



Trust In Artificial Intelligence

Australian insights
2020



October 2020

[KPMG.com.au](https://www.kpmg.com.au)



University of Queensland Researchers

Professor Nicole Gillespie, Dr Steve Lockey
and Dr Caitlin Curtis

KPMG Advisors

James Mabbott, Richard Boele, Ali Akbari,
Rossana Bianchi and Rita Fentener van Vlissingen.

Acknowledgements

We are grateful for the insightful input, expertise and feedback provided by members of the Trust, Ethics and Governance Alliance at The University of Queensland, particularly Dr Ida Someh, Associate Professor Martin Edwards and Professor Matthew Hornsey, KPMG Partners Phillip Sands, Scott Guse, Joel Di Chiara and Leo Zhang, as well as domain expert input by Greg Dober, Mike Richmond and Professor Monica Janda.

Citation

Lockey, S., Gillespie, N., & Curtis, C. (2020). Trust in Artificial Intelligence: Australian Insights. The University of Queensland and KPMG Australia. doi.org/10.14264/b32f129

Contents

Executive summary.....	2
Introduction.....	5
How we conducted the research.....	6
Do Australians trust AI?.....	8
Who do Australians trust to develop and regulate AI?.....	16
What expectations do Australians have about AI regulation?.....	22
What principles are important for Australians to trust AI systems?.....	28
How do Australians feel about AI at work?.....	32
How do Australians view key AI challenges?.....	36
How well do Australians understand AI?.....	40
What are the key drivers of trust and acceptance of AI?.....	46
Conclusion and implications.....	50
Appendix.....	52
Endnotes.....	53



Executive summary

Artificial Intelligence (AI) is the cornerstone technology of the Fourth Industrial Revolution and is enabling rapid innovation with many potential benefits for Australian society (e.g. enhanced healthcare diagnostics, transportation optimisation) and business (e.g. enhanced efficiency and competitiveness). The COVID-19 pandemic has accelerated the uptake of advanced technology, and investment in AI continues to grow exponentially¹.

AI also poses considerable risks and challenges to society which raises concerns about whether AI systems are worthy of trust. These concerns have been fuelled by high profile cases of AI use that were biased, discriminatory, manipulative, unlawful, or violated privacy or other human rights. Without public confidence that AI is being developed and used in an ethical and trustworthy manner, it will not be trusted and its full potential will not be realised². To echo the sentiment of Dr Alan Finkel AO, Australia's Chief Scientist, acceptance of AI rests on "the essential foundation of trust"³. Are we capable of extending our trust to AI?

This national survey is the first to take a deep dive into answering this question and understanding community trust and expectations in relation to AI. To do this, we surveyed a nationally representative sample of over 2,500 Australian citizens in June to July 2020.

Our findings provide important and timely research insights into the public's trust and attitudes towards AI and lay out a pathway for strengthening trust and acceptance of AI systems. Below, we summarise the key findings. In the conclusion to the report, we draw out the implications of these insights for government, business and NGOs.



Trust is central to the acceptance of AI, and is influenced by four key drivers

Our results confirm that trust strongly influences AI acceptance. There are four key drivers that influence citizens' trust in AI systems: 1) beliefs about the adequacy of current regulations and laws to make AI use safe, 2) the perceived uncertain impact of AI on society, 3) the perceived impact of AI on jobs, and 4) familiarity and understanding of AI.

Of these drivers, the perceived adequacy of current regulations and laws is clearly the strongest. This demonstrates the importance of developing adequate regulatory and legal mechanisms that people believe protect them from the risks associated with AI use.



Australians have low trust in AI systems but generally 'accept' or 'tolerate' AI

Trust in AI systems is low in Australia, with only one in three Australians reporting that they are willing to trust AI systems. Almost half of the public (45%) are unwilling to share their information or data with an AI system and two in five (40%) are unwilling to trust the output of an AI system (e.g. a recommendation or decision). Many Australians are not convinced AI systems are trustworthy. However, they are more likely to perceive AI systems as competent than designed to operate with integrity and humanity.

While many in the community are hesitant to trust AI systems, Australians generally accept (42%) or tolerate (28%) AI, but few approve (16%) or embrace (7%) AI.

