of familiarity and knowledge about AI are found to be patterned primarily by five specific variables (see Figure 2). Comparing the likelihood of people claiming to 'know a lot' about AI than having 'never heard about it' highlights the significant factors[1]:

- **Computer science/programming background:** Those who have a computer science/programming background are 8 times more likely to describe themselves as 'knowing a lot' rather than 'never heard of it', than those without a background in computer science/programming.
- **Disability:** People who identify as having a disability are almost three times more likely to describe themselves as 'knowing a lot' rather than 'never heard of it', than those who do not identify as having a disability.
- **Gender:** Males are twice as likely to describe themselves as 'knowing a lot' rather than 'never heard of it' than females.
- **Age**[2]**:** Respondents over 35 are almost twice less likely to describe themselves as 'knowing a lot' rather than 'never heard of it' than those in the 18–24 years age range.

> **"25.5% of respondents described themselves as knowing 'a fair bit' or 'a lot' about AI. These are most likely to be people with computer science or programming expertise, aged between 18–24 years, and/or from households where a language other than English is spoken"**

1. An association with household income is also apparent but relates primarily to differences between those earning less than $20,000 and those earning between $20,000 and $39,999. Households earning less than $20,000 are likely to include large numbers of retired adults and students. Higher levels of income are not linked to higher levels of familiarity with AI.
2. One caveat of this finding is that there is little variability within cells. For example, few respondents in the 55–59 and 60–64 age groups selected the option of 'know a lot about it'. This means that for that specific age-group/familiarity level combination our sample size is very small, increasing the uncertainty associated with our estimations, which translates into very wide confidence intervals.

Figure 2. Relative risk ratio and 95% confidence intervals of selecting each level of familiarity over 'Have never heard of AI before taking this survey' by respondents' demographic characteristics.



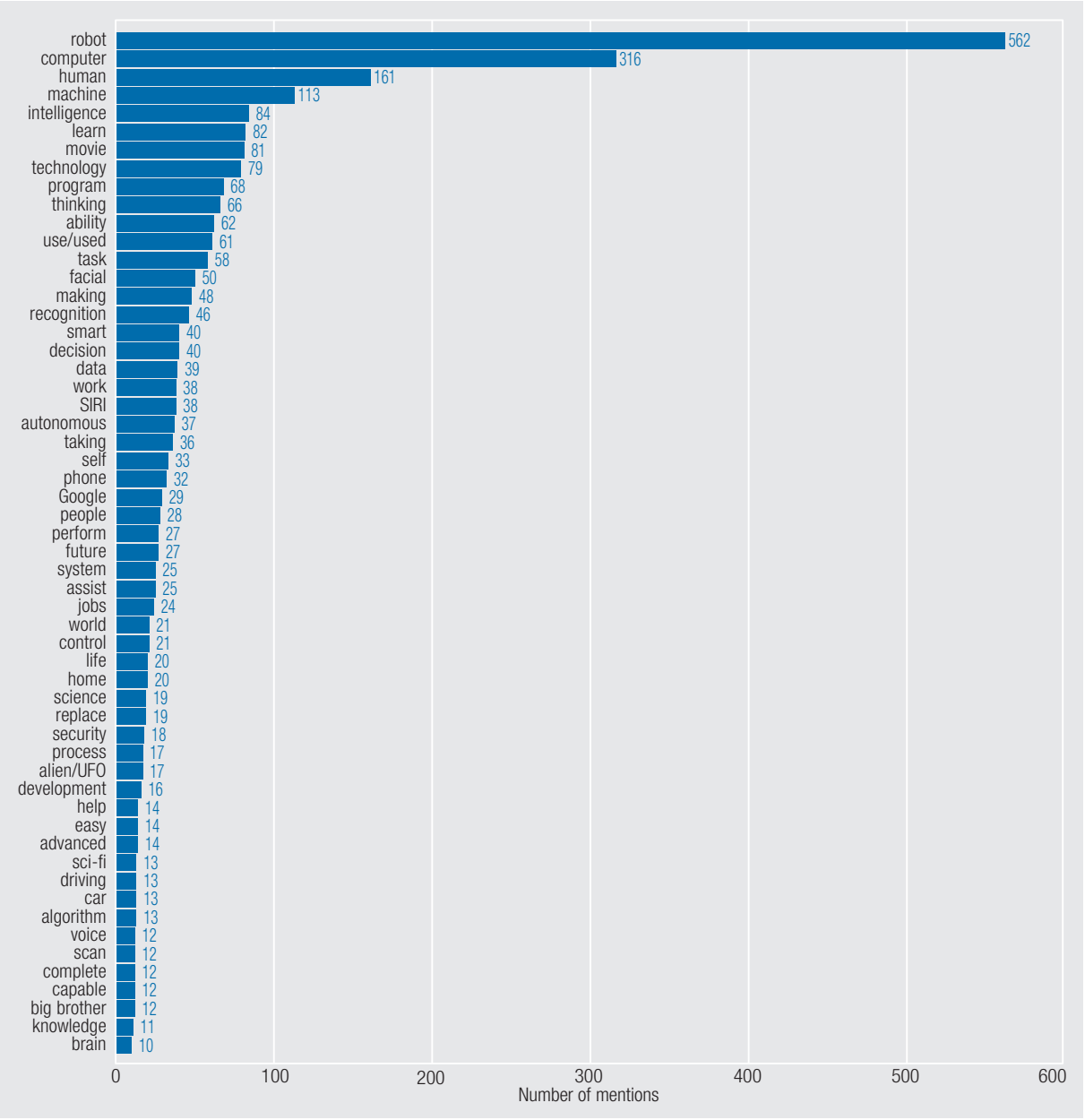
NB. Figure shows multinomial logistic estimation results.

1.2 What do people understand by 'AI'?

In order to substantiate these self-reported levels of knowledge and familiarity, this opening section of the survey also explored the ways in which respondents conceptualised AI. Respondents were asked to respond to a straightforward open-ended question: *'What's the first thing that comes to mind when you think about Artificial Intelligence?'.*

This prompt elicited 7547 words from 1763 respondents. As can be seen in Figure 3, nearly one-third of respondents offered responses mentioning 'Robots' (31.9%), followed by 'Computers' (17.9%) and 'Humans' (9.1%).

Figure 3. Most frequently used terms associated with the phrase 'Artificial Intelligence'.



NB. Data are number of respondents mentioning each term in their response to the question: *"What's the first thing that comes to mind when you think about Artificial Intelligence?"* Responses were provided by 1763 respondents.

The most popular ways in which AI is described reflect ideas of robots 'taking on work' and/or 'taking over the world'

A better sense of the meanings attached to these responses is gained by considering how different terms were associated with each other within the same response. Table 3 shows the eight most cited terms, alongside other associated terms. Terms were judged to be associated when at least 25% of the occasions that a term was used was in relation to one of the eight most cited terms. This shows, for example, that 36 out of 66 mentions of 'Thinking' were made in responses that mentioned 'Computer'. Similarly, 21 out of 36 mentions of 'Taking' were made in responses that mentioned 'Robots'. Mapping these associations therefore helps us develop a more detailed sense of the ways in which the main terms associated with AI were being used.

Table 3. Eight most-cited terms – with other associated terms that have 25% or more instances in same response.

Main term	Associated terms (loading 25% or more)
Robot (562)	Complete (4/12); Driving (4/13); Jobs (10/24); Replace (6/19); Taking (21/36); Work (16/38); World (9/21)
Computer (316)	Ability (26/62); Capable (5/12); Complete (3/12); Data (14/39); Decision (22/40); Develop (8/16); Human (56/161); Learn (35/82); Making (21/48); Perform (20/27); Program (31/68); Science (8/19); Self (10/33); Smart (12/40); Task (30/58); Thinking (36/66); Work (10/38)
Human (161)	Capable (5/12); Complete (8/12); Control (6/21); Develop (9/16); Intelligence (35/84); Machine (31/113); Making (12/48); Perform (20/27); Process (5/17); Replace (11/19); Science (7/19); Task (30/58); Thinking (17/66)
Machine (113)	Capable (4/12); Intelligence (22/84); Learn (25/82); Science (5/19)
Intelligence (84)	Capable (4/12); Develop (7/16); Perform (12/27); Science (7/19)
Learn (82)	Capable (4/12); Knowledge (5/11); Process (5/17); Program (18/68)
Movie (81)	Sci-Fi (6/13)
Technology (79)	–

For these eight clusters of responses, a number of observations can be made about the understandings of AI conveyed by our respondents. First, the numerous responses relating to 'Robots' conveyed relatively broad understandings of artificial intelligence. While some of these responses mentioned high-profile examples of robotics (such as 'Sophia' and 'Boston Dynamics'), most tended to describe non-specific automated technology 'taking over' various aspects of everyday life – specifically taking over 'jobs', 'work' and even the possibility of 'taking over the world'. Responses ranged from a sense of AI as technology controlled by humans, through to AI as agentic technology that is capable of superseding humans. For example:

> ***"Robots doing tasks that are menial to relieve people of boring jobs"*** *[Respondent #527, M, 65+ years, NSW]*
>
> ***"Robotic implements which are programmed to respond to certain commands and provide answers as long as the required information has been put into the machine"*** *[#537, F, 65+, NSW]*
>
> ***"AI Robots taking over the world as sentient beings who don't feel physical pain and have the ability to learn and process information incredibly fast"*** *[#2571, M, 18–24, QLD]*
>
> ***"Scary robots taking over the world"*** *[#4257, M, 40–44, NSW]*

Responses involving 'Computers' rather than robots tended to be more nuanced. Here, descriptions often related to computers being 'built', 'programmed', 'trained' or 'used'. Some mention was made of specific existing technologies such as SIRI, Alexa and Google Search. AI technology was generally described as only 'appearing' to be intelligent, while actually involved in relatively low-level tasks – as one respondent put it, *"computers doing the grunt work for everything"* [#727, M, 45–49, VIC]. For example:

> ***"A form of computer programming that enables advanced tasks to be undertaken in shorter time"*** *[#100, M, 65+, QLD]*
>
> ***"Using the power of computers to predict, model and provide solutions to future problems"*** *[#1731, M, 45–49, ACT]*
>
> ***"Computer learning to work through the logic of tasks"*** *[#314, M, 55–59, ACT]*

Similarly, responses relating to '**Human**', '**Machine**' and '**Intelligence**' all tended to convey a sense of machines being given some form of human-level capacity – for example, *"It's a computer science that tries to put human intelligence into machines"* [#307, F, 60–64, VIC]. As with responses that made connections between AI and computers, these descriptions tended to acknowledge the bounded and 'synthetic' nature of these capabilities. Nevertheless, emphasis here was often placed on technology capable of reaching human-level intelligence. For example:

> ***"Programming capable of operating somewhat like a human, i.e. mimicking human thought/ decision making, potentially capable of learning"*** *[#951, F, 25–29, WA]*

That said, responses relating to these themes did occasionally stray into more extreme scenarios:

> ***"Complete abrogation of human intelligence to computerised algorithms, causing a lot of chaos in society"*** *[#601, F, 60–64, VIC]*

Finally, were three smaller clusters of responses that each appear to convey distinctly different sets of understandings. These included responses relating to the term '**Learn**' – often conveying a sense of 'autonomous learning' and technology that is programmed to be capable of 'self-learning'. For instance, AI was described as:

> ***"Technology that is programmed to be able to learn and develop by itself"*** *[#1854, M, 18–24, QLD]*
>
> ***"Any system that learns on its own"*** *[#233, F, 45–49, VIC]*
>
> ***"AI makes me think about how computers can develop their own intellect by learning like us humans where they become autonomous and independent to solve a particular function"*** *[#1938, M, 18–24, NSW]*

A small cluster of responses (4.6%) related to fictional portrayals of AI from popular culture – specifically representations of AI in '**Movies**'. Some of these associations were non-specific, for example: *"Movies where the AI has taken control"* [#2867, M, 18–24, NSW] and *"Out of control robots in science fiction movies"* [#612, M, 65+, ACT]. Specifically identified movies included multiple mentions of 'Terminator', 'the movie AI' and 'I, Robot'. Other associations were made with 'HAL 9000' in '2001: A Space Odyssey', 'Blade Runner', 'The Matrix' and even 'Back to the Future'. These responses often simply mentioned the movie, and occasionally a reference to a dystopian aspect of its storyline.

Finally, another small number of responses related to more general notions of AI as '**Technology**' that makes various tasks and aspects of life 'easier'. These responses ranged from vague illusions to *"very clever technology"* [#3114, M, 18–24, VIC] through to specific mention of high-profile (but not necessarily AI-related) technology genres such as "smart technology", "block chain technology" and "nanotechnology".

> **"Many respondents have more sophisticated understandings, making immediate associations with computers being programmed to perform tasks, learning from data, and displaying human-like thinking"**

2. Public support for AI development

At this point in the survey, respondents were provided with the following definition (and five randomly chosen examples of AI technology from the list in Appendix (Figure A1) to frame their subsequent responses:

> ***"Artificial Intelligence (AI) refers to computer systems that perform tasks, make decisions and predictions that usually require human intelligence. AI can perform these tasks or make these decisions without explicit human instructions".***

Having read this description, respondents were asked: *"In general, how much do you support or oppose the development of AI?"*. As shown in Table 4, nearly two-thirds of respondents (62.4%) indicated support for the development of AI, with only 10.5% expressing any opposition.

Given that a large majority of respondents professed to having little familiarity or knowledge of AI prior to taking the survey, we were interested in how these levels of support/opposition might alter in light of having taken the whole survey. In other words, once having spent 20 minutes or so considering issues and arguments about the rights and wrongs of the technology, how might people's overall opinions towards AI alter?

The survey therefore concluded by asking respondents: "Having taken this survey (and thought through your previous answers) how much do you now support or oppose the development of AI?". As also shown in Table 4, the overall levels of initial support/opposition did not appear to have altered notably after engaging with the survey topics – with 63.5% of respondents expressing support for the development of AI after taking the survey.

Table 4. Respondents' support or opposition to the development of AI – comparison of responses at the beginning, and the end of the survey.

	Oppose	Neutral	Support	Don't know
Beginning of survey	10.5	23	62.4	4.1
End of survey	12	18.7	63.5	5.8

NB. Data are percentage of respondents expressing an opinion

That said, while the overall levels of support/opposition remained broadly similar, these were not necessarily the same respondents. Indeed, we found 45.3% of respondents to have altered their opinion (on the six-point Likert scale from 'strongly support' to 'strongly oppose', alongside 'don't know') between the beginning and the end of the survey. Specifically, 20.5% of respondents moved their opinion from being initially more opposed to then being less opposed (i.e. making a positive shift in opinion). Conversely, 24.8% of respondents moved their opinion from being initially more supportive to than being less supportive (i.e. making a negative shift in opinion).

More significant shifts in opinion are detailed in Table 5. This shows, for example, that 36.7% of respondents who initially stated an initial 'neutral' opinion subsequently moved to a 'support' position. Similarly, 25% of respondents who initially stated an 'oppose' opinion subsequently shifted to a 'support' opinion. In turn 17.1% of respondents with an initial 'support' position then moved to either a 'neutral', 'oppose' or 'don't know'. These flows are also represented in Figure 4.

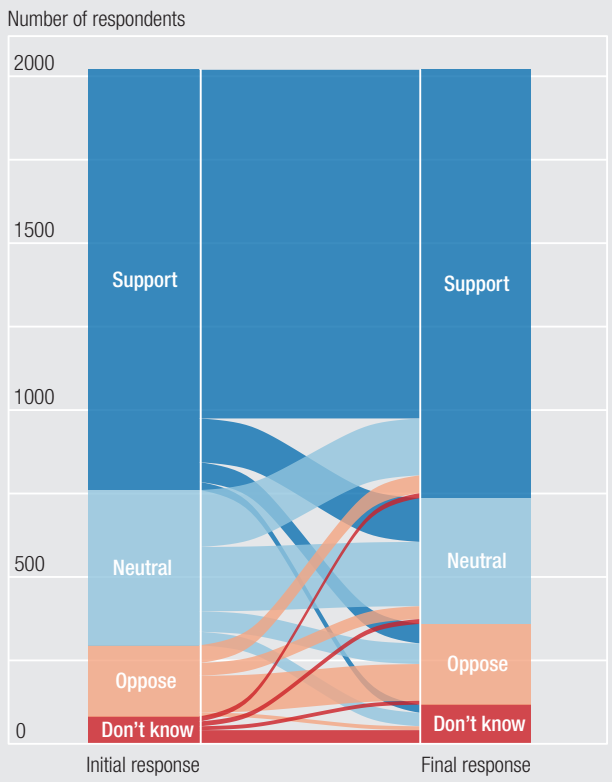
> **"63.5% of respondents stated that they 'somewhat support' or 'strongly support' the development of AI"**

Table 5. Respondents' support or opposition to the development of AI.

	Final response			
Initial response	**Support**	**Neutral**	**Oppose**	**Don't know**
Support	82.9	10.4	4.7	2.0
Neutral	36.7	41.4	13.2	8.7
Oppose	25.0	18.0	51.6	5.4
Don't know	18.3	17.7	14.1	49.8

NB. Data are proportion of respondents according to their initial opinion.

Figure 4. Changes in support or opposition to the development of AI between the beginning and end of the survey.



> **43% of the respondents who initially considered themselves 'opposed' to the development of AI to shift subsequently to either a 'neutral' or 'supportive' stance once having engaged with all the survey questions.**

As such, it is interesting to consider which respondents had their opinion altered by the experience of engaging with the survey questions. However, multivariate analysis suggests that there is no clear pattern in terms of demographic characteristics. In other words, there were no clear trends in terms of respondents' age, gender, socio-economic status and so on.

Indeed, estimation of a multinomial logit model (detailed in Appendix) shows that respondents who did not change their mind are very similar in all their demographic characteristics to those respondents who changed their mind from their initial answer to opposing the use of AI at the end of the survey. The only specific trend that was apparent were signs that respondents with experience of programming and/or computer science/ engineering at university were *less* likely to change their mind to a 'Don't Know' answer at the end of the survey. Otherwise, no clear trends were evident.

3. Public opinions on potential applications of AI for social good

Table 6. Descriptions of potential applications of AI for social good used in survey. Some descriptions are adapted from McKinsey (2018) discussion paper: 'Applying artificial intelligence for social good'.