Australians have the most confidence in Australian universities and research institutions, as well as defence organisations, to develop and use (51 – 55%) and regulate and govern AI (45 – 46%). In contrast, Australians have the least confidence in commercial organisations to do this. Only about a quarter (24%) have high or complete confidence in commercial organisations to develop and use AI, and less than one in five (19%) have high or complete confidence in them to regulate and govern AI. This may be due to the fact that most (76%) believe commercial organisations innovate with AI for financial gain, rather than for societal benefit.



Australians expect AI to be regulated and carefully managed

The large majority of Australians (96%) expect AI to be regulated, but most either disagree (45%) or are ambivalent (20%) that current regulations and laws are sufficient to make the use of AI safe. This powerfully highlights the importance of strengthening the regulatory and legal framework governing AI.

Most Australians (66%) expect government oversight of AI, with co-regulation between industry and government also a popular option (60%).

All Australians expect AI governance challenges to be carefully managed. The public view data challenges such as fake online content (70%), surveillance (69%), data privacy (69%), and cyber-attacks (67%) to be the most likely to impact large numbers of Australians in the near future. More than half also viewed disease misdiagnosis (56%), HR bias (56%), and technological unemployment (51%) as likely to impact Australian society.



Australians expect organisations to uphold the principles of trustworthy AI

The public has very clear expectations of the principles and related practices they expect organisations deploying AI systems to uphold in order to be trusted. Australians almost unanimously expect AI systems to meet high standards of:

- performance and accuracy
- data privacy, security and governance
- transparency and explainability
- accountability
- risk and impact mitigation
- fairness
- human oversight.

Most Australians (more than 70%) would be more willing to use AI systems if assurance mechanisms were in place, such as independent AI ethics reviews, AI ethics certifications, national standards for transparency, and AI codes of conduct. Organisations can directly build trust and consumer willingness to use AI systems by supporting and implementing these mechanisms.



Australians feel comfortable with some but not all uses of AI at work

Most Australians (65 – 79%) are comfortable with the use of AI at work for the purposes of monitoring organisational security, and task automation and augmentation. However, they are less comfortable with the use of AI for monitoring and evaluating employees, or in recruitment and selection.

Most Australians (59%) disagree that AI will create more jobs than it will eliminate. In the event that their jobs are automated, Australians clearly expect advanced notice (93%), retraining opportunities (92%), and redeployment (89%). This suggests that where AI has negative impacts, Australians expect support. Meeting these expectations will require strategic long-range workforce planning and investments in re-skilling by organisations and government.



Australians want to know more about AI but currently have low awareness and understanding of AI and its uses

Only 51% of the public have heard about AI in the past year, and most (61%) report a low understanding of AI, including how and when it is used in everyday applications. For example, even though 78% of Australians report using social media, 59% of them are unaware that social media apps use AI. The good news is that most Australians (86%) want to know more about AI. Considered together, the results suggest there is both a need for, and an appetite for, a public AI literacy program.

Introduction

Artificial Intelligence (AI) is an increasingly ubiquitous part of the everyday lives of Australians that is transforming the way we live and work⁴. AI is used in a range of applications, such as calculating the best travel route to take in real-time, predicting what customers will buy, identifying credit card fraud, helping diagnose disease, identifying people from photos, and enabling self-driving vehicles.

All sectors of the Australian economy are embracing AI. In the words of Klaus Schwab, Chairman of the World Economic Forum, we are entering a fourth industrial revolution characterised 'by a fusion of technologies that is blurring the lines between the physical, digital, and biological spheres'⁵.



What is AI?

Artificial Intelligence (AI) refers to computer systems that can perform tasks or make predictions, recommendations or decisions that usually require human intelligence. AI systems can perform these tasks and make these decisions based on objectives set by humans but without explicit human instructions.

The benefits and promise of AI for society and business are undeniable. AI is helping people make better predictions and informed decisions, is enabling innovation, productivity gains and improved efficiency, and lowering costs. It is helping protect physical and financial security (e.g. through fraud detection) and facilitating the current global fight against COVID-19.

The risks and challenges that AI poses for society are equally undeniable. These include the risk of codifying and reinforcing unfair biases, infringing on human rights such as privacy, spreading fake online content, technological unemployment and the dangers stemming from mass surveillance technologies, critical AI failures and autonomous weapons. These issues are causing public concern and raising questions about the trustworthiness and regulation of AI systems⁶.

The public's trust in AI technologies is vital for its continued acceptance. If AI systems do not prove to be worthy of trust, their widespread acceptance and adoption will be hindered, and the potential societal and economic benefits will not be fully realised.

Despite the central importance of trust for the widespread use and acceptance of AI in society, to date little is known about the Australian community's trust in AI or what influences it. Instead, current thinking has been informed by supposition, speculation and surveys in other jurisdictions.

This national survey is designed to understand and quantify Australians' trust in and support of AI, and to benchmark these attitudes over time. By taking the first deep dive into the question of trust, this research provides a comprehensive and nuanced understanding of Australians' overall trust in AI systems, as well as in specific AI applications in the domains of healthcare, policing,

HR and financial investment. These domains represent common applications of AI that relate to citizens, employees and consumers.

This research provides insights into the key drivers of trust, community expectations and confidence in the regulation of AI, expectations of the management of societal challenges associated with AI, as well as Australians' current understanding and awareness of AI. Importantly, the findings provide a clear understanding of the practices and principles Australians expect organisations to use to responsibly develop and ethically deploy AI in society and the workplace.

Collectively the research insights provide an evidence-based pathway for building and maintaining the trust and acceptance of AI systems by the Australian public. The insights are relevant for informing policy and practice across all three sectors of government, business and non-profits.

How we conducted the research

We used a research panel to collect the data. Research panels are commonly used in survey research to obtain a group of people that are representative of the Australian population on specific demographic attributes.


Our sample of 2,575 respondents was nationally representative on gender, age and state matched against Australian Bureau of Statistics (ABS) data, and broadly representative on income and downloading of the COVIDSafe app⁷. We collected data between 24th June and 21st July 2020.

Who completed the survey?

Sample

 2,575

Gender

 51%
Female

 49%
Male

Age Groups

13%
Generation Z
(18 – 23)

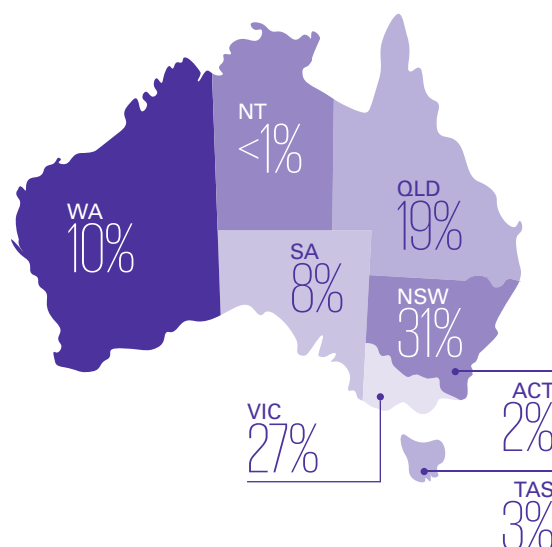
29%
Millennial
(24 – 39)

25%
Generation X
(40 – 55)

27%
Baby Boomer
(56 – 74)

6%
Silent Generation
(75 – 89)