Application	Description
Crisis response	AI can be used to address specific crisis-related challenges, such as responses to natural and human-made disasters in search and rescue missions, as well as the outbreak of disease. Examples include using AI on satellite data to map and predict the progression of wildfires, earthquakes, hurricanes, floods and other natural disasters. Drones with AI capabilities can also be used to find missing persons in wilderness areas.
Economic empowerment	With an emphasis on currently vulnerable populations, AI can be used to open access to economic resources and opportunities, including jobs, the development of skills, and market information. For example, AI can be used to detect plant damage early through low-altitude sensors, including smartphones and drones, to improve yields for small farms.
Educational challenges	AI uses in education include maximizing student achievement and improving teachers' productivity. For example, adaptive-learning technology can base recommended content to students on past success and engagement with the material.
Environmental challenges	AI can be used to sustain biodiversity and combat the depletion of natural resources, pollution, and climate change. For example, AI tools are used in conservation efforts across the world – detecting illegal logging in vulnerable forest areas, or over-fishing in restricted areas.
Equality and inclusion	AI can be used to address challenges to equality, inclusion, and self-determination (such as reducing or eliminating bias based on race, sexual orientation, religion, citizenship, and disabilities). For example, AI has been used to automate the recognition of emotions and to provide social cues to help individuals along the autism spectrum interact in social environments.
Health	AI can be used to address health challenges, including early-stage diagnosis and modelling the outbreak and transmission of infectious disease. Applications include a disease-detection AI system – using the visual diagnosis of natural images, such as images of skin lesions, to determine if they are cancerous. AI-enabled wearable devices can already detect people with potential early signs of diabetes.
Medicine	AI can be used to develop new drugs and medicines, thanks to its ability to scan millions of chemical compounds at high speed in order to detect antibiotic and other medicinal properties.
Culture	AI can be used to generate culture for popular consumption. It has been used to create art, including painting poetry, novels, music, and screenplays, potentially developing new cultural forms and aesthetic sensibilities.
Information verification and validation	AI can be used to facilitate the provision, validation, and recommendation of helpful, valuable, and reliable information to all. It focuses on filtering or counteracting misleading and distorted content, including false and polarizing information disseminated through the relatively new channels of the internet and social media. For example, AI has been used to actively present opposing views to ideologically isolated pockets in social media.
Infrastructure management	AI can be used to address infrastructure challenges in the areas of energy, water and waste management, transportation, real estate, and urban planning. For example, traffic-light networks can be optimized using real-time traffic camera data and sensors to maximize vehicle throughput. AI is being used to schedule predictive maintenance of public transportation systems, such as trains and public infrastructure (including bridges), to identify potentially malfunctioning components.
Public and social-sector management	AI can be used to support efficiency and the effective management of public- and social-sector entities, including strong institutions, transparency, and financial management, are included in this domain. For example, AI is being used to identify tax fraud using alternative data such as browsing data, retail data, or payments history.
Security and justice	AI can be used to prevent crime and other physical dangers, as well as tracking criminals and mitigating bias in police forces. These applications of AI focus on security, policing, and criminal-justice issues. For example, AI has been used to predict likely areas of criminal activity, and to identify traffic violations.

> **"Particularly strong support was expressed for the use of AI to address challenges in areas such as health and medicine, alongside environmental challenges and crisis response"**

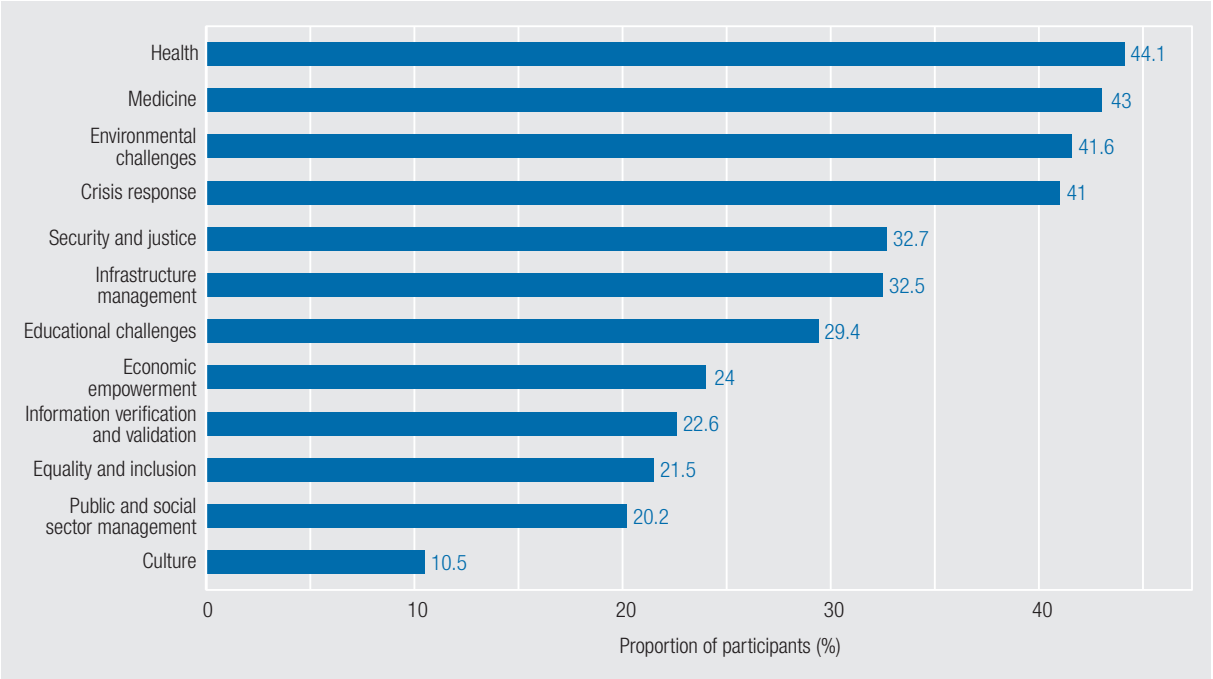
3.1 Public support for using AI to address social, humanitarian and environmental challenges.

The next section of the survey explored respondents' opinions toward the application of AI for 'social good' – described in the survey as *"the application of artificial intelligence to social, humanitarian and environmental challenges"*. Each respondent was asked to rate their support/opposition to the use of AI in 5 different areas of possible AI implementation (chosen randomly from an overall list of 12 areas outlined in Table 6).

Levels of support for the use of AI across all these 12 areas was consistently high. It is perhaps most useful, then, to examine the relative proportions of respondents expressing 'strong support'. As shown in Figure 5 and Table 7, the highest proportions of 'strong support' for AI use were in the areas of health and medicine. The other two highest strongly supported areas of development each relate to recent areas of public concern in Australia – 'environmental challenges' (in the aftermath of the 2019/2020 summer bushfires), and 'crisis response' (in the midst of the COVID-19 pandemic). Conversely, a third high-profile topic at the time of taking the survey ('information management and verification' – relating to issues of disinformation and fake news) received considerably lower levels of strong support.

In contrast to support for health and medicine, there were notably fewer respondents expressing strong support for the use of AI for cultural purposes, public sector management, and equity and inclusion issues. This might well relate to lower levels of public interest in the cultural, public and social sectors, or perhaps a perception of these areas as more human-led. Lower levels of 'strong' support might also be compounded by these areas being less familiar when compared to health and environment.

Figure 5. Self-reported support for the use of AI in different areas of society.



NB. Data in bar chart are percentage of respondents indicating that they 'Strongly Support'. Survey asked: *"How much do you support or oppose the use of AI in…"*

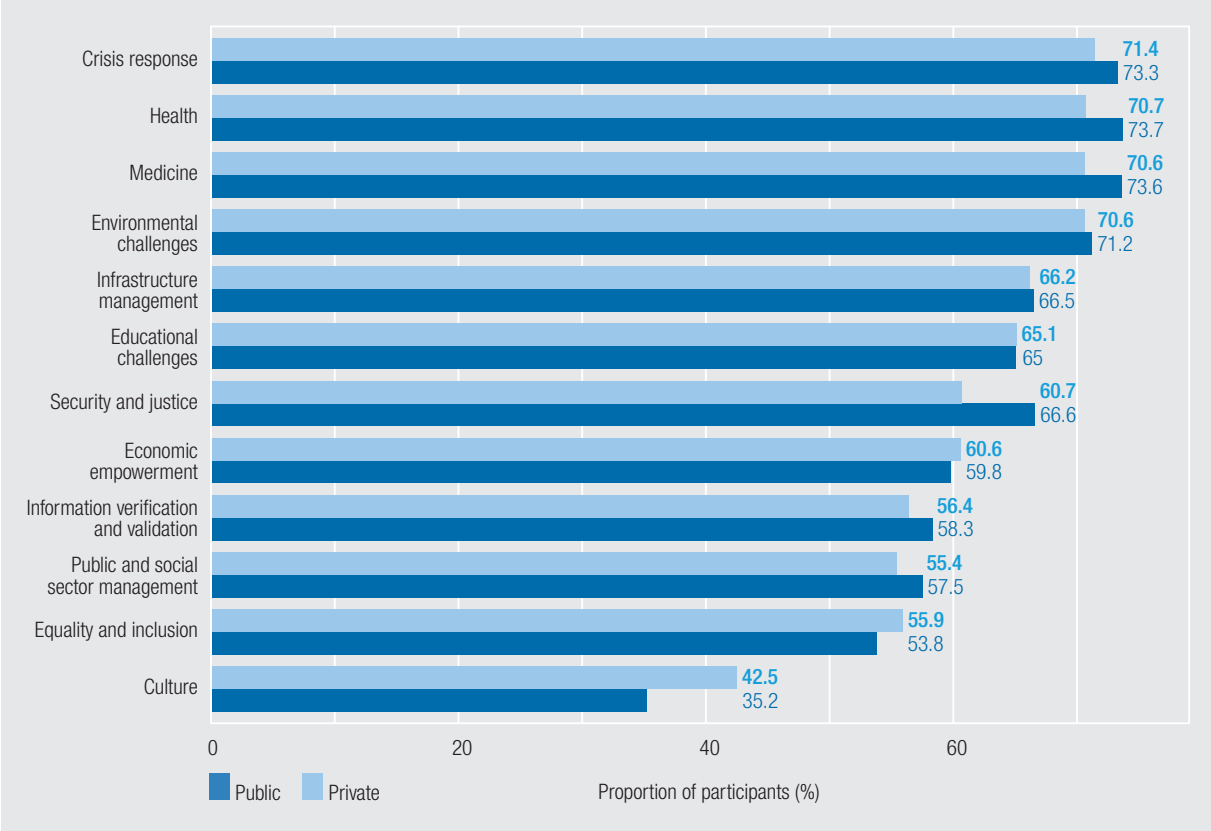
Table 7. Percentage of respondents according to their answer to: *'How much do you support or oppose the use of AI in…'*

	Strongly oppose	Somewhat oppose	Neither support nor oppose	Somewhat support	Strongly support	Don't know
Health	2.6	3.5	13.5	33.3	44.1	2.9
Medicine	2.8	5.1	12.7	33.2	43	3.2
Environmental challenges	2.9	3.6	12.1	35.9	41.6	3.8
Crisis response	3.1	3.4	12	37.3	41	3.2
Security and justice	5.9	6.7	15	35.6	32.7	4.2
Infrastructure management	3.1	4.2	15.7	38.9	32.5	5.6
Educational challenges	2.7	5.2	17	41.9	29.4	3.8
Economic empowerment	4.2	5.9	21.2	39.4	24	5.2
Information verification and validation	3.7	7.2	19.8	42.5	22.6	4.2
Equality and inclusion	4.5	7.9	24.5	37.6	21.5	4
Public and social sector management	4.5	7.7	20.9	41.5	20.2	5.2
Culture	7.4	14.2	31.6	31	10.5	5.2

As shown in Figure 6, these orderings were broadly replicated in terms of respondents' opinions regarding whether these areas merited funding by government and commercial sectors. Indeed, there was surprisingly little difference between respondents' support for state and commercial involvement in most areas.

> **"The only area with notably lower levels of support is the use of AI to generate culture for popular consumption (such as films, books, music or art)"**

Figure 6. Self-reported support for the development of AI in each sector.

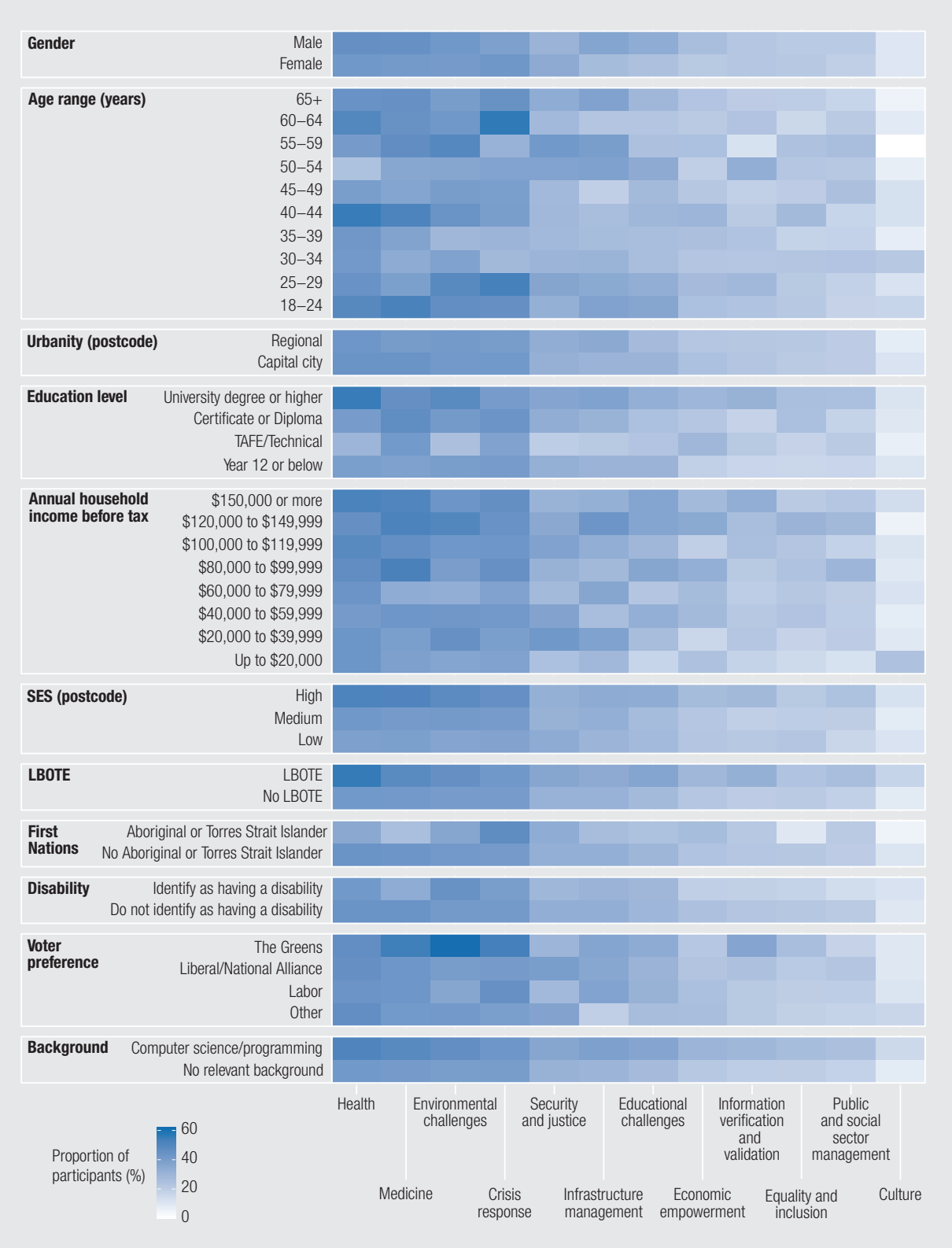


NB. Survey item asked: *"How much do you support or oppose…"* Data are percentage of respondents indicating that they 'Strongly Support' or 'Somewhat Support'.

Given the sometimes-contentious nature of issues such as the environment and matters of justice in Australian media and political debate, we might expect public support for the use of AI in these different areas to vary along lines of age, socio-economic status or political preference. However, there was surprisingly little sense of differentiation, suggesting that perhaps AI is not seen as an especially politicised, divisive or contentious issue.

Figure 7 shows the proportion of people who strongly support the use of AI in different areas, according to their demographic characteristics. This shows, for example, that there are no differences by income group or urbanity (regional/capital city) in the level of support for any of the AI application areas in the survey. Indeed, the main trend appears to be a link between having a background in computer science/programming and higher levels of support across seven of the twelve AI application areas (the exceptions being: health, medicine, environmental challenges, crisis response and public & social sector management).

Figure 7. Proportion of people who strongly support the use of AI by area and demographic characteristics.



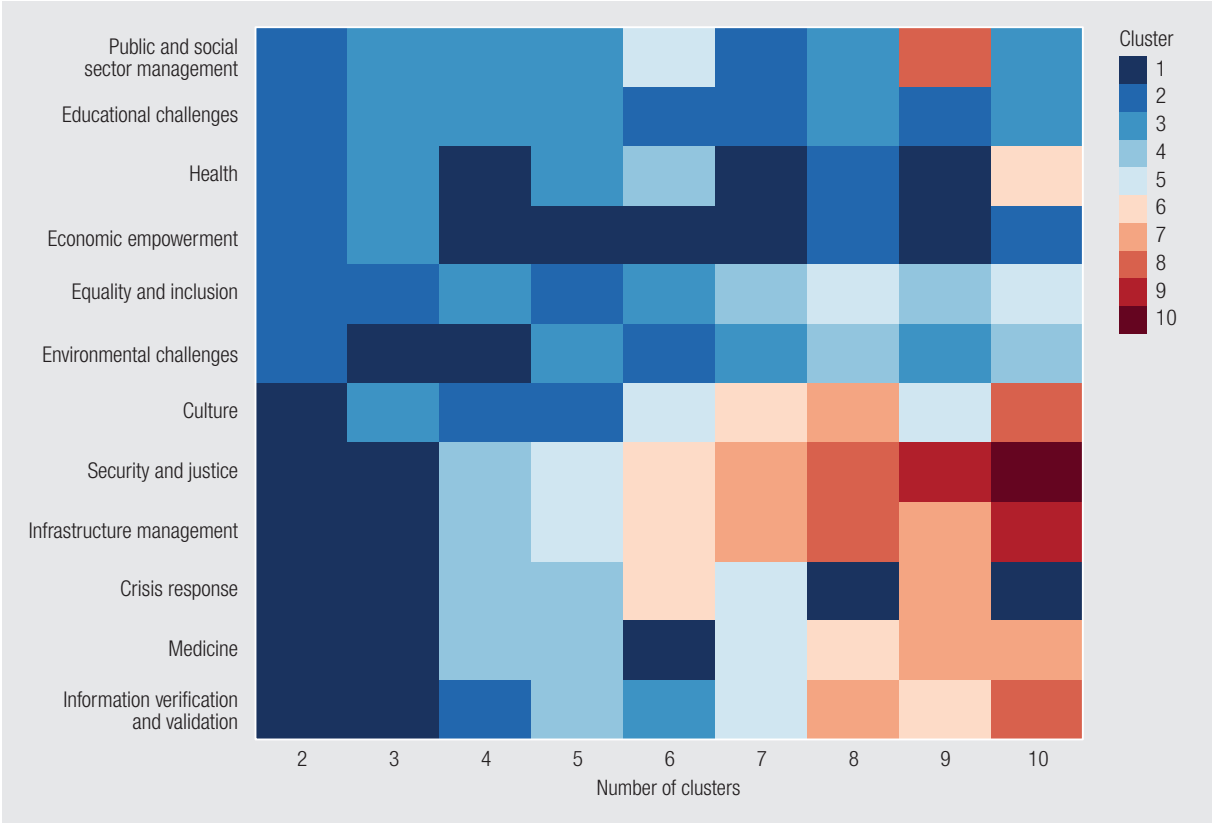
NB. Survey item asked: *"How much do you support or oppose the use of AI in…"* Data are percentage of respondents indicating that they 'Strongly Support'.

This analysis also shows that eleven of the twelve areas fall into three subgroups – each with different levels of support amongst all demographic groups:

- **Very high support:** Medicine, health, environmental challenges and crisis response;
- **High levels of support:** Security and justice, infrastructure management and educational challenges;
- **Relatively high level of support:** Economic empowerment, information verification and validation, equality and inclusion and public and social sector management.