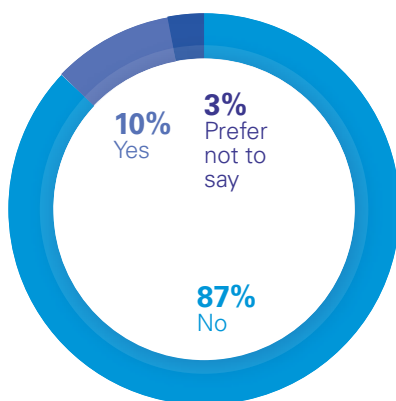
State



Area of Residence

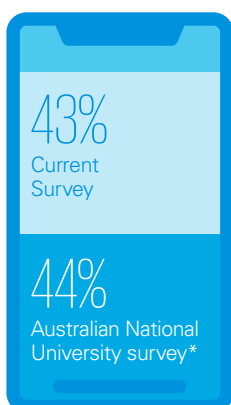


Identification as an ethnic minority group



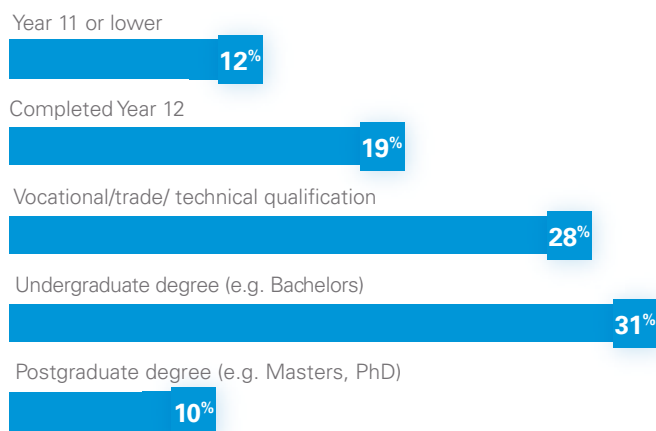
2% identify as Aboriginal or Torres Strait Islander

% Downloaded Covidsafe app



*ANU Survey conducted in May 2020 with a nationally representative sample of 3,249 Australians

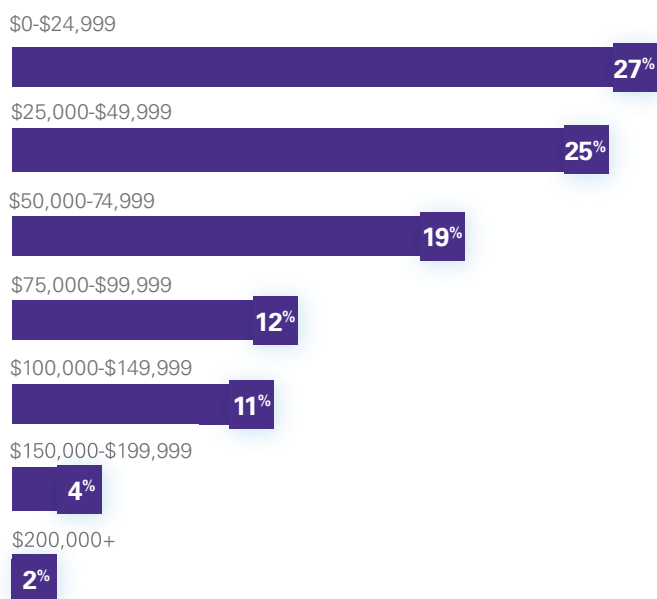
Education



13% have taken at least one university-level course in computer science


7% have a computer science or engineering degree

Income





Do Australians trust AI?



**To answer this question, we asked
Australians how much they trust, accept
and support AI in general, as well as
specific applications of AI.**

Overall, the results indicate the Australian public are ambivalent about trusting AI. In particular, the public question whether AI is designed to operate with integrity and humanity. Further, the public’s support for the development and use of AI systems depends

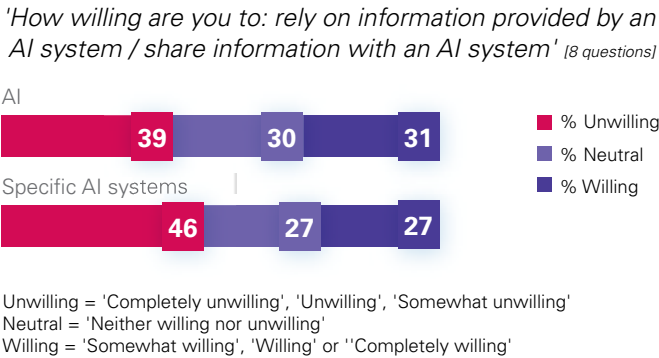
on the specific purpose of the AI system. The public is somewhat supportive of AI use in general. Of the specific AI applications we examined, the public is most supportive of its use in healthcare (e.g. disease diagnosis) and least supportive of its use in human resources.