However, cluster analysis does not confirm any statistical basis for grouping together these different AI application areas according to their level of support. For example, when three clusters are created, health and medicine are placed in the same group, but this is not the case when any other number of clusters are created. For completeness, Figure 8 shows the different groups that would be achieved using K-means clustering.

Figure 8. Cluster of areas of AI development according to the number of clusters in which the areas are divided. Areas with the same colour belong to the same group (cluster).



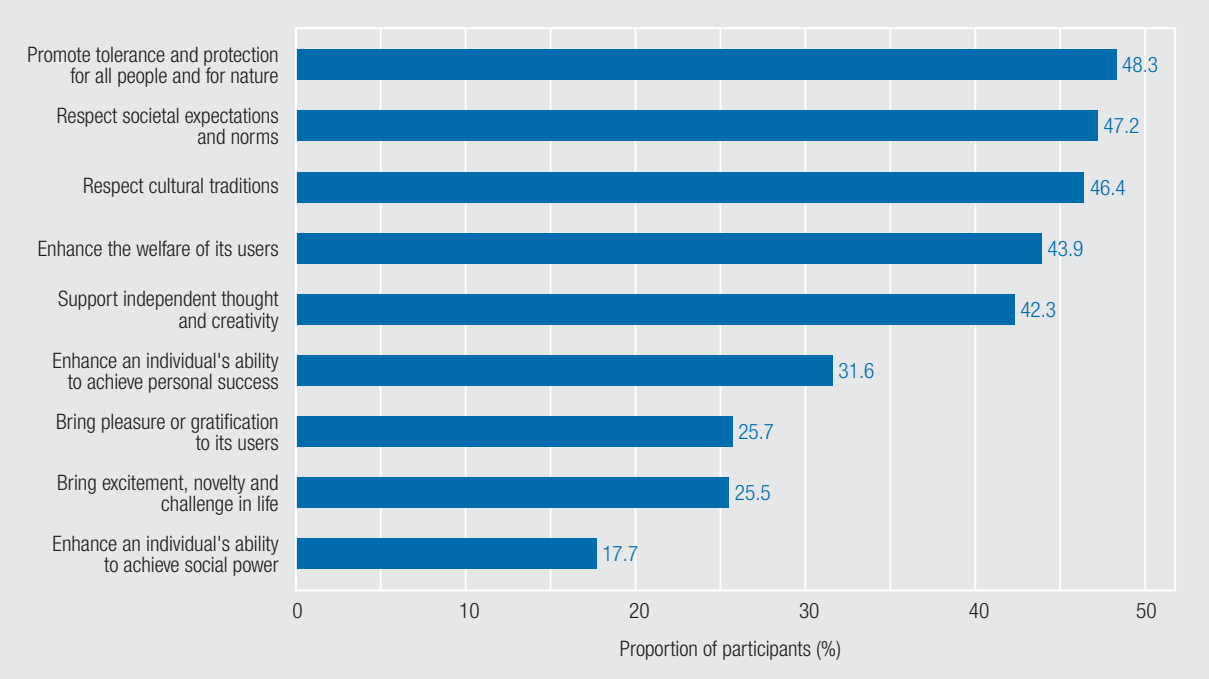
NB. Survey item asked: *"How much do you support or oppose the use of AI in…"* Data are percentage of respondents indicating that they 'Strongly Support'

3.2 Public opinions on the social values that should underpin AI use

The survey then explored respondents' views of the nature of these possible 'social goods' associated with AI development. Respondents were presented with a series of 10 statements relating to value statements regarding AI and society. As shown in Figure 9 and Table 8, the three most supported sets of values (in terms of respondents indicating that they 'strongly agree') are those that might be described as concerned with continuity rather than change – i.e. related to ideas of respecting and protecting established conditions, traditions and norms.

Conversely, the only value statement that provoked a notably divided response related to the use of AI to "enhance an individual's ability to achieve social power". This may well reflect a general discomfort amongst many respondents with measures that challenge or alter the status quo. In this respect, it is telling that the prospect of AI being used to "enhance an individual's ability to achieve personal success" attracted notably higher levels of support.

Figure 9. Self-reported agreement with value statements about AI.



NB. Data in bar chart are percentage of respondents indicating that they 'Totally agree' with each value statement. Survey asked: *"How much do you agree or disagree with the following statements? AI should…"*

Table 8. Levels of respondent dis/agreement with statements relating to AI and social good.

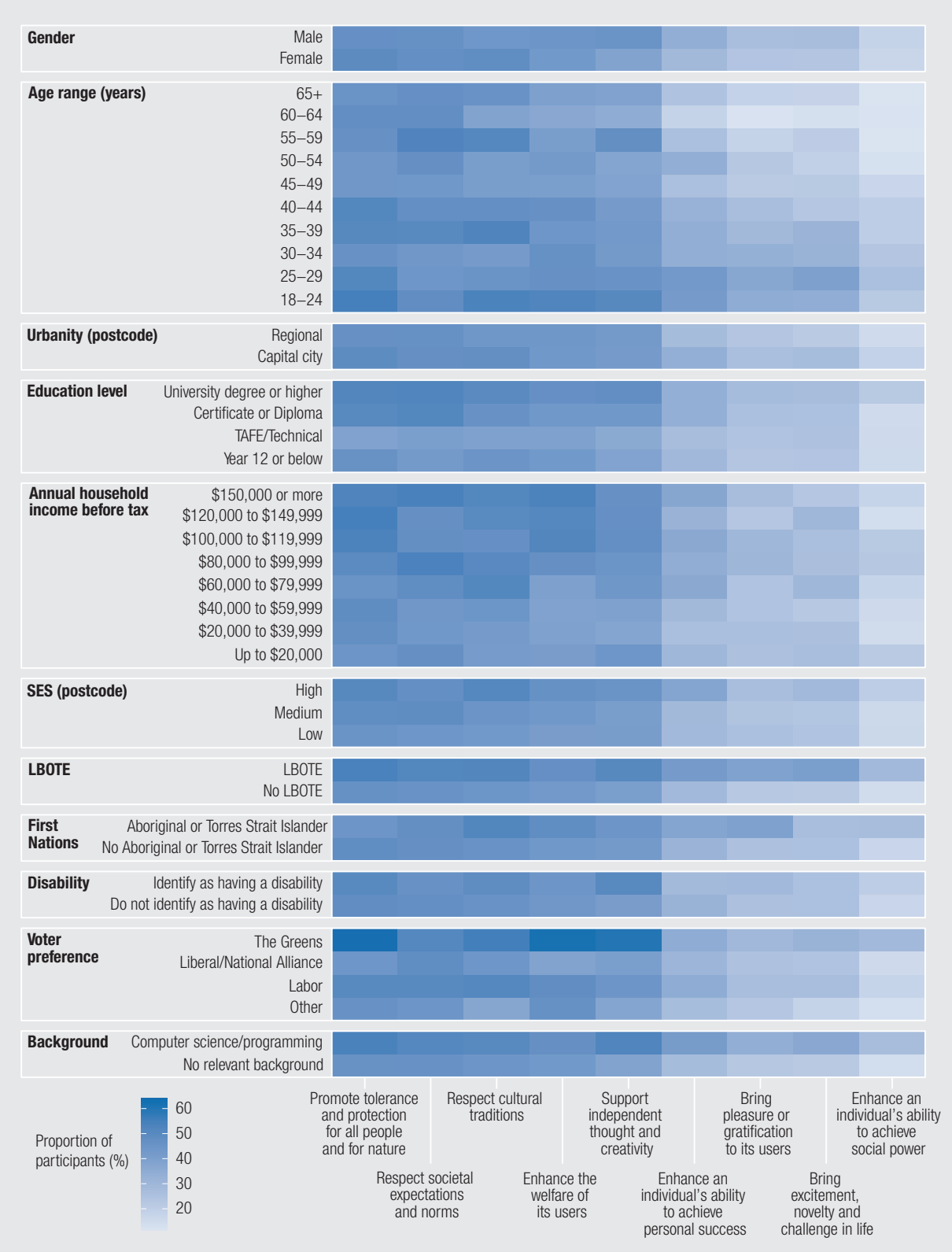
	Totally disagree	Tend to disagree	Tend to agree	Totally agree	Don't know
Promote tolerance and protection for all people and for nature	1.7	4.0	37.9	48.3	8.0
Respect societal expectations and norms	1.3	2.9	40.5	47.2	8.1
Respect cultural traditions	1.8	5.2	37.7	46.4	8.9
Enhance the welfare of its users	1.3	3.5	43.0	43.9	8.3
Support independent thought and creativity	1.7	5.1	41.6	42.3	9.3
Enhance an individual's ability to achieve personal success	2.7	8.9	45.6	31.6	11.3
Bring pleasure or gratification to its users	2.9	10.2	45.6	25.7	15.6
Bring excitement, novelty and challenge in life	2.5	11.7	46.7	25.5	13.5
Enhance an individual's ability to achieve social power	9.7	23.0	33.9	17.7	15.7

NB. Data are percentage of respondents indicating that they 'Totally agree' with each value statement. Survey asked: *"How much do you agree or disagree with the following statements? AI should…"*

As with respondents' support for different uses of AI, further analysis finds support for these values to be shared by most respondents, with few trends or differences among demographic groups (see Figure 10). At best, five specific areas of difference can be identified:

- Respondents with a computer science/programming background differ in terms of being more likely to 'totally agree' with the idea that AI should 'promote tolerance and protection for all people and for nature'. Otherwise, having a computer science/programming background makes no difference to how people responded.
- Those with an income of $80,000 or above are more likely to 'totally agree' with the idea that AI should 'enhance the welfare of its users' than those with lower levels of income.
- Those who speak a language other than English at home are more likely to 'totally agree' with the idea that AI should be 'promoting tolerance and protection for all people and for nature' and 'supporting independent thought and creativity', than respondents only speaking English at home.
- Those living in high SES postcode areas are more likely to 'totally agree' with the idea that AI should 'enhance an individual's ability to achieve personal success' and 'support independent thought and creativity, than those living in medium and low SES postcode areas.
- Respondents who voted for Labor and The Greens are also more likely to totally agree with five out of the nine value statements, in comparison to those who voted for other parties:
  - AI should respect societal expectations and norms
  - AI should enhance an individual's ability to achieve personal success
  - AI should bring excitement, novelty and challenge in life
  - AI should respect cultural traditions
  - AI should enhance an individual's ability to achieve social power (The Greens only)

Figure 10. Proportion of people who 'totally agree' with each value statement: 'AI should…' by demographic characteristics.



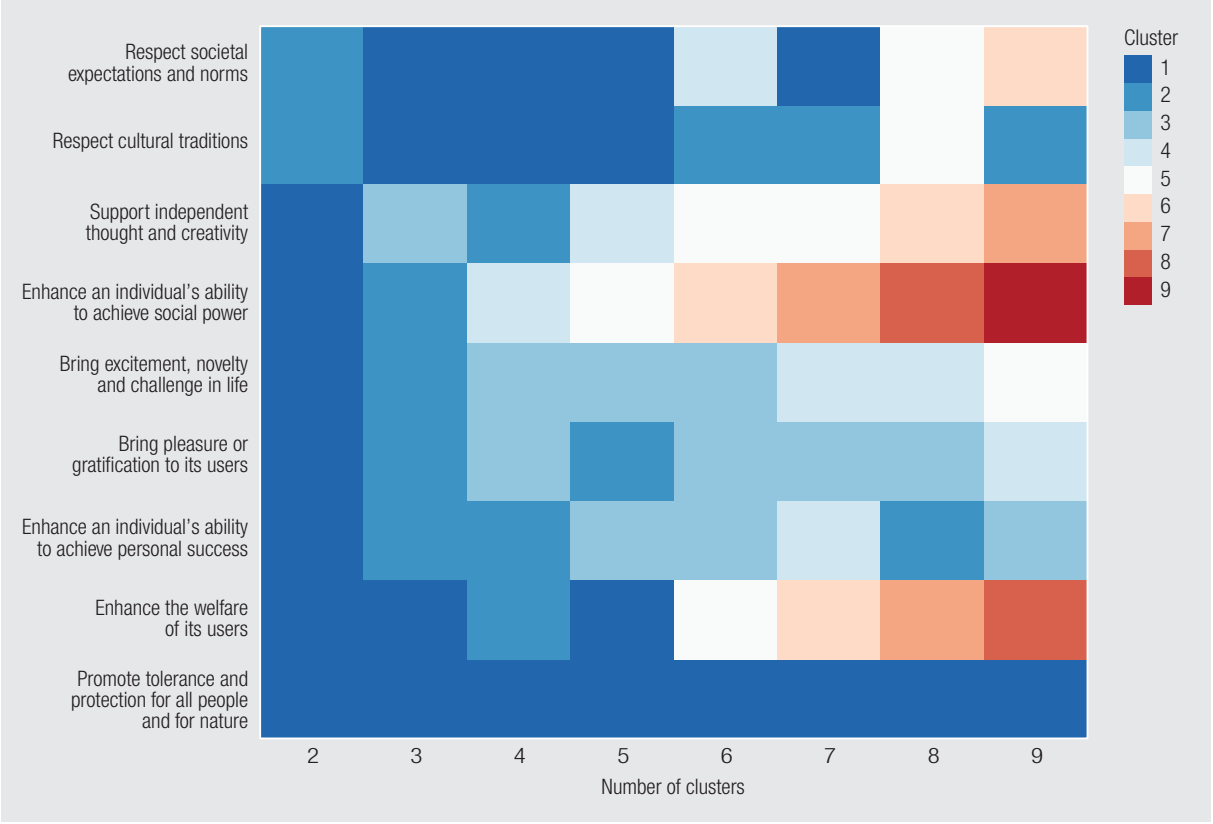
NB. Survey asked: *"How much do you agree or disagree with the following statements? AI should…"*

As before, we were interested to see if there is any distinct patterning or grouping to these responses. For example, Figure 10 appears to show congruence around two sub-groups of values:

- AI should … Promote tolerance and protection for all people and for nature; Respect societal expectations and norms; Respect cultural traditions; Enhance the welfare of its users; Support independent thought and creativity.
- AI should … Enhance an individual's ability to achieve personal success; Bring pleasure or gratification to its users; Bring excitement, novelty and challenge in life.

Nonetheless, these distinctions are not statistically confirmed by further cluster analysis. In short, fit indicators show no robust grouping for these two sets of 'social goods' (see Appendix A3). Figure 11 shows how support for different values are (dis)associated as the numbers of clustered are changed in the modelling.

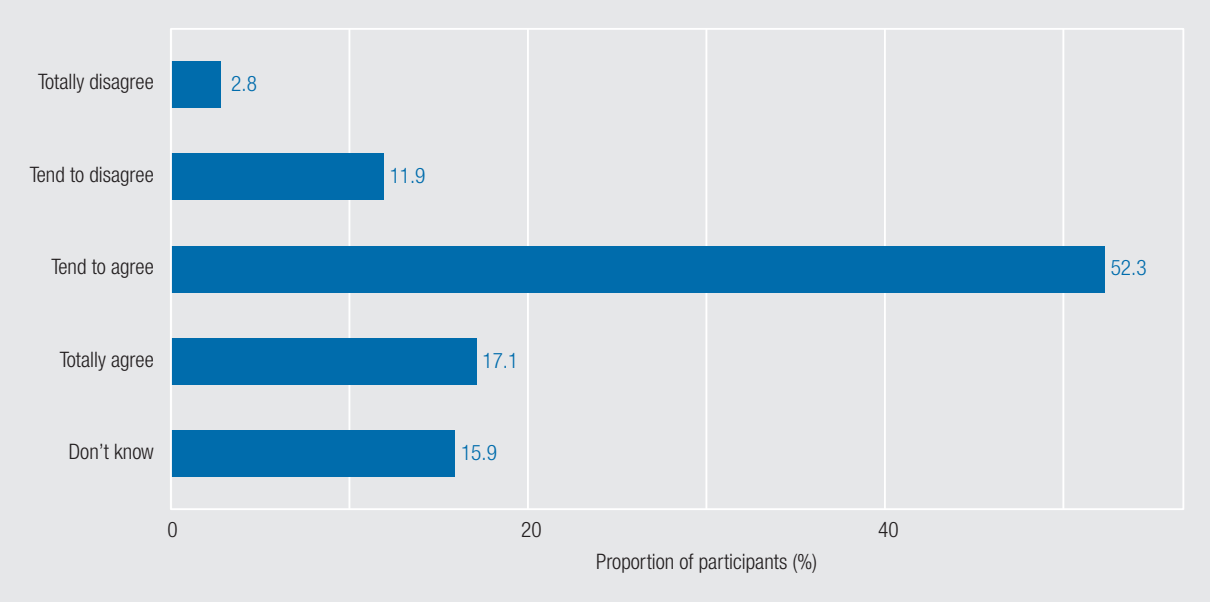
Figure 11. Cluster of value statements according to the number of clusters in which the statements are divided. Values with the same colour belong to the same group (cluster).



NB. Survey asked: *"How much do you agree or disagree with the following statements? AI should…"*

Finally, having considered these different applications and issues, respondents were then asked whether they anticipate that AI will 'on the whole' do more 'social good' than 'social harm'. In line with the earlier indications of overall support/opposition to the development of AI, a sizable majority of respondents (69.4%) agreed that AI will do more social good than harm (see Figure 12).

Figure 12. Self-reported levels of agreement with the statement: "on the whole, AI is a technology that will do more 'social good' than 'social harm'".



NB. Survey item asked: *"How much do you agree or disagree that on the whole, AI is a technology that will do more 'social good' than 'social harm?"* Data are percentage of respondents.

4. Public opinions on potential AI harms and challenges

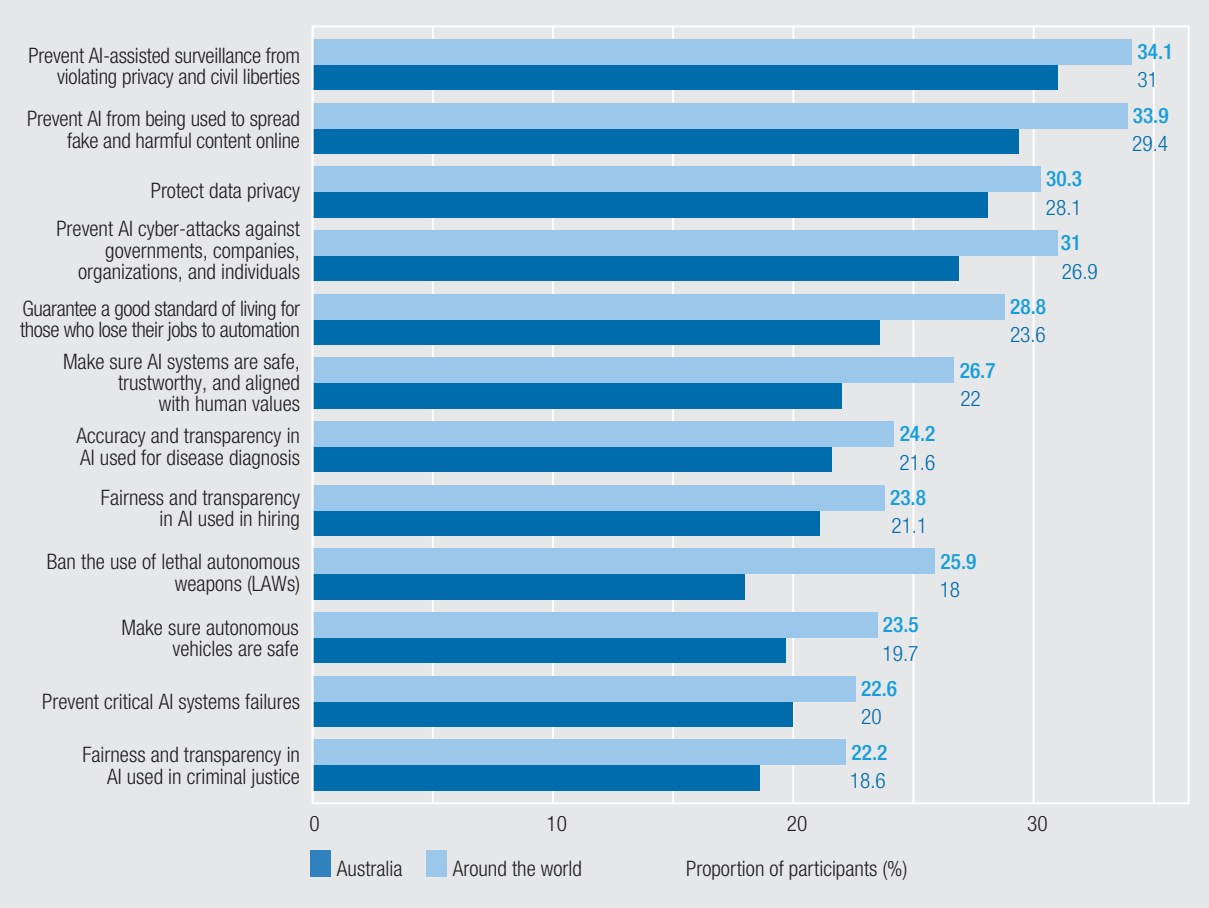
Table 9. Descriptions of potential AI 'harms' and challenges used in survey. Taken from Zhang & Dafoe (2019).

Harms and challenges	Description
Fairness and transparency in AI used in hiring	Increasingly, employers are using AI to make hiring decisions. AI has the potential to make less biased hiring decisions than humans. But algorithms trained on biased data can lead to hiring practices that discriminate against certain groups. Also, AI used in this application may lack transparency, such that human users do not understand what the algorithm is doing, or why it reaches certain decisions in specific cases.
Fairness and transparency in AI used in criminal justice	Increasingly, the criminal justice system is using AI to make sentencing and parole decisions. AI has the potential to make less biased sentencing and parole decisions than humans. But algorithms trained on biased data could lead to discrimination against certain groups. Also, AI used in this application may lack transparency such that human users do not understand what the algorithm is doing, or why it reaches certain decisions in specific cases.
Accuracy and transparency in AI used for disease diagnosis	Increasingly, AI software has been used to diagnose diseases, such as heart disease and cancer. One challenge is to make sure the AI can correctly diagnose those who have the disease and not mistakenly diagnose those who do not have the disease. Another challenge is that AI used in this application may lack transparency such that human users do not understand what the algorithm is doing, or why it reaches certain decisions in specific cases.
Protect data privacy	AI can be used to sustain biodiversity and combat the depletion of natural resources, pollution, and climate change. For example, AI tools are used in conservation efforts across the world – detecting illegal logging in vulnerable forest areas, or over-fishing in restricted areas.
Make sure autonomous vehicles are safe	Companies are developing self-driving cars and trucks that require little or no input from humans. Some worry about the safety of autonomous vehicles for those riding in them as well as for other vehicles, cyclists, and pedestrians.
Prevent AI from being used to spread fake and harmful content online	AI has been used by governments, private groups, and individuals to harm or manipulate internet users. For instance, automated bots have been used to generate and spread false and/or harmful news stories, audios, and videos.
Prevent AI cyber-attacks against governments, companies, organizations, and individuals	Computer scientists have shown that AI can be used to launch effective cyber-attacks. AI could be used to hack into servers to steal sensitive information, shut down critical infrastructures like power grids or hospital networks, or scale up targeted phishing attacks.
Prevent AI-assisted surveillance from violating privacy and civil liberties	AI can be used to process and analyse large amounts of text, photo, audio, and video data from social media, mobile communications, and CCTV cameras. Some worry that governments, companies, and employers could use AI to increase their surveillance capabilities.
Make sure AI systems are safe, trustworthy, and aligned with human values	As AI systems become more advanced, they will increasingly make decisions without human input. One potential fear is that AI systems, while performing jobs they are programmed to do, could unintentionally make decisions that go against the values of its human users, such as physically harming people.
Ban the use of lethal autonomous weapons (LAWs)	Lethal autonomous weapons (LAWs) are military robots that can attack targets without control by humans. LAWs could reduce the use of human combatants on the battlefield. But some worry that the adoption of LAWs could lead to mass violence. Because they are cheap and easy to produce in bulk, national militaries, terrorists, and other groups could readily deploy LAWs.
Guarantee a good standard of living for those who lose their jobs to automation	Some forecast that AI will increasingly be able to do jobs done by humans today. AI could potentially do the jobs of blue-collar workers, like truckers and factory workers, as well as the jobs of white-collar workers, like financial analysts or lawyers. Some worry that in the future, robots and computers can do most of the jobs that are done by humans today.
Prevent critical AI systems failures	As AI systems become more advanced, they could be used by the military or in critical infrastructure, like power grids, highways, or hospital networks. Some worry that the failure of AI systems or unintentional accidents in these applications could cause 10% or more of all humans to die.

The next section of the survey explored respondents' opinions of likely societal harms and challenges associated with the increased development of AI technologies and their implementation across society. Respondents were presented with 5 randomly-selected issues from a long-list of 12 likely challenges (outlined in Table 9).

In terms of respondents' expectations of the likelihood that these issues impact on people within the next 10 years, the most likely issues were seen to relate to violations relating to privacy and the spread of fake online content. Conversely, issues relating to lethal autonomous weapons, autonomous vehicles, critical system failures and criminal justice were anticipated to be the least likely issues to emerge over the next 10 years. All these issues were slightly more likely to be seen to be a problem in countries other than Australia (see Figure 13).

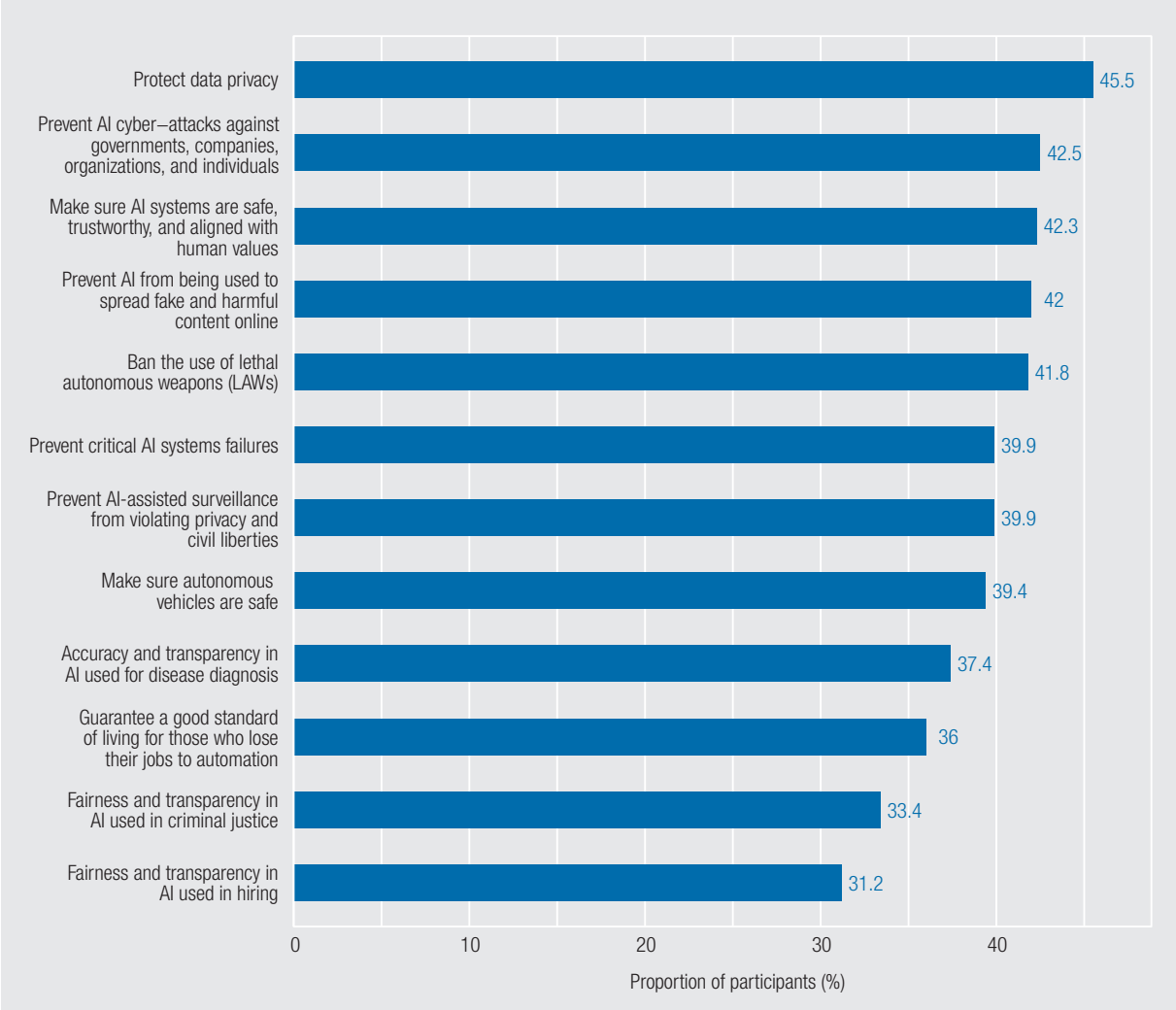
Figure 13. Perceived risk for each future scenario in Australia and the rest of the world.



NB. Data are percentage of respondents indicating that they think scenario is 'Very likely'. Survey asked: *"In the next 10 years, how likely do you think it is that this will impact large numbers of people…?"*

The survey then asked respondents to indicate which of these issues merit careful management by government and technology companies. As before, issues relating to data privacy, cyber-attacks, safety and fake online content were among the issues most likely to be 'strongly supported' to require government and industry management (see Table 10 and Figure 14). Conversely, issues relating to 'fairness and transparency' in the use of AI in employment hiring and criminal justice decisions were least likely to be 'strongly supported' as requiring government and industry management. That said, these are relative differences in terms of the levels of expressed support. Very few respondents indicated that careful regulation and management was not important for any of these issues.

Figure 14. Perceived importance of managing AI challenges.



NB. Data are percentage of respondents indicating that they 'strongly support'. Survey item asked: *"In the next 10 years, how important is it for tech companies and governments to carefully manage this challenge?"*

Table 10. Respondent support for each of the areas.

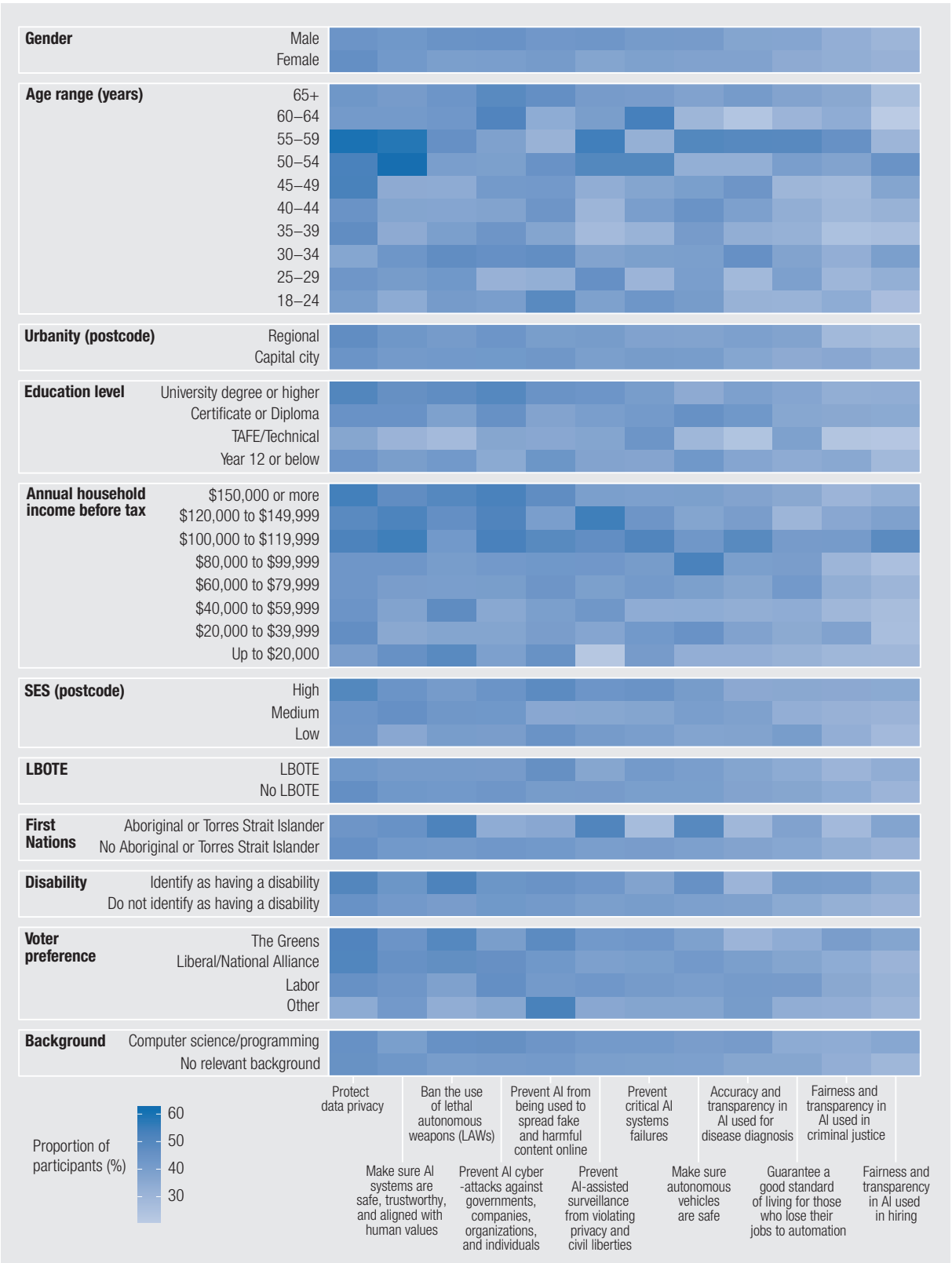
	Not at all important	Not too important	Somewhat important	Important	Very important	Don't know
Protect data privacy	0.6	1.0	8.5	38.4	45.5	6.1
Prevent AI cyber-attacks against governments, companies, organizations, and individuals	0.4	0.5	7.3	43.0	42.5	6.3
Make sure AI systems are safe, trustworthy, and aligned with human values	0.5	0.5	6.9	44.2	42.3	5.7
Prevent AI from being used to spread fake and harmful content online	0.4	1.0	8.7	41.8	42.0	6.0
Ban the use of lethal autonomous weapons (LAWs)	0.7	1.6	7.9	41.1	41.8	7.0
Prevent AI-assisted surveillance from violating privacy and civil liberties	0.2	0.5	10.9	41.1	39.9	7.4
Prevent critical AI systems failures	0.6	1.3	9.8	43.5	39.9	4.9
Make sure autonomous vehicles are safe	0.4	1.2	7.6	46.4	39.4	5.0
Accuracy and transparency in AI used for disease diagnosis	0.6	1.3	9.3	45.7	37.4	5.7
Guarantee a good standard of living for those who lose their jobs to automation	0.8	0.7	10.6	44.1	36.0	7.7
Fairness and transparency in AI used in criminal justice	0.7	1.8	11.0	46.1	33.4	7.0
Fairness and transparency in AI used in hiring	0.4	1.0	12.4	47.5	31.2	7.5

NB. Survey item asked: *"In the next 10 years, how important is it for tech companies and governments to carefully manage this challenge?"*

Again, there seems to be considerable consensus across all demographic groups with regards to the relative importance of managing these different challenges. In short, these are all issues that people are likely to agree are important – suggesting no notable divisions in public opinion. At best, the following specific differences can be noted from the bivariate analysis (in Figure 15):

- Those with a background in computer science/programming are slightly more likely to agree that eight of the twelve challenges should be addressed. The four areas where this was not the case are: 'making sure autonomous vehicles are safe'; 'preventing AI from being used to spread fake and harmful content online'; 'guaranteeing a good standard of living for those who lose their jobs to automation'; and 'preventing critical AI systems failures'.
- Respondents living in capital cities are slightly more likely to indicate 'very important' than those living in regional areas the need to address issues relating to 'fairness and transparency in AI' when used in hiring *and* in criminal justice.
- Those who voted for Labor and The Greens are more likely to indicate 'very important' the challenge of 'preventing critical AI systems failures'. Otherwise, there are no differences in opinions about these AI challenges by voting preference.

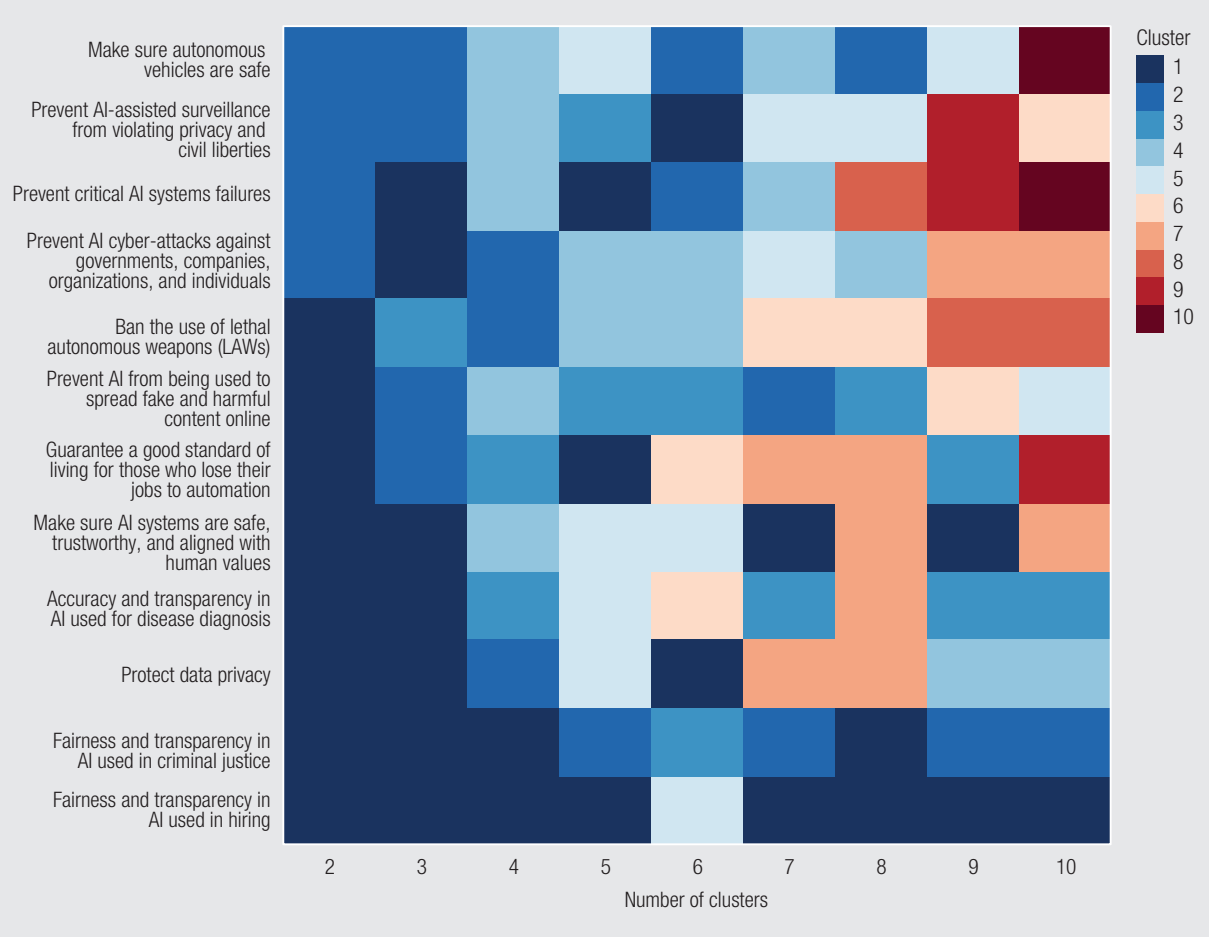
Figure 15. Proportion of people who indicate 'very important' for the existence of harms and challenges of AI by challenge and demographic characteristics.