Most Australians are unwilling or ambivalent about trusting AI systems

Many Australians are wary about trusting in AI systems. About two in five indicate they are unwilling to trust in AI systems in general and about a third report ambivalence. Only about a third of Australians report they are willing to trust AI systems.

Australians are slightly less trusting of specific AI applications (e.g. in healthcare, policing, human resources, or financial investment, see Appendix) than in AI systems in general, with 46% reporting an unwillingness to trust such specific AI applications and 27% reporting ambivalence in their trust (see Figure 1).

Figure 1. Willingness to trust AI systems in general and specific AI systems



Australians are slightly less willing to share their information with an AI system, than to rely on the output of an AI system.

We drilled down into two underlying components of trust: reliance and information sharing.

Reliance

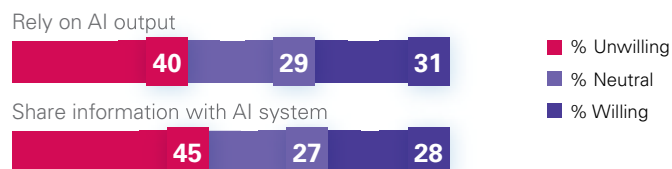
Assesses people’s willingness to rely on an AI system’s output, such as a recommendation or decision (i.e. to trust that it is accurate). If people are not willing to rely on AI system output, the system will not be used.

Information sharing

Relates to the willingness to share information or data with an AI system (i.e. to provide data to enable the system to work or perform a service for you). All AI systems are trained on large databases, but only some require the specific user to share information as input to function.

As shown in Figure 2, Australians are slightly less willing to share information with AI systems (45% unwilling, mean 3.5/7), than to rely on the output of AI systems (40% unwilling, mean 3.7/7). The same pattern emerged for trust in AI systems in general as for specific AI systems.

Figure 2. Willingness to rely on and share information with AI systems



Many Australians do not view AI systems as trustworthy. However, they are more likely to perceive AI as competent than to be designed to operate with integrity and humanity.

We assessed the key components of trustworthiness.



Ability

Relates to the perceived reliance, performance and accuracy of AI output.



Integrity and humanity

Relates to perceptions that the AI is developed based on sound ethical principles (e.g. fairness), is transparent about the data it collects and how it is used, and upholds the rights of users and societal interests.

As shown on the top half of Figure 3, when it comes to AI systems in general, most Australians disagree (32%) or are ambivalent (35%) about whether these systems operate with integrity and humanity. In contrast, most Australians either agree (56%) or are ambivalent (29%) about the ability of AI systems to produce reliable, accurate output.

For specific applications of AI (bottom of Figure 3), the difference between ability and integrity/humanity is less pronounced, with a majority of Australians either agreeing or feeling ambivalent about the ability of AI to produce accurate output (79%) and operate with integrity and humanity (71%). This most likely reflects the fact that specific applications of AI are more tangible and enable people to evaluate trustworthiness better. In contrast, for AI systems in general, people are more likely to be influenced by stereotypes about AI systems.