NB. Data are percentage of respondents indicating that they think scenario is 'very important'. Survey item asked: *"In the next 10 years, how important is it for tech companies and governments to carefully manage this challenge?"*

As in previous sections, we were interested to see if these twelve different challenges fell into distinct groups or clusters with similar levels of perceived importance. Nevertheless, K-means cluster analysis proved to be fairly unstable and does not reveal a robust grouping of challenges (see Figure 16).

Figure 16. Cluster of challenges of AI according to the number of clusters in the analysis. Challenges with the same colour belong to the same group (cluster).

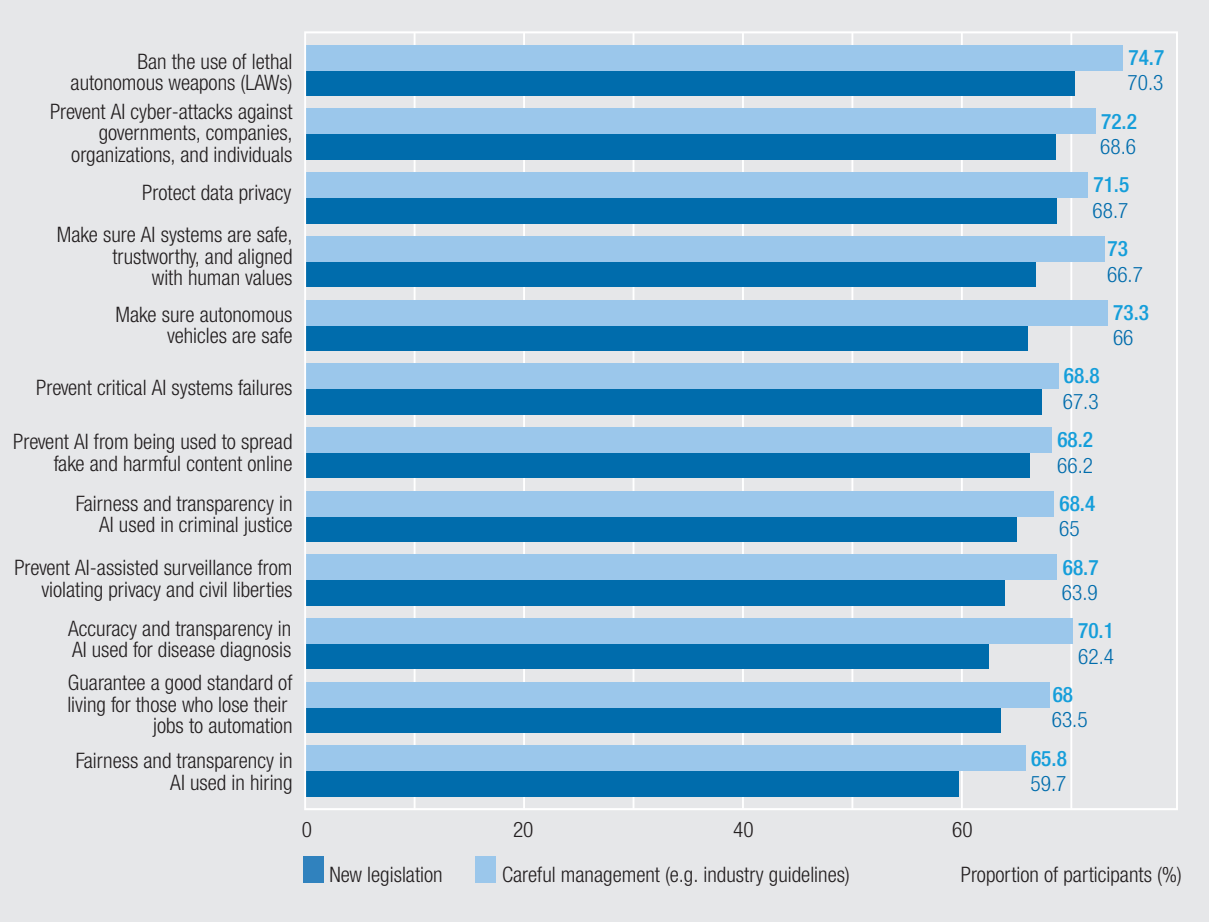


NB. Survey item asked: *"In the next 10 years, how important is it for tech companies and governments to carefully manage this challenge?"*

The majority of respondents indicate that they 'totally agree' with the need for new legislation to address these different AI challenges (see Figure 17). Despite being acknowledged earlier in the survey as something that people consider to be less likely to arise in the next ten years, banning the use of lethal autonomous weapons was the issue that was most likely seen to require industry management and government legislation. Similarly, ensuring the safety of autonomous vehicles was also raised as a high priority as requiring industry guidelines.

These consistently high levels of recognition might point to a public tendency to see 'AI' as a homogeneous issue regardless of context of specific application. These findings might also point to a lack of distinction between different forms of AI technology.

Figure 17. Self-reported agreement with the management of AI challenges by the government and private companies.



NB. Data are percentage of respondents indicating that they 'Totally Agree'. Survey item asked: *"To what extent do you agree or disagree with the following statements?"*

> **"The majority of the public agrees with the need for industry guidelines and/or new legislation to address all possible harms associated with AI"**

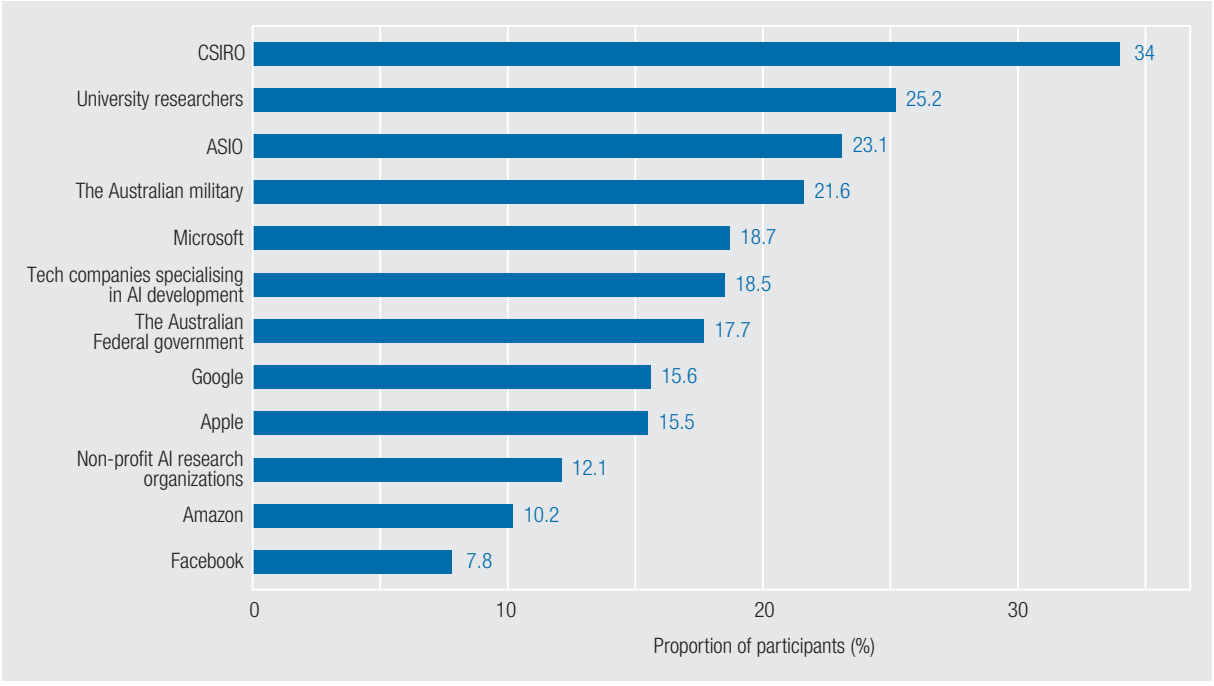
5. Public trust in AI development and governance

The survey then asked respondents to consider which institutions, organisations and groups they have most confidence in to address the challenges associated with AI and society. Following Zhang and Dafoe (2019), the survey offered a set of questions that first related to confidence in organisations to develop AI in best interests of the public, followed by questions relating to confidence in organisations to manage the development and use of AI.

In terms of the issue of AI development, survey respondents were most likely to express 'a great deal of confidence' in researchers and scientists – whether employed by Australia's federal science agency (CSIRO) or by universities. There were also relatively high levels of confidence in the security (ASIO) and military to develop AI in the best interests of the public.

In contrast, public confidence in industry actors was lower and more varied (Table 11 and Figure 18). For instance, 18.7% of respondents expressed a 'great deal of confidence' in Microsoft's ability to develop AI in the best interests of the public, compared with 7.8% expressing a 'great deal of confidence' in Facebook (for whom 35.3% of respondents expressed 'no confidence').

Figure 18. Self-reported level of confidence in different organisations to develop AI in the best interests of the public.



N.B. Data in bar chart are percentage of respondents indicating 'A great deal of confidence'. Survey asked: *"How much confidence, if any, do you have in each of the following to develop AI in the best interests of the public?"*

> **"The Australian public has high levels of confidence in CSIRO and university researchers to develop AI in the best interests of the public"**

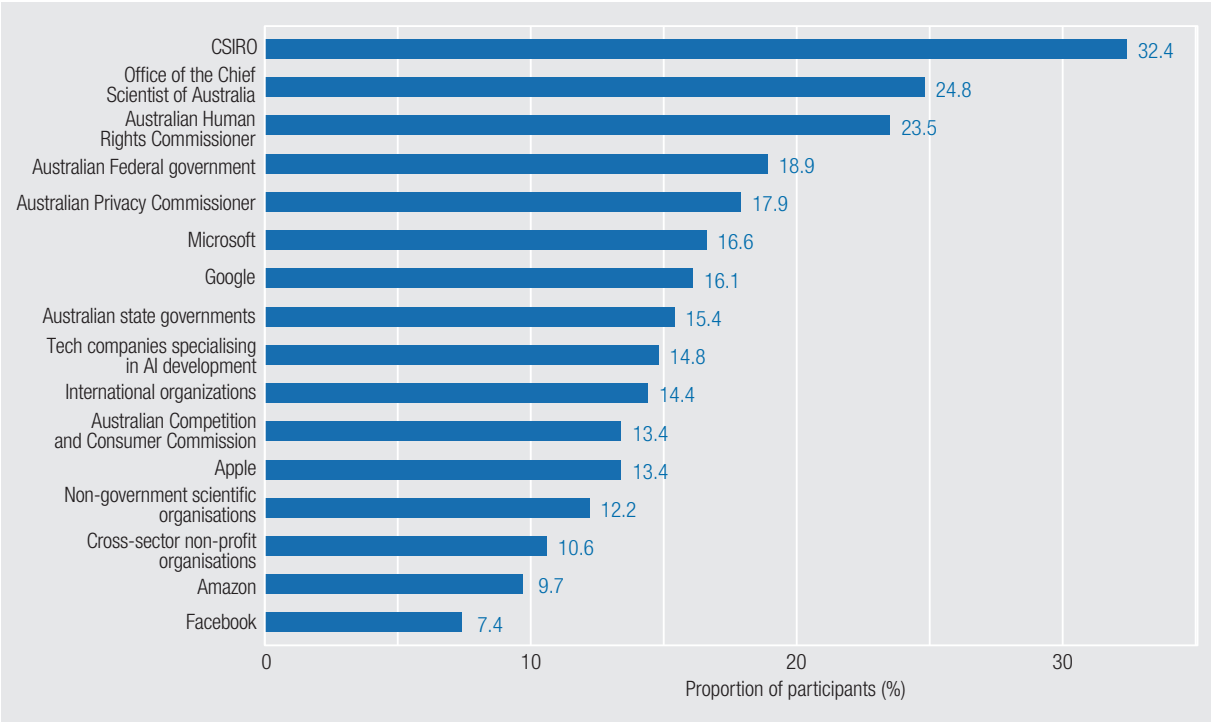
Table 11. Proportion of respondents according to their level of confidence in different organisations to develop AI in the best interests of the public.

	No confidence	Not too much confidence	A fair amount of confidence	A great deal of confidence	Don't know
CSIRO	2.6	8.4	43.5	34.0	11.5
University researchers	3.4	11.8	49.0	25.2	10.7
ASIO	6.5	15.1	44.5	23.1	10.8
The Australian military	8.4	18.7	41.0	21.6	10.2
Microsoft	6.9	24.1	39.1	18.7	11.2
Tech companies specialising in AI development	6.0	21.8	40.4	18.5	13.3
The Australian Federal government	8.0	21.1	42.3	17.7	10.8
Google	13.7	30.9	30.3	15.6	9.3
Apple	13.1	25.8	33.4	15.5	12.2
Non-profit AI research organizations	5.3	20.9	40.3	12.1	21.5
Amazon	18.3	31.0	28.0	10.2	12.6
Facebook	35.3	26.9	21.8	7.8	8.3

N.B. Survey asked: *"How much confidence, if any, do you have in each of the following to develop AI in the best interests of the public?"*

In terms of public confidence in institutions and organisations who might take responsibility for managing the development and use of AI, the highest levels of confidence were again placed in CSIRO researchers and scientists, with nearly one-third of respondents (32.4%) indicating that they had 'a great deal of confidence' (see Table 12 and Figure 19). This was followed by two independent Federal government bodies – the Office of the Chief Scientist of Australia and the Australian Human Rights Commissioner. As before, trust in IT industry actors varied according to the specific companies.

Figure 19. Self-reported confidence in different organisations to manage the development and use of AI in the best interests of the public.



N.B. Data in bar chart are percentage of respondents indicating 'A great deal of confidence'. Survey asked: *"How much confidence, if any, do you have in each of the following to manage the development and use of AI in the best interests of the public?"*

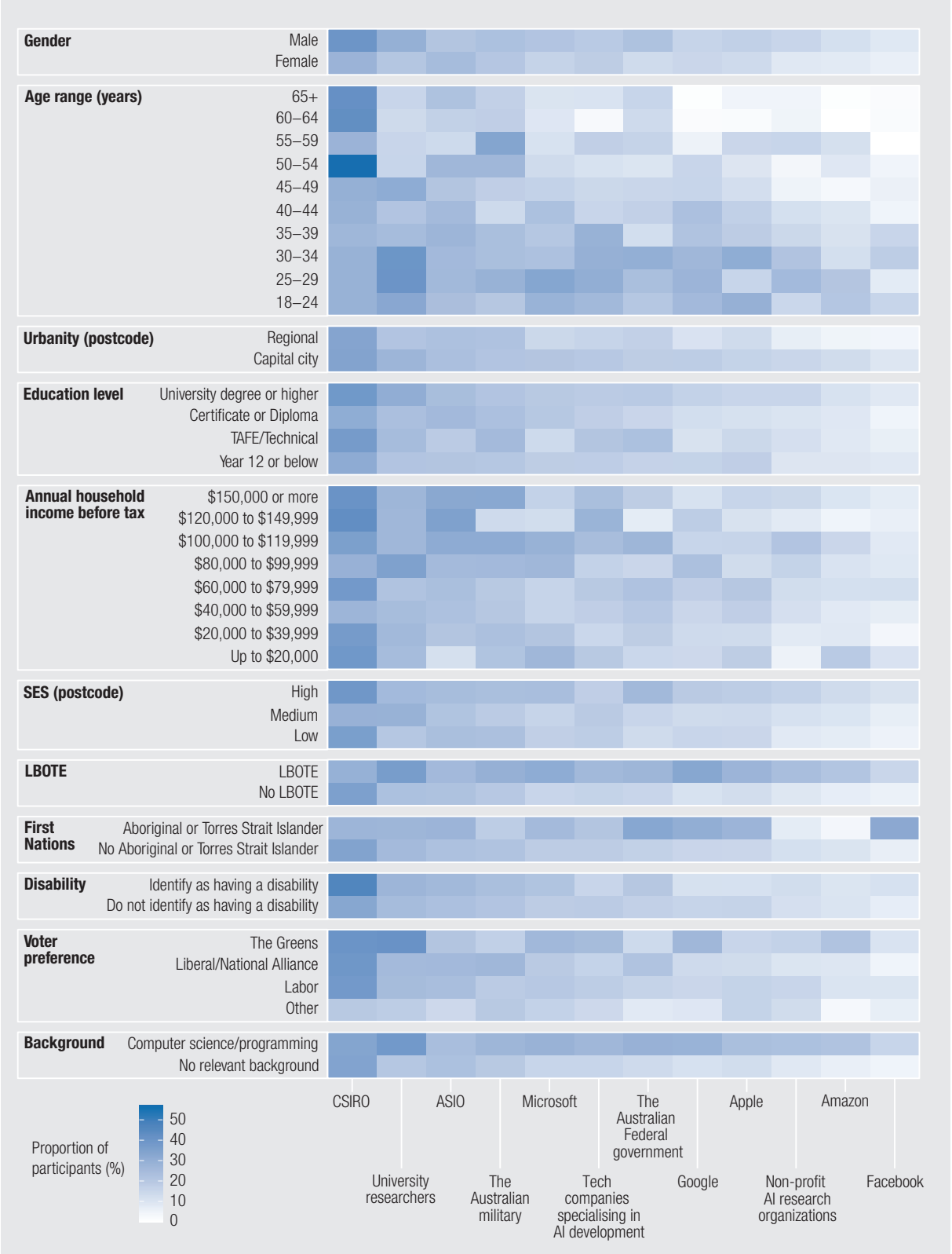
Table 12. Proportion of respondents according to their level of confidence in different organisations to manage the development and use of AI in the best interests of the public.

	No confidence	Not too much confidence	A fair amount of confidence	A great deal of confidence	Don't know
CSIRO	3.6	8.5	45.0	32.4	10.5
Office of the Chief Scientist of Australia	4.9	10.7	45.7	24.8	13.9
Australian Human Rights Commissioner	5.8	15.4	42.3	23.5	12.9
The Australian Federal government	7.3	23.3	42.9	18.9	7.6
Australian Privacy Commissioner	6.3	20.1	43.9	17.9	11.8
Microsoft	10.4	25.6	36.5	16.6	10.9
Google	19.0	26.3	29.9	16.1	8.6
Australian state governments	8.0	26.5	39.8	15.4	10.2
Tech companies specialising in AI development	7.4	23.7	44.3	14.8	9.9
International organizations	11.2	22.3	40.5	14.4	11.7
Australian Competition and Consumer Commission	6.6	21.6	46.6	13.4	11.8
Apple	15.6	27.6	34.1	13.4	9.3
Non-government scientific organisations	5.3	29.1	38.6	12.2	14.8
Cross-sector non-profit organisations	7.2	22.3	41.9	10.6	18.1
Amazon	24.3	30.8	21.7	9.7	13.4
Facebook	36.4	31.2	14.3	7.4	10.7

Bivariate analysis of these responses shows a number of patterns for the development of AI (see Figure 20):

- Support for CSIRO is uniformly high, with no significant differences in approval by demographic characteristics.
- Having a background in computer science/programming is not a significant influence on people's approval of any of these organisation.
- Respondents in capital cities are more likely to have a great deal of confidence in Google, Amazon and non-profit AI research organizations, and a fair amount of confidence in general (non-named) technology companies specialising in AI development.
- Those respondents living in households where languages other than English are spoken are more likely to have a great deal of confidence in the Australian Federal Government. Conversely, respondents in this group are less likely to select 'I don't know' for trust in the Australian military and less likely to have 'no confidence' in tech companies specialising in AI development.
- Females are more likely to select the 'I don't know' option for the following organisations: Australian Federal Government, ASIO, Google, Apple, Microsoft, Non-profit AI research organizations, and university researchers.
- Those who identify as having a disability are more likely to have 'no confidence' in the Australian Federal Government to develop AI in the best interests of the public.
- Support for university researchers was high across voters of all three main political parties, but notably higher amongst respondents who identified themselves as voting for The Greens.

Figure 20. Proportion of people with a great deal of confidence in the development of AI by organisation and demographic characteristics.



N.B. Data are percentage of respondents indicating 'A great deal of confidence'. Survey asked: *"How much confidence, if any, do you have in each of the following to manage the development and use of AI in the best interests of the public?"*

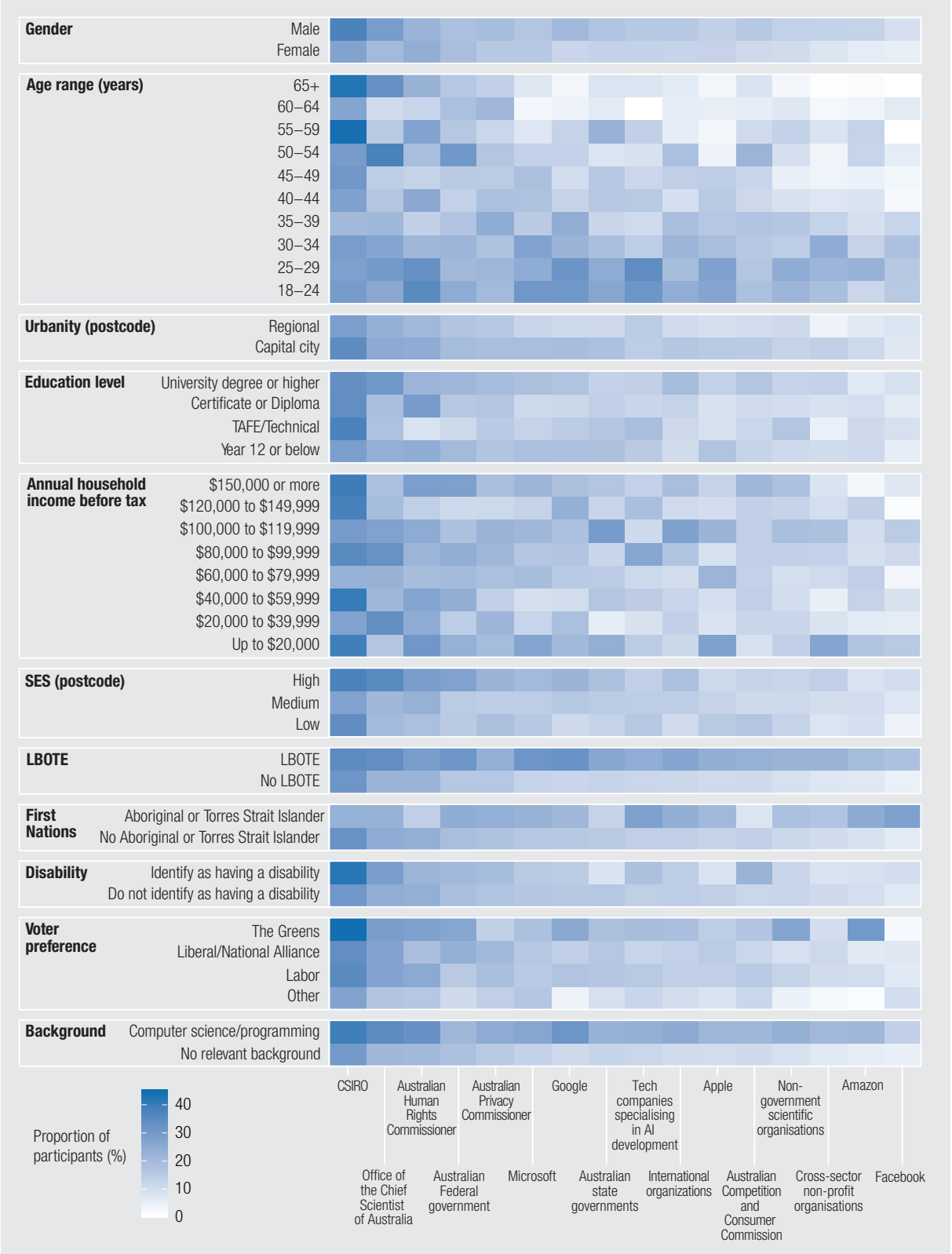
Some of these levels of confidence in organisations to manage AI were notably patterned according to respondents' backgrounds. For example, bivariate analysis of these responses shows the following points of note with regards to the management of AI (see Figure 21):

- Those with a background in computer science/programming are more likely to have a great deal of confidence in Australian state governments and general tech companies specialising in AI development. These respondents are more likely to respond 'I don't know' to other organisations.
- Female respondents are more likely to state 'I don't know' in terms of their level of trust for Australian state governments, International organizations, Australian Human Rights Commissioner, Office of the Chief Scientist of Australia, Australian Competition and Consumer Commission.
- Those living in regional areas are less likely to have 'a great deal of confidence' in Google, Amazon and cross-sector non-profit organisations.
- Those living in low SES postcode areas are less likely to have 'a great deal of confidence' in Google and more likely to have 'no confidence' in the Australian federal government to manage AI in the best interests of the public.
- Those who identify as having a disability and/or from households where languages other than English are spoken are more likely to have 'a great deal of confidence' in the Australian Competition and Consumer Commission.
- Support for CSIRO was high across voters of all three main political parties, but notably higher amongst but notably higher amongst respondents who identified themselves as voting for The Greens.

> **"The lowest levels of confidence to develop AI in the best interests of the public are expressed for Amazon and Facebook"**

> **"In terms of managing the development of AI, high levels of confidence are expressed for CSIRO, as well as independent government bodies such as the Office of The Chief Scientist, and the Australian Human Rights Commissioner"**

Figure 21. Proportion of people with 'a great deal of confidence' in the management of AI by organisation and demographic characteristics.



N.B. Data in bar chart are percentage of respondents indicating 'A great deal of confidence'. Survey asked: *"How much confidence, if any, do you have in each of the following to manage the development and use of AI in the best interests of the public?"*

Regardless of demographic background or political preference, there was very strong support for the establishment of a new regulatory body to address AI development. Specifically, 87.3% of the sample agreed with the statement that 'Australia requires a new regulatory body to govern and promote responsible innovation in the area of AI' (see Table 13).

Table 13. Respondents' dis/agreement over Australian involvement in the development of AI.

	Totally disagree	Tend to disagree	Tend to agree	Totally agree	Don't know
It is important for Australia to play a leading role in the international development of AI	1.3	5.6	42.2	39.2	11.8
Australia requires a new regulatory body to govern and promote responsible innovation in the area of AI	0.9	3.0	35.1	52.2	8.8

NB. Survey item asked: *"To what extent do you agree or disagree with the following statements?"* Data are percentage of respondents.

> **87.3% of respondents support establishing a new regulatory body in Australia to govern and promote responsible innovation in the area of AI**

6. Public expectations of the future development of AI

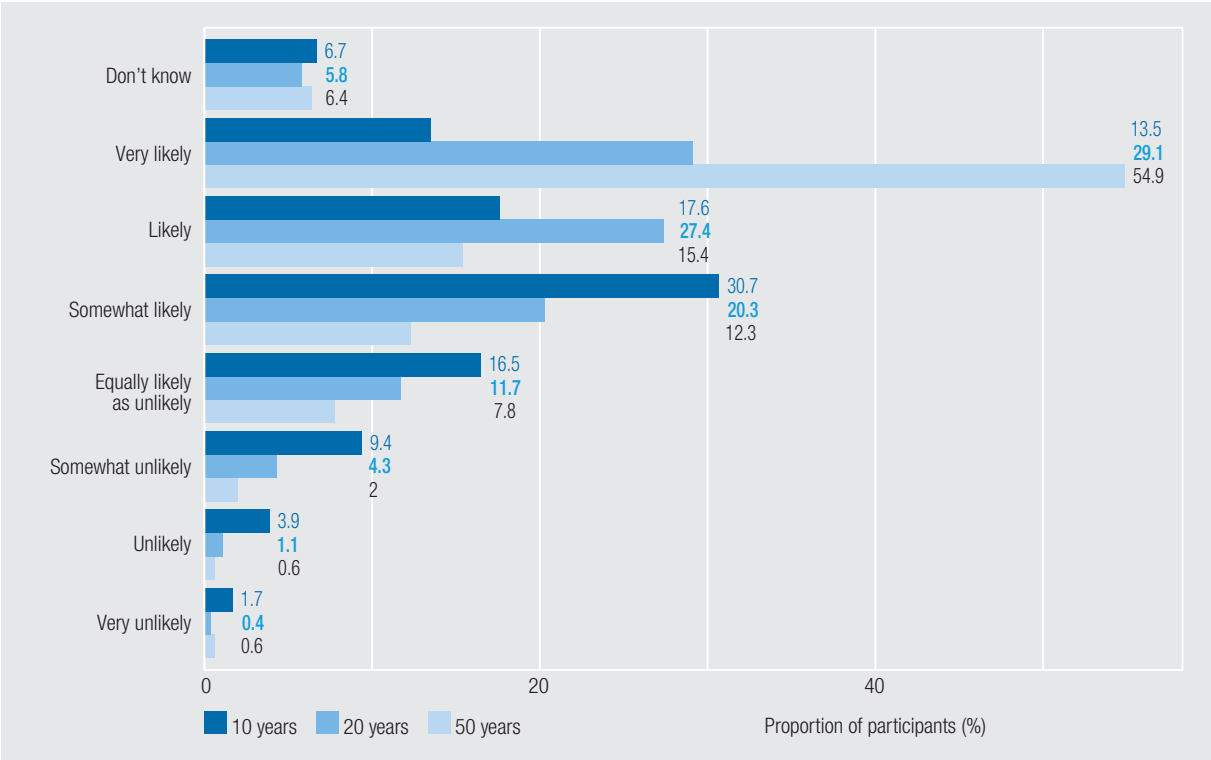
The next section of the survey replicated a set of items from Zhang and Dafoe's (2019) survey to explore respondents' opinions about the future development of AI. In particular, these questions relate to respondents' expectations of the likelihood of AI being developed to the point of being able to understand specific tasks better than the average human. Zhang and Dafoe (2019) describe this as 'high-level machine intelligence', corresponding with what is sometimes referred to as 'artificial general intelligence'.

Following Zhang and Dafoe, the survey first provided respondents with the following definition and proviso regarding the hypothetical nature of the concept:

> ***"The following questions are about high-level machine intelligence. This refers to machines having the capacity to learn or understand economically relevant tasks (i.e. jobs) better than the average human. This would include language and communication skills, for example asking subtle context specific questions.***
>
> ***High-level machine intelligence doesn't exist yet".***

Given the speculative nature of the concept, it is perhaps most useful to look at the respondents who anticipate the existence of high-level machine intelligence to be 'very likely'. As shown in Table 14 and Figure 22, 13.5% of respondents anticipate the existence of high-level machine learning within the next 10 years, rising to 54.9% expecting it to exist within the next 50 years.

Figure 22. Perceived likelihood of the existence of high-level machine intelligence in different time frames.



NB. Survey asked: *"In your opinion, how likely is it that high-level machine intelligence will exist in 10 years, 20 years and 50 years?"*

Table 14. Proportion of respondents according to their perceived likelihood of the existence of high-level machine intelligence in different time frames.

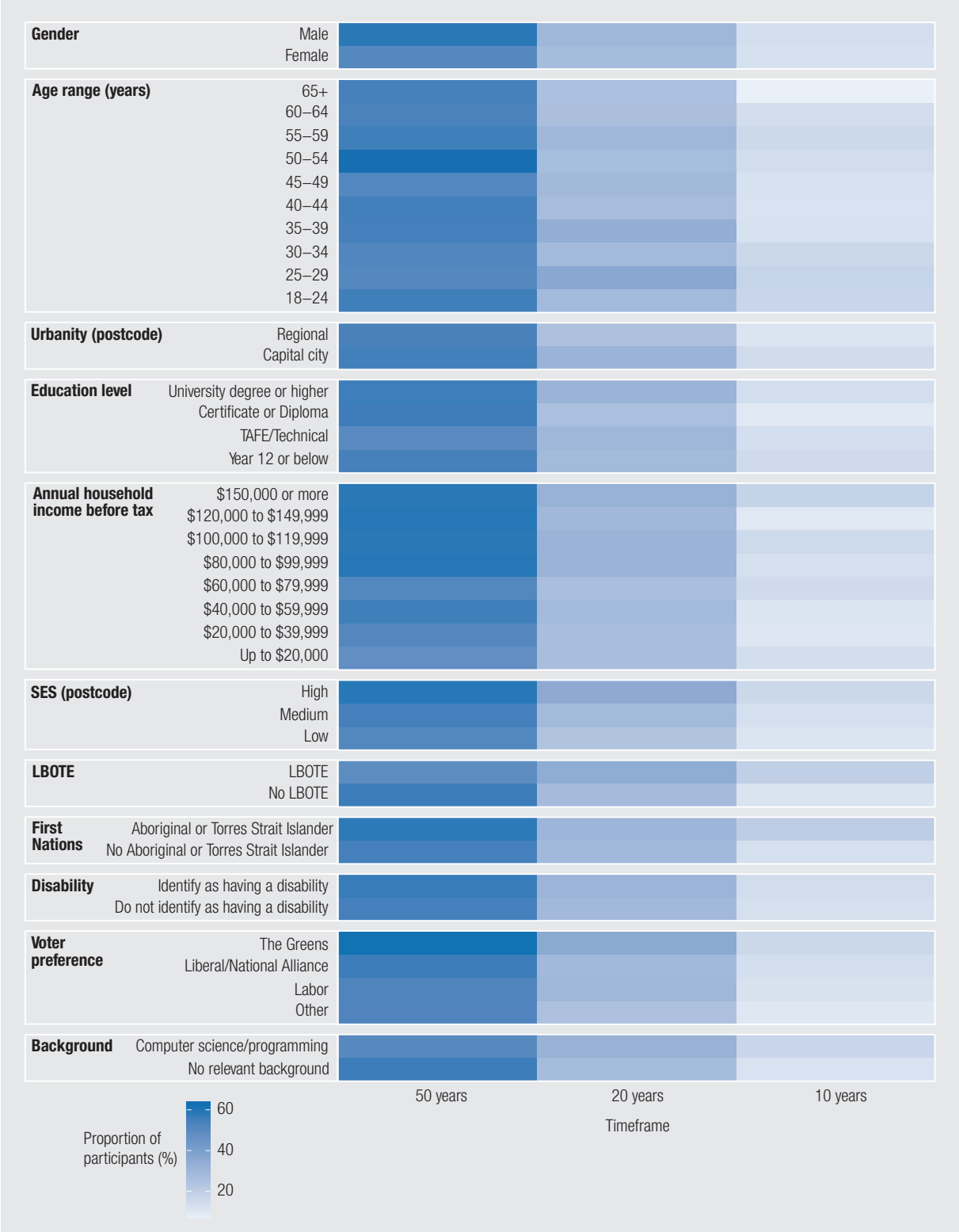
Time frame	Very unlikely	Unlikely	Somewhat unlikely	Equally likely as unlikely	Somewhat likely	Likely	Very likely	Don't know
10 years	1.7	3.9	9.4	16.5	30.7	17.6	13.5	6.7
20 years	0.4	1.1	4.3	11.7	20.3	27.4	29.1	5.8
50 years	0.6	0.6	2.0	7.8	12.3	15.4	54.9	6.4

NB. Survey asked: *"In your opinion, how likely is it that high-level machine intelligence will exist in 10 years, 20 years and 50 years?"*

We were interested to see whether this belief in the likelihood of high-level machine intelligence was more prevalent for different demographic groups. In terms of expectation of high-level machine intelligence within the next 10 years, the only significant subgroup that could be found was in terms of being in the 'high' socio-economic-status (SES) group – i.e. living in a postcode area with a high SES. Beyond postcode area, there are is no robust association between respondents' demographic characteristics and how likely they saw high-level machine intelligence being in the next 10 years (see Figure 6.2). In other words, people from almost all backgrounds are equally likely to anticipate the possible development of AI in the same way – regardless of their experience of computer science/programming, educational background, age or other characteristic.

When asked for the expectation of high-level machine intelligence within the next 20 years, those living in a high-SES postcode area are also more likely to think this is 'very likely'. For this time-frame, males and those who speak a language other than English at home are also more likely to think that high-level machine intelligence is very likely. In contrast, expectations for the next 50 years are homogeneous (see Figure 23).

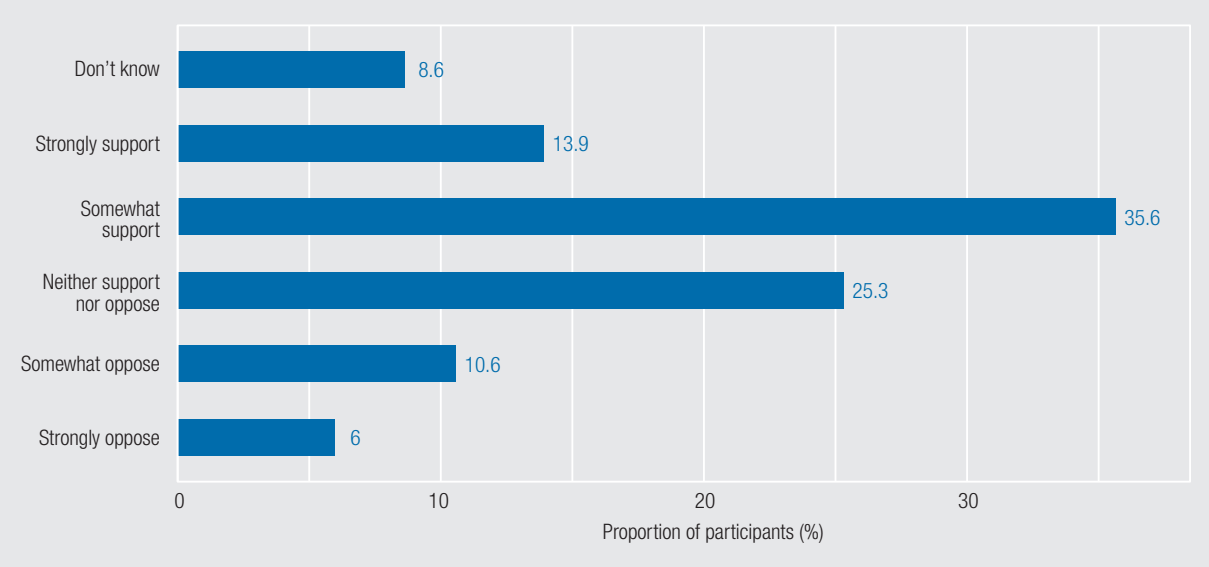
Figure 23. Proportion of people who think the development of high-level machine learning is 'very likely' in the next 10, 20 and 50 years by demographic characteristics.