Figure 3. Perception of the ability, integrity and humanity of AI systems

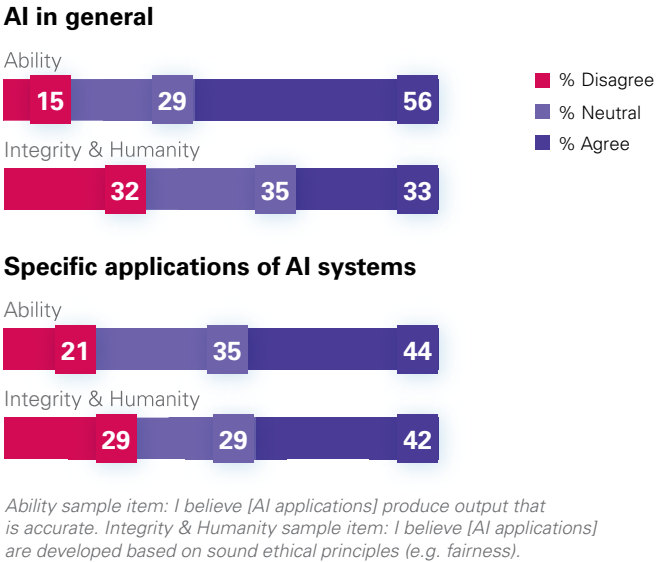
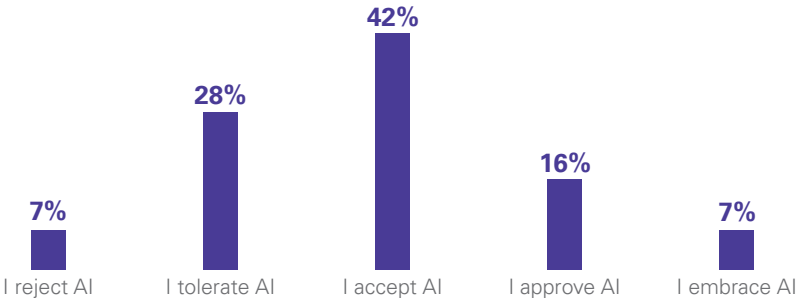


Figure 4. AI Acceptance

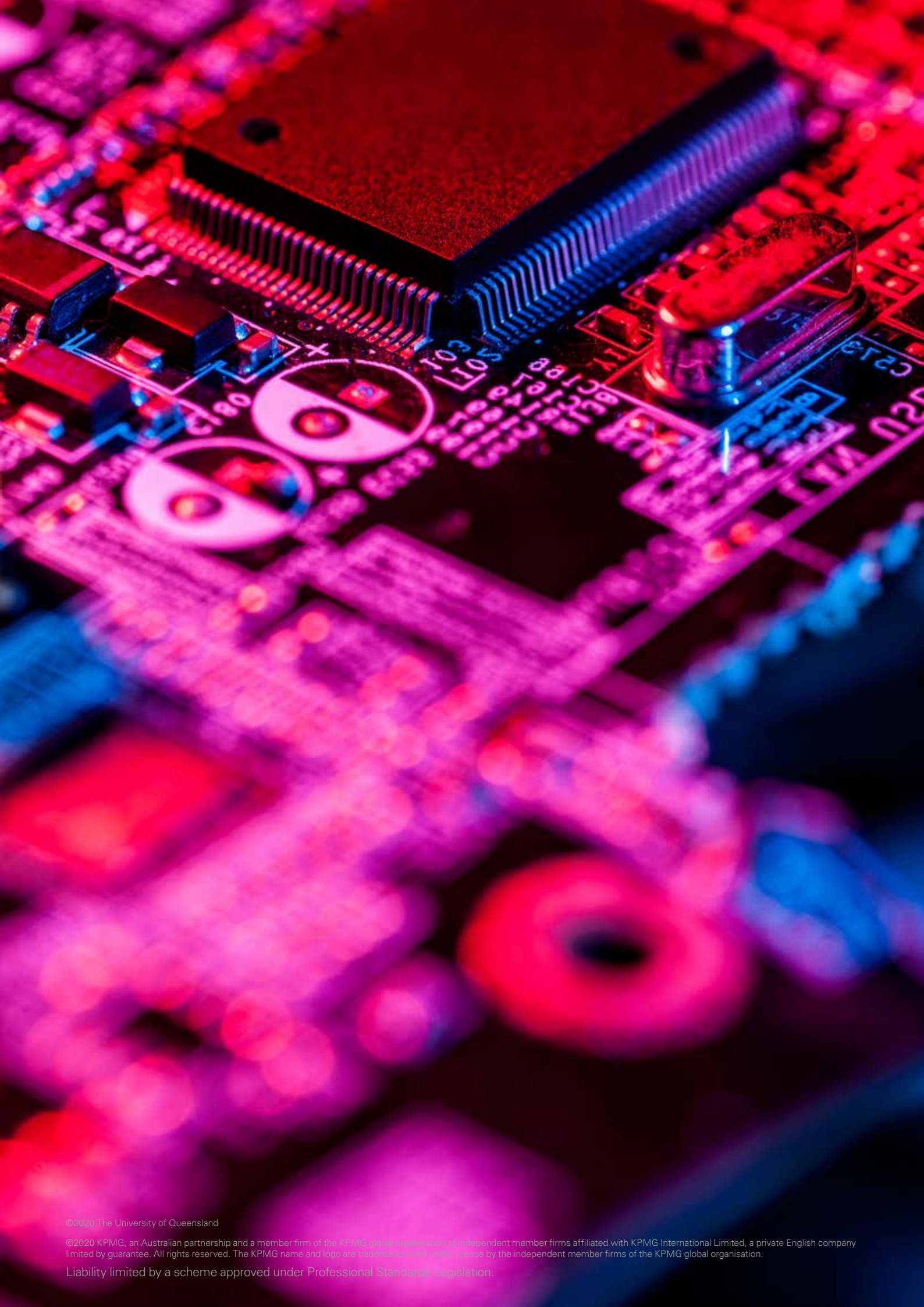
'In thinking about AI, which of the following best represents your view?'