NB. Survey asked: *"In your opinion, how likely is it that high-level machine intelligence will exist in 10 years, 20 years and 50 years?"*

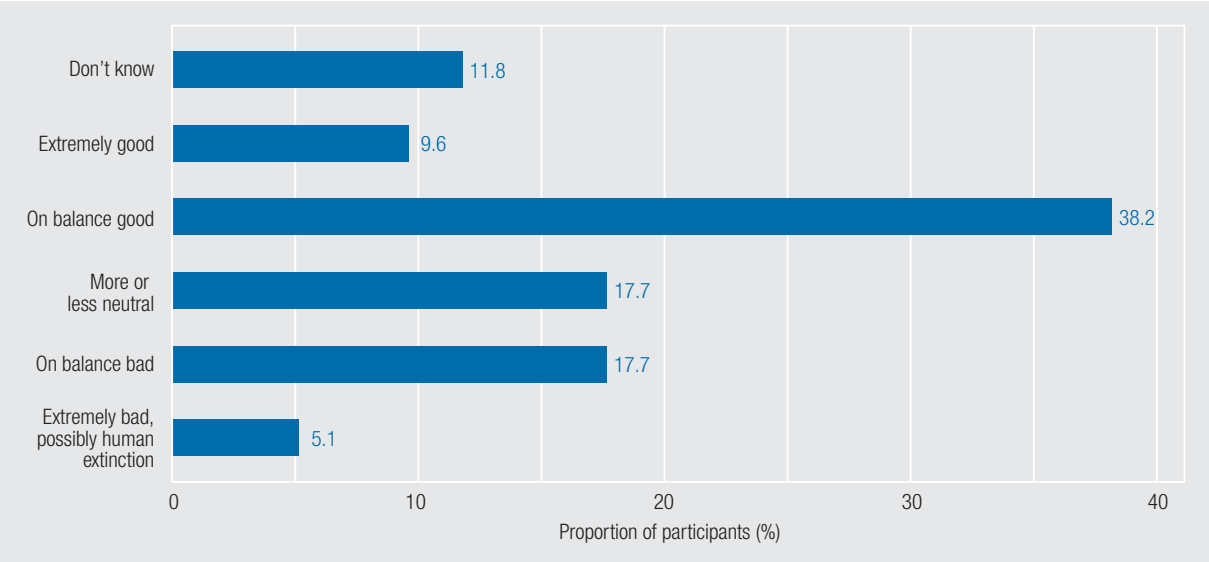
On the whole, the prospect of high-level machine intelligence was seen in fairly optimistic terms. Half of the respondents express support for the development of high-level machine intelligence (see Figure 24). Similarly, 47.8% expect it to have a 'good' impact on humanity, with 22.8% speculating that it would be 'bad' (see Figure 25). Comfortingly, perhaps, only 5.1% of respondents consider there to be a possibility of the development of high-level AI leading to human extinction.

Figure 24. Self-reported levels of support for high-level machine intelligence.



NB: Survey asked: *"How much do you support or oppose the development of high-level machine intelligence?"* Data are percentage of respondents.

Figure 25. Perceived impact of high-level machine intelligence.



NB. Survey asked: *"Suppose that high-level machine intelligence could be developed one day. How positive or negative do you expect the overall impact of high-level machine intelligence to be on humanity in the long run?"*

7. Public hopes and fears for AI and society

The final section of the survey asked respondents to reflect on all the topics and issues that they had considered while taking the survey, and write short statements describing their hopes and fears relating to AI and society.

The first of these questions – *"What aspects of AI do you feel most optimistic/hopeful about?"* – elicited 1998 responses (16,070 words in total). As can be seen in the word-count visualisation below (Figure 26), one of the most prevalent areas of AI development seen in optimistic terms related to advances in health and medicine (specifically terms relating to 'health', 'healthcare', 'medicine', 'medical', 'diseases'). For example:

> ***"If they can develop cures for diseases and anything to help people with grave illnesses"*** *[#633, F, 65+, NSW]*
>
> ***"AI in healthcare. Advanced treatment and diagnostics is possible with great use of AI in health"*** *[#2193, M, 30–34, VIC]*
>
> ***"New medicines for incurable diseases"*** *[#2183, F, 25–29, QLD]*
>
> ***"The ability to make safe medicines [and] know who is likely to become unwell"*** *[#136, F, 55–59, WA]*
>
> ***"In the health care space it could be an incredible assistance to doctors and nurses"*** *[#774, M, 25–29, VIC]*
>
> ***"Medicine and the ability to provide quicker and more appropriate solutions to issues such as pandemics, cancer, nutrition and rehab, to name a few"*** *[#1535, M, 45–49, VIC]*

Another prominent set of responses coalesced broadly around employment and matters relating to 'jobs', 'tasks' and 'work'. For example:

> ***"Making the harder jobs a bit easier"*** *[#1609, M, 18–24, NSW]*
>
> ***"AI can be really helpful in mining and all other risky jobs where humans can be harmful"*** *[#2224, M, 30–34, NSW]*
>
> ***"That it can reduce time in mundane jobs and free up humans for more complicated tasks that are urgently required"*** *[#1392, F, 30–34, VIC]*
>
> ***"I think that fact that machines can do work we don't have to, which could bring a better work/life balance. We could focus more on health and family instead of doing shit work"*** *[#1682, M, 18–24, WA]*

The ways in which these hopes and optimism was expressed included terms such as 'improve' (lives, society, everyday life, health), 'make', 'easier', 'good', 'ability', 'advancement' (in medicine, health and science), 'management' (of the environment, disasters, health, traffic) and 'automation' (tasks, transport and often unspecified forms).

> **"The Australian public is especially hopeful about the potential benefits of AI in healthcare and medicine for advanced diagnosis, development of medicine and disease treatments"**

Figure 26. Word count visualisation for optimistic/hopeful feelings about AI.

NB. Size of word relates to relative number of mentions in the corpus. Survey asked: *"What aspects of AI do you feel most optimistic/hopeful about?"*

The second of these questions – *"What aspects of AI do you feel most concerned/fearful about?"* – elicited 1997 responses (13,895 words in total). As can be seen in the second word-count visualisation (Figure 27), one of the most prevalent areas of concern related to privacy and surveillance issues (reflected in terms such as: 'privacy', 'surveillance', and 'monitoring'). For example:

> ***"The lack of privacy. It's not the technologies' fault, it's the people manipulating the technology. I fear that it has to be closely monitored or it will get out of hand very quickly!"*** *[#2575, M, 30–34, NSW]*
>
> ***"Misuse such as we are witnessing in China: to keep the whole population under surveillance, punish independent thought and to try ensure the population remains ignorant of the brutality and corruption and cruelty going on around them"*** *[#761, F, 50–54, QLD]*
>
> ***"Total surveillance. I'd rather kill myself than live in that kind of society"*** *[#880, M, 35–39, TAS]*

Another associated set of prominent responses coalesced broadly around direct references to governments, military, security, and war. For example:

> ***"That it'll be misused. That it'll be used by draconian governments to exert control over citizens"*** *[#3424, M, 35–39, QLD]*
>
> ***"Governments and private companies using AI to control or persecute populations of people"*** *[#2735, M, 25–29, NSW]*
>
> ***"Use by government to monitor and control citizens and by companies to manipulate people for economic purposes"*** *[#2304, M, 18–24, VIC]*
>
> ***"Cyber-attacks; network warfare; use by foreign military"*** *[#121, F, 65+, NSW]*
>
> ***"The 'rogue' military/armed forces/police/mercenary uses in the development of ruthlessly merciless killing capabilities..."*** *[#358, M, 65+, NSW]*

A third prominent set of responses centred broadly around employment matters relating to jobs, unemployment and redundancies. Indicative responses here include:

> ***"How reliable the AI will be, how industries will deal with this change and whether this would call redundancies"*** *[#2131, F, 18–24, NSW]*
>
> ***"Taking over people's jobs, people and their occupations becoming redundant as a result of AI"*** *[#458, F, 18–24, QLD]*
>
> ***"Potential to have a negative short-term impact on the job market, particularly for labour focused jobs. Those affected have low employment mobility, so could be a big issue"*** *[#2804, M, 25–29, WA]*
>
> ***"The fear of the next generation not having jobs"*** *[#192, F, 65+, WA]*
>
> ***"Skynet replacing jobs & not creating new avenues for humanity to work or explore"*** *[#2844, M, 40–44, VIC]*
>
> ***"The effect that it taking over jobs may have, and how sudden this may be for some industries. We have already seen content streaming put an end to most video/DVD rental stores... and e-reader books effectively closing many bookstores"*** *[#54, F, 45–49, QLD]*
>
> ***"OUR JOBS of course. But I always say. We still need humans to fix computers nowadays. But I think it may change in the future. I say computers will be able to fix themselves"*** *[#44, F, 50–54, NSW]*

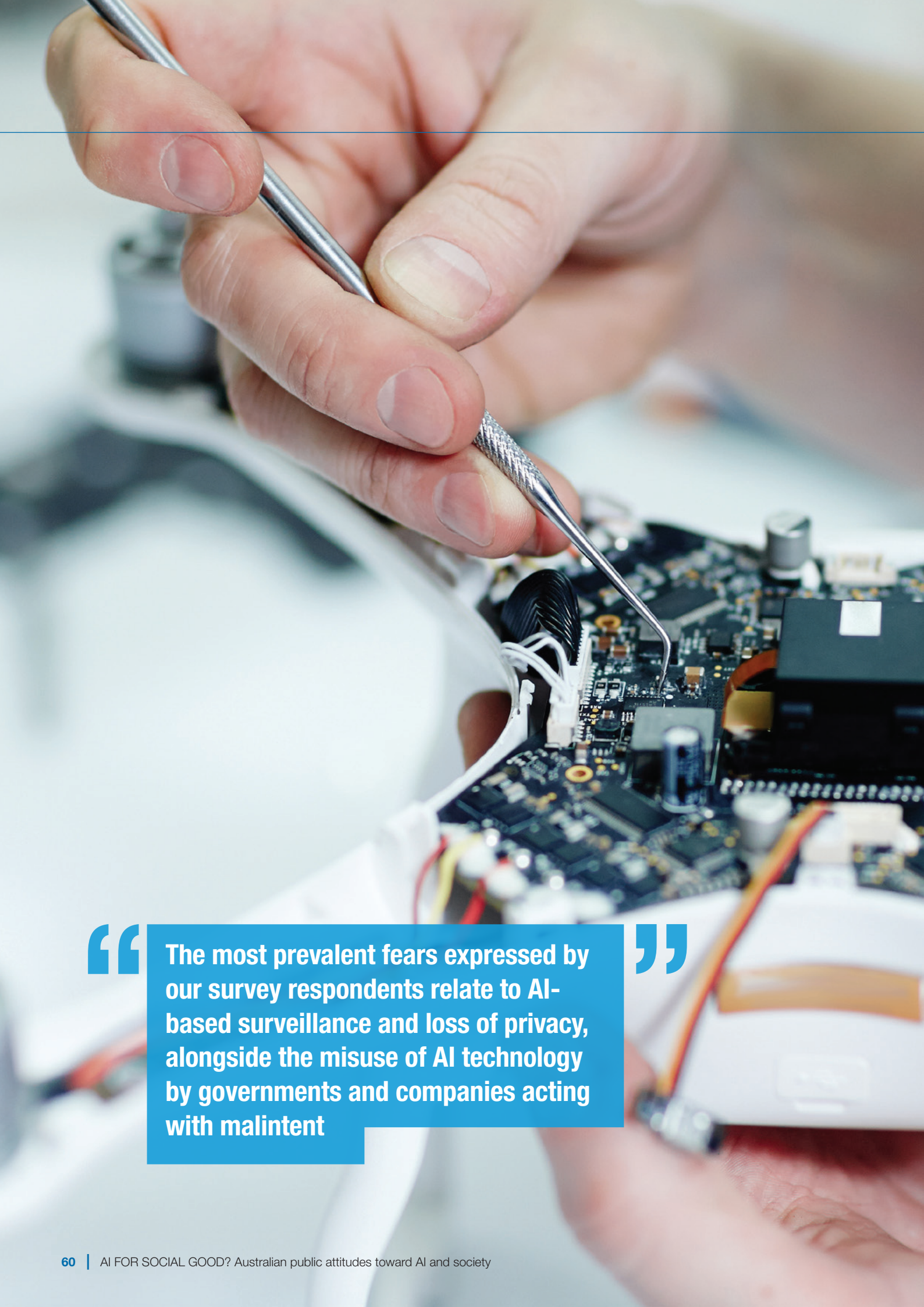
The ways in which these concerns and fears were expressed included terms such as 'loss' and 'losing' (primarily relating to 'jobs', but also 'control' and 'privacy'), 'wrong' (getting into the wrong hands, wrong uses, wrong people, things going wrong), 'misuse' and 'abuse' (by individuals, those in authority, governments, companies), and 'lack' (of …privacy, freedom, human intelligence, humanity, control).

Interestingly, mention of both 'Terminator' and 'Skynet' was prominent in these 'fearful' open-ended responses – echoing the small cluster of movie-related responses to the initial survey question of *"What's the first thing that comes to mind when you think about Artificial Intelligence?"* For example:

> ***"That it will literally take away from everything we are as humans. It will (already has) taken away jobs e.g. the self-serve registers in all supermarkets. So many jobs have been lost. And that powerful people will use it for wrong doing and to control the people. That one day I may become so advanced there is the possibility it could become like Terminator with machines believing they know better and are better than humans. Call me crazy ... but I completely disagree with this overall"*** *[#2268, F, 25–29, WA]*

Figure 27. Word count visualisation for concerns/fearful feelings about AI.

NB. Size of word relates to relative number of mentions in the corpus. Survey asked: *"What aspects of AI do you feel most concerned/fearful about?"*



> **The most prevalent fears expressed by our survey respondents relate to AI-based surveillance and loss of privacy, alongside the misuse of AI technology by governments and companies acting with malintent**

Discussion

This survey provides a relatively encouraging portrayal of opinions and attitudes amongst the Australian public relating to AI. While the majority of the Australian public considers themselves to have little knowledge and familiarity with the topic of AI, we find a considerable range of relatively sensible and well-grounded basic understandings of what AI is. Whereas some people's imaginations are still taken by vague ideas of robots 'taking over' and dystopic portrayals of AI in movies such as Terminator, there are also large numbers of immediate associations with computers being programmed to perform tasks, learn from data, and display human-like thinking. That said, there is still a tendency for people to equate AI with technology capable of approaching and/or exceeding human levels of intelligence – perhaps indicating a persistence of what Seymour Papert termed the 'superhuman human fallacy' (McCorduck 1979). Significantly, however, we find a willingness among many people to change their opinions once having received further information about AI and being asked to think through issues relating to AI and society.

All told, the Australian public retains a relatively optimistic outlook on the prospect of increased AI use across society. Support for AI development is expressed by nearly two-thirds of respondents, and there are consistently high levels of agreement with the idea of using AI to address social, humanitarian and environmental challenges (what might be termed 'AI for social good'). In some ways, then, these high levels of consensus could be seen as encouraging for those working in the area of AI development, management and governance. However, it is perhaps surprising to find consistently high levels of support across such a broad range of issues and applications – especially topics (such as the environment, justice and national security) that are usually contentious areas of general public and political debate. It is also perhaps surprising to find such strong consensus regarding the need for new legislation and industry guidelines. If these questions were asked about a topic other than AI then a wider range of opinions might well be evident.

These inconsistencies might point to a lack of distinction in the minds of many of our respondents – perhaps reflecting their low levels of direct experience with a diversity of AI technologies. In other words, whereas many people now recognise the general idea of 'AI' to be important, it is still perceived in rather homogeneous, non-specific terms. Thus, while respondents do have specific opinions regarding how they imagine AI might be used in areas such as healthcare and employment (relatively familiar domains to most adults), the possible uses of AI in areas such as public sector services or culture are less easily imagined. Moreover, the high proportion of people whose opinions shifted upon taking the survey, suggests that these views remain largely unformed, perhaps because of relatively low levels of knowledge about the technology and its consequences. The instability of viewpoint indicates the importance of greater education about, and discussion of, the issues raised by the technology.

Notwithstanding this lack of detailed understanding, these responses from the Australian public compare favourably to similar recent surveys elsewhere in the world. In particular, our survey partially replicated an earlier study of US public opinion (Zhang & Dafoe 2019). On one hand, we find notably higher levels of support for the development of AI amongst the Australian public (62.4%) than was reported for US public opinion (41%). We also find the Australian public to be generally more optimistic about the impact of AI on their lives and society than was the case with the US sample, as well as being more optimistic about the possible impact of the future development of high-level machine intelligence. Our survey also finds Australian public opinion to be less strongly patterned by factors such as household income, gender and educational attainment – all of which were key points of delineation in the US survey.

On the other hand, there are many similarities between our Australian survey and Zhang and Dafoe's US study. For example, three of the top four AI 'governance challenges' reported in the 2018 US public survey are replicated in our Australian findings – i.e. data privacy; fake/harmful content; and cyber-attacks. Moreover, the US public survey found similarly high levels of confidence in researchers, scientific organisations and the military to develop and manage AI in the best interests of the public. As we found to be the case in Australia, the US data also shows low levels of public approval for the involvement of Facebook in AI development and management.

Indeed, looking beyond the Zhang and Dafoe study, there are interesting points of congruence between our Australian findings and other recent studies elsewhere. For example, our survey chimes with findings that public perceptions of AI technology are susceptible to change through exposure to further information about the technology (Robb *et al.* 2020). Surveys in Europe also highlight public fears around privacy, job concerns and the rise of government and businesses acting with malintent (Lobera *et al.* 2020). Our findings of Australian public confidence in scientists and researchers to develop AI in the best interests of the public mirrors recent studies elsewhere. For example, Kerr and colleagues' (2020) research with the Irish public also notes a preference for academic and industry AI researchers to be in control of the trajectory of innovation, alongside the expectation that public authorities should be responsible for impact. As Kerr *et al.* (2020) point out, these public expectations of state responsibility for controlling AI impact are understandable, but complicated by the collaborative nature of AI research and development.

Other recent research also confirms some of the trends arising from our survey that might be considered surprising. For example, we found less evidence of having computer science/programming experience to influence people's opinions about AI. This chimes with other recent studies that suggest having a background in computing is not necessarily an indication of having different (or more complex) attitudes toward AI and society. For example, one recent sentiment analysis of 0.6 million tweets relating to AI found no discernible links between attitudes toward AI and working in computer-related occupations and/or having computer-related expertise (Wang 2019).