Australians generally accept or tolerate AI, but few approve or embrace it

As shown in Figure 4, about two out of every five Australians 'accept' AI. However, about a third of Australians (35%) report they either 'tolerate' or 'reject' AI, and only about a quarter of Australians (23%) 'approve' of AI or 'embrace' it.

It is interesting to note that only a small proportion of Australians position themselves on the extreme poles of either 'rejecting' or 'embracing' AI.



©2020 The University of Queensland

©2020 KPMG, an Australian partnership and a member firm of the KPMG global organisation of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. The KPMG name and logo are trademarks used under license by the independent member firms of the KPMG global organisation.

Liability limited by a scheme approved under Professional Standards Legislation.

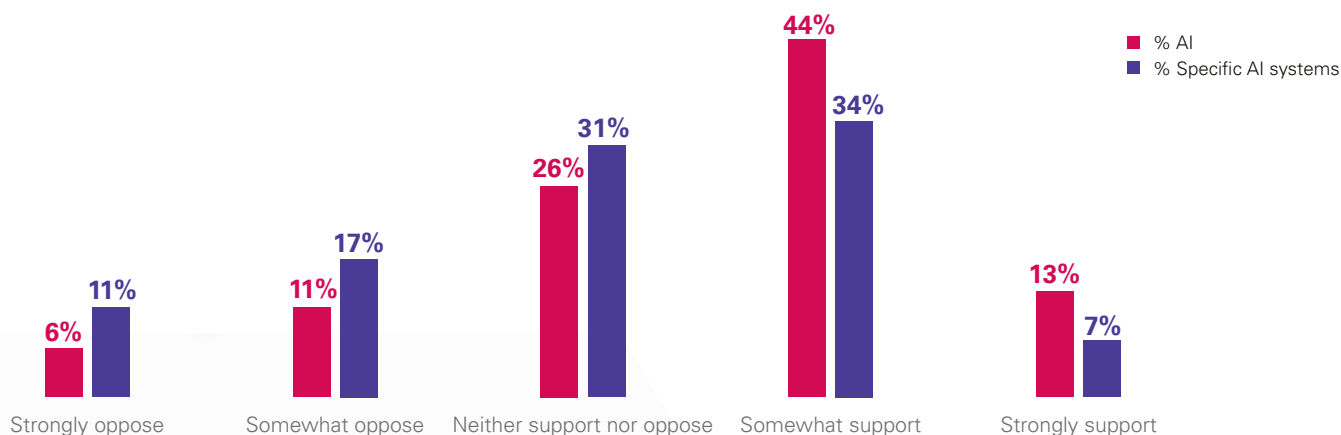
More Australians support than oppose the development and use of AI

As shown in Figure 5, when it comes to AI use in general, more Australians (57%) 'somewhat' or 'strongly' support the development and use of AI than 'somewhat' or 'strongly' oppose it (28%). However, a significant proportion (26%) are ambivalent about its development and use.

We note that fewer Australians (41%) are supportive of the development and use of specific applications of AI. When we examine public support across the four specific AI applications (see Appendix), we find Australians are more supportive of the development and use of AI for healthcare diagnosis and treatment (the most supported application at 49% support) compared to AI use for HR analytics (the least supported application with 34% support). This suggests that public support for AI varies significantly depending on its specific use application.