Within these broad conclusions, however, a number of specific points require further consideration. For example, the patterning of our survey findings along demographic lines was less pronounced than might be expected – especially in terms of the apparent lack of sustained differences highlighted in the US survey according to household income, socio-economic status, gender and level of educational attainment. That said, we did find some differences in terms of those respondents from younger age groups and/or with computer science/programming expertise. More surprising, perhaps, were the occasional differences in particular aspects of our questioning relating to specific differences – for example, involving respondents from households where a language other than English is spoken, reporting a disability, voting for a particular political party and so on. While it is possible to speculate on the reasons underpinning each of these specific differences, our data is limited in terms of providing robust indications of what broader issues they point to. As such, further research is required to continue to explore and clarify these possible patterns and associations. As one of the first studies of its type in Australia, our work is best seen as a starting-point for more detailed conversations about the future roles of that AI can play in Australian society.

> **Notwithstanding specific concerns, our survey finds the Australian public to be generally optimistic about the impact of AI on their lives and society. 69.4% of the public agrees that AI will do more 'social good' than 'social harm'**

Conclusions

This survey was conducted as a scoping study rather than an exercise in providing specific recommendations or agendas for further action. Nevertheless, our findings do raise a few issues that might now merit further consideration amongst policymakers and the broader Australian 'AI community'. For example, our findings show very strong public support for the establishment of a new regulatory body to address AI development, as well as support for the increased governance and oversight of AI development through legislation and industry guidelines. These are issues that clearly merit further consideration.

Similarly, our findings point to high levels of public trust in CSIRO, university researchers and bodies such as the Office of the Chief Scientist of Australia and Australian Human Rights Commissioner – therefore suggesting the benefits of these organisations playing prominent public-facing roles in any efforts to manage and oversee AI development in the future. In contrast, high-profile multi-national IT corporations such as Facebook and Amazon clearly need to work harder in order to gain confidence amongst the Australian public. For whatever reasons, trust in these two corporations was notably lower than for Microsoft and Google.

Yet perhaps the most immediate priority arising from our findings is the **development of public education efforts to enhance what might be termed 'public understandings of AI'**. Indeed, our survey suggests that many people will change their initial opinions and preconceptions about AI when provided with further information, examples and questions. This raises the likely benefits of initiating a sustained nationwide public education program to address 'public understandings of AI', 'AI literacy' and other 'data futures' related issues (see, for example, Long and Magerko 2020).

Public education around AI is certainly a topic of emerging interest elsewhere in the world (see, for example, Abbott *et al.* 2019, Balaram *et al.* 2018), and our survey findings suggest that there is a clear scope to develop such work in Australia. For example, we are seeing other countries begin to develop various forms of public engagement with AI – from online videos, campaigns through broadcast media and news media, through to exhibitions in museums and galleries, citizens juries and public workshops (see Adams & Burall 2019, Abbott *et al.* 2019, Balaram *et al*. 2018).

Existing models of how such activities might be developed and coordinated include the public-facing work of organisations such as the 'AI Now' Foundation in the US, and the 'Partnership on Artificial Intelligence to Benefit People and Society' (PAI). Developing similar Australian-specific efforts along these lines might benefit from involving organisations that are most trusted by the public – our survey suggests universities and CSIRO, as well as companies such as Google and Microsoft. This might point to the need to better promote national organisations such as 'Data61' and the '3AInstitute amongst non-expert public audiences in Australia.

Our survey suggests a number of issues that should be factored into any such public education work in Australia. For example:

- While there might be a desire amongst the AI community to promote public understandings of AI around ethical issues (such as so-called 'FAT' issues of fairness, accountability and transparency), first there is a need to develop basic public awareness of different types of AI – especially what might be termed 'narrow AI' and forms of AI already embedded into familiar technologies.
- A priority for any program of public education should be reaching the minority of public with little or no knowledge of AI at all. We found these people to be from older age groups, lower-income households and/or lower levels of education attainment. These are all demographic groups that have historically been amongst the most likely to directly experience adverse consequences of the deployment of automated technology in society.

- It makes sense for public education efforts to start by building on existing public support and interest in the use of AI in areas such as health and medicine, as well as environmental challenges and crisis response.
- Conversely, there is also a need to work on developing public understandings of the possible use of AI for in the area of culture and the arts. This seems to be a particular area where people are less able to conceive the possible benefits of AI use. Such efforts might be led by the Australian arts and cultural sector.

> **While not necessarily convinced that such technology will ever exist, most people imagine that AI capable of exceeding human intelligence would have an overall 'good' or 'more or less neutral' impact on humanity**

References

Adams, E. and Burall, S. (2019) *How to stimulate effective public engagement on the ethics of artificial intelligence.* London, Involve.

Azoulay, A. (2020). Towards an ethics of artificial intelligence. *UN Chronicle* www.un.org/en/chronicle/article/towards-ethics-artificial-intelligence.

Balaram, B. Greenham, T. and Leonard, J. (2018) *"Artificial intelligence: real public engagement"* London, RSA.

Cave, S., Coughlan, K. and Dihal, K. (2019). Scary robots: examining public responses to AI. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society.* ACM, 331–337.

Fast, E. and Horvitz, E. (2017) Long-term trends in the public perception of artificial intelligence. In: *31st AAAI conference on Artificial Intelligence,* San Francisco, California, USA, February 4–9.

Ipsos/MORI (2017) *Public views of machine learning.* London, Royal Society.

Kerr, A., Barry, M. and Kelleher, J. (2020). Expectations of artificial intelligence and the performativity of ethics. *Big Data & Society,* 7(1), 2053951720915939.

Lobera, J., Fernández Rodríguez, C. and Torres-Albero, C. (2020). Privacy, values and machines: predicting opposition to artificial intelligence. *Communication Studies,* 71(3):448-465.

Long, D. and Magerko, B. (2020, April). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16).

McCorduck, P. (1979). *Machines Who Think.* New York, Freeman.

McKinsey (2018) '*Applying artificial intelligence for social good*' November 28th, www.mckinsey.com/featured-insights/artificial-intelligence/applying-artificial-intelligence-for-social-good

New York Times (2020). I could solve most of your problems. *New York Times,* 2nd May.

Robb, D., Ahmad, M., Tiseo, C., Aracri, S., McConnell, A., Page, V. and Ardón Ramírez, P. (2020, March). Robots in the danger zone: exploring public perception through engagement. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 93–102).

Royal Society (2017). *Machine learning: the power and promise of computers that learn by example.* London, Royal Society.

Tomboulides, D., Dooney, J., Mcintyre, L. and Abbott, T. (2019). Strategies to inform the Swiss Public on artificial intelligence. https://digitalcommons.wpi.edu/iqp-all/5583/

Vinuesa, R., Azizpour, H., Leite, I. et al. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications* 11(233). https://doi.org/10.1038/s41467-019-14108-y

Wang, X. (2019, June). Using sentiment analysis for comparing attitudes between computer professionals and laypersons on the topic of artificial intelligence. In *Proceedings of the 2019 3rd International Conference on Natural Language Processing and Information Retrieval* (pp. 5–8).

Zhang, B. and Dafoe, A. (2019). *Artificial intelligence: American attitudes and trends.* Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3312874

Zhang, B. and Dafoe, A. (2020). US public opinion on the governance of artificial intelligence. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 187–193).

> **Only 5.1% of respondents see the possible future development of high-level machine intelligence as 'extremely bad' and/or maybe even 'leading to human extinction'**

Appendix

Table A1. Representativeness of survey sample in terms of adult population of Australia.

	Population	Survey percentage	Difference	Survey weighted percentage	Difference
Age range (years)					
18–24	12	14	-2.2	12	-0.1
25–29	10	11	-0.5	10	0.3
30–34	10	13	-3.2	10	0.4
35–39	9	12	-2.6	9	0.3
40–44	8	9	-0.7	8	-0.4
45–49	9	12	-2.8	9	0.4
50–54	8	3	5	8	0
55–59	8	4	3.9	8	0.1
60–64	7	5	2.1	7	0
65+	20	18	2	20	0.2
Gender					
Male	49	50.1	-1.1	49	-0.2
Female	51	49.9	1.1	51	0.2
State					
New South Wales (NSW)	32	31.7	0.3	30	1.7
Australian Capital Territory (ACT)	2	3.5	-1.5	3	-1.2
Victoria (VIC)	26	25.8	0.2	26	0.1
Queensland (QLD)	20	19.9	0.1	21	-0.9
Northern Territory (NT)	7	1.6	5.4	2	5.5
Western Australia (WA)	10	8.5	1.5	8	1.7
South Australia (SA)	2	7	-5	7	-5.4
Tasmania (TAS)	1	2.1	-1.1	3	-1.6
Socioeconomic status (derived from postcode)					
Low	32	30.8	1.2	30	1.6
Medium	41	40.1	0.9	42	-0.7
High	28	29.1	-1.1	28	0.1

Figure A1. Description of AI given to respondents from Section Two of the survey onwards

Artificial Intelligence (AI) refers to computer systems that perform tasks, make decisions and predictions that usually require human intelligence. AI can perform these tasks or make these decisions without explicit human instructions.

Today, AI has been used in the following applications:
[Respondents are shown five items randomly selected from the list below.]

- Translate over 100 different languages
- Predict one's Google searches
- Identify people from their photos
- Diagnose diseases like skin cancer and common illnesses
- Predict who are at risk of various diseases
- Help run factories and warehouses
- Block spam email
- Play computer games
- Help conduct legal case research
- Categorize photos and videos
- Detect plagiarism in essays
- Spot abusive messages on social media
- Predict what one is likely to buy online
- Predict what movies or TV shows one is likely to watch online

Figure A2. Comparison of the demographic characteristics of Language Background Other than English (LBOTE) and non LBOTE respondents.

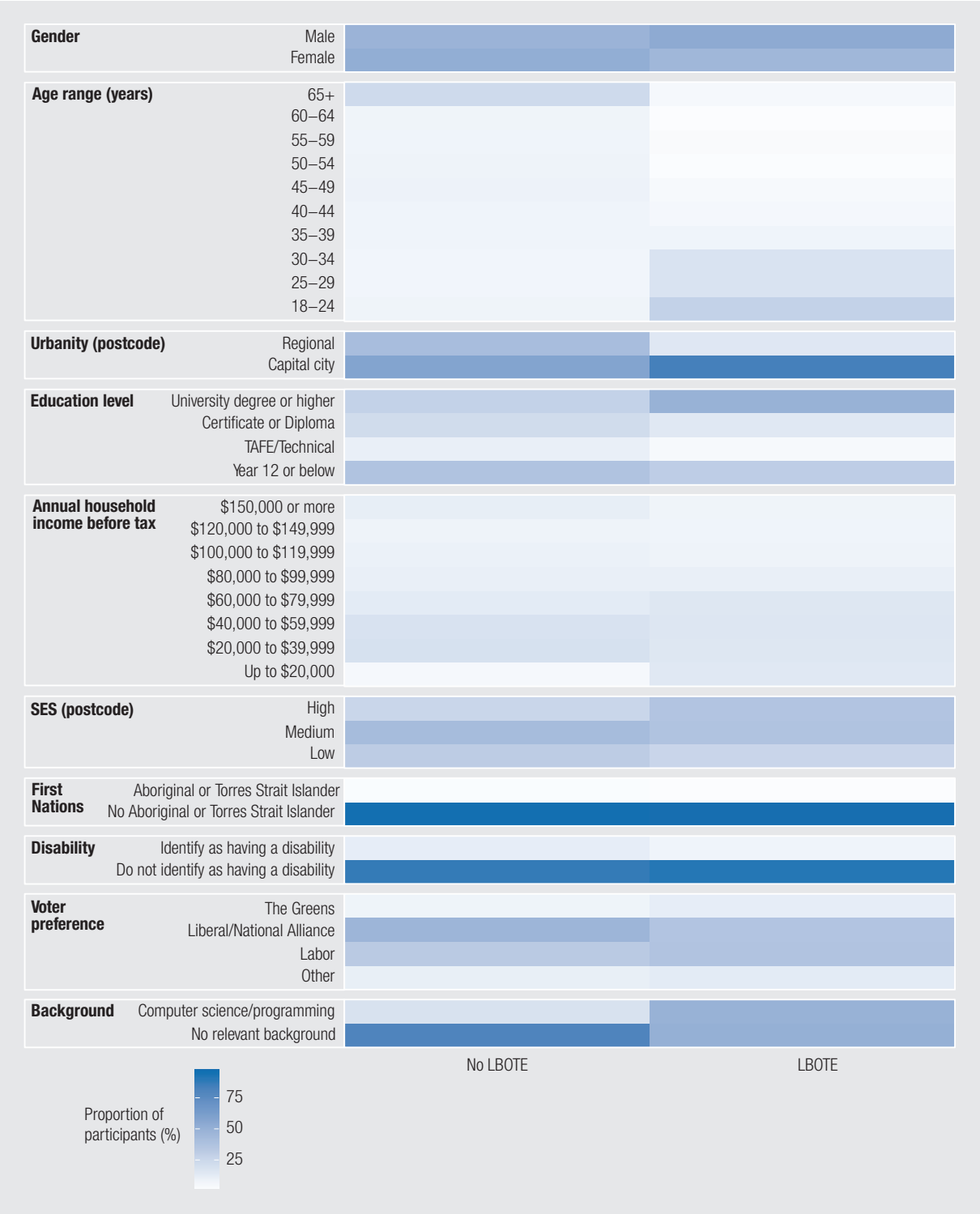
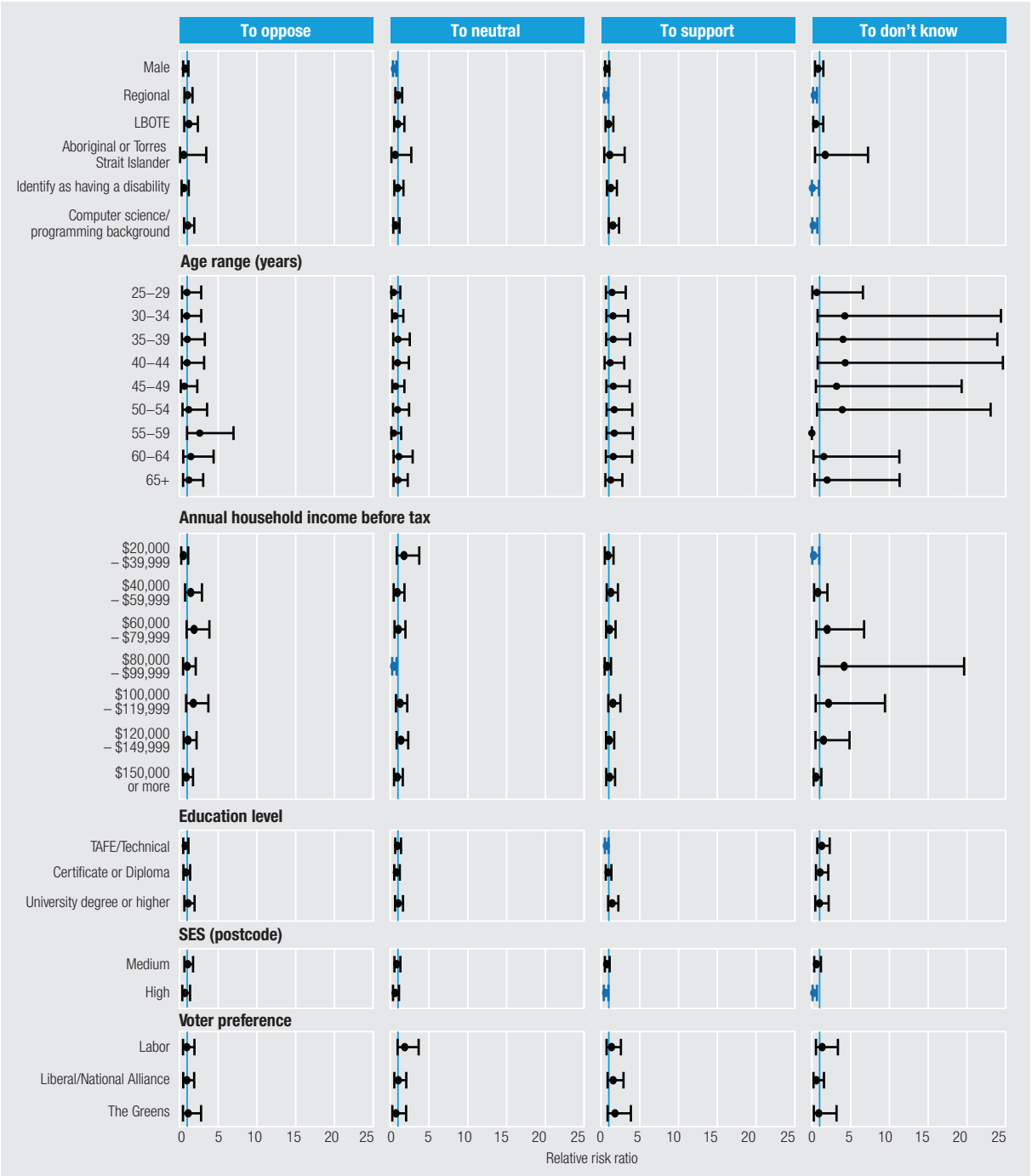


Table A2. Respondents' support or opposition to the development of AI – comparison of responses at the beginning, and the end of the survey – disaggregated.

	Strongly oppose	Somewhat oppose	Neither support nor oppose	Somewhat support	Strongly support	Don't know
Beginning of survey	3.1	7.4	23	40.5	21.8	4.1
End of survey	3.8	8.2	18.7	46.8	16.7	5.8

NB. Data are percentage of respondents expressing an opinion

Figure A3. Relative risk ratio and 95% confidence intervals of changing the opinion in each of these directions versus not changing opinion by respondents' demographic characteristics. Multinomial logistic estimation results.



Other notes from the multinomial regression model:

- Females were 46.8% more likely to change their mind to 'neutral' over not changing their mind at all than males.
- Those living in regional areas were 38.3% less likely to change their mind to 'support' over not changing their mind at all than those living in the capital cities.
- Those in TAFE/technical education are 30.9% less likely to change their mind to 'support' (over not changing their mind at all) than those with Year 12 or below educational attainment.
- Those living in high-SES postcode areas are 41.3% less likely to change their mind to 'support' (over not changing their mind at all) than those living in low-SES postcode areas.
- It is easier to find differences for those who change their minds to 'I don't know' at the end of the survey. In this case the following groups were X% less likely to change to 'I don't know' than to not change their minds at all (e.g. they do know, even when they change their minds):
  - Those living in regional areas (vs. those living in capital cities): 68.1%
  - Those identifying as having a disability (vs those who do not): 93.4%
  - Those with experience of programming and/or computer science/engineering at university (vs those without): 75%

Figure A4. Fit indicators for cluster analysis for the support of the use of AI in different areas.

NB. Survey item asked: *"How much do you support or oppose the use of AI in…?"*

Figure A5. Fit indicators for cluster analysis for the agreement with values related to the use of AI.

NB. Survey item asked: *"How much do you agree or disagree with the following statements? AI should…"*

Figure A6. Fit indicators for cluster analysis for the agreement with challenges related to the use of AI.

NB. Survey item asked: "*In the next 10 years, how important is it for tech companies and governments to carefully manage this challenge?"*

MONASH
University

MONASH
DATA
FUTURES
INSTITUTE

https://www.monash.edu/ai-data-science