Figure 5. Support for AI and specific AI systems

'How much do you support or oppose the development and use of AI?'



Population segments vary in their acceptance, support and trust of AI systems

Younger people, notably Gen Z and Millennials, as well as those with computer science (CS) experience are more likely to trust, accept and support the development of AI (see Figure 6 and 7). However, even among these populations, trust is quite low.

Those with a university education report greater acceptance (see Figure 7), but do not differ substantively on other measures. There are no gender, income or regional differences.

Figure 6. AI support and trust by population segments

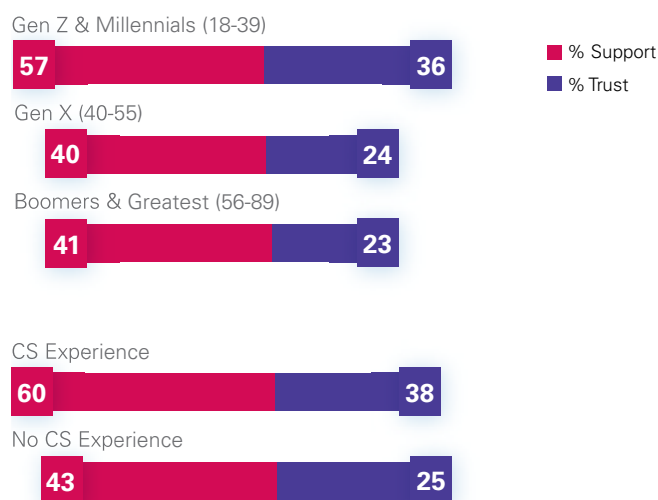
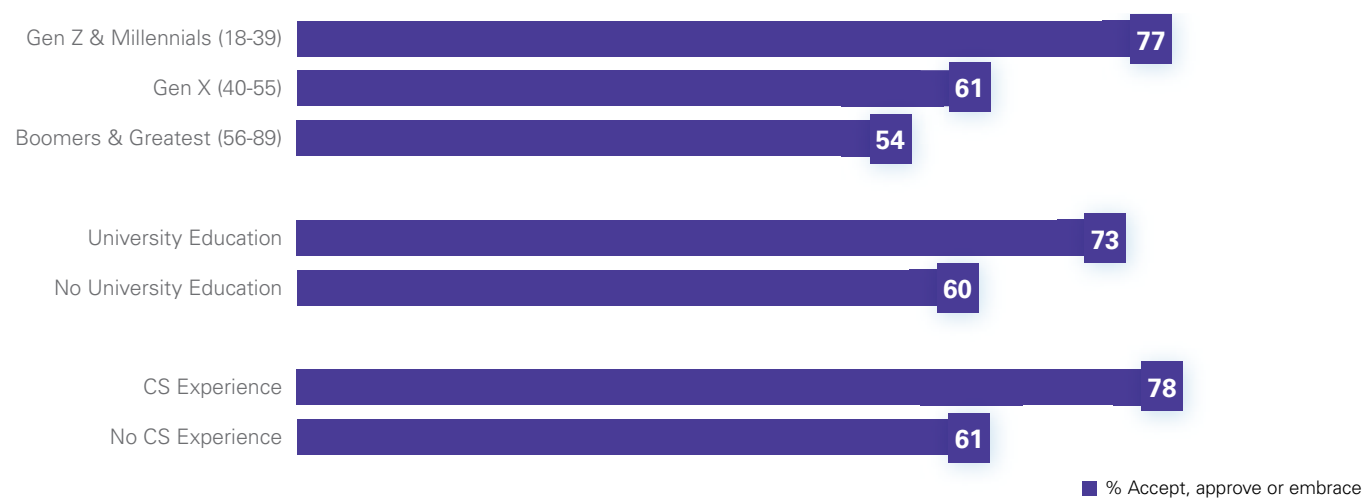



Figure 7. AI acceptance by population segments





Who do Australians trust to develop and regulate AI?



**To answer this question, we asked
Australians how much confidence they
have in different entities to develop and
use AI, as well as regulate and govern AI.**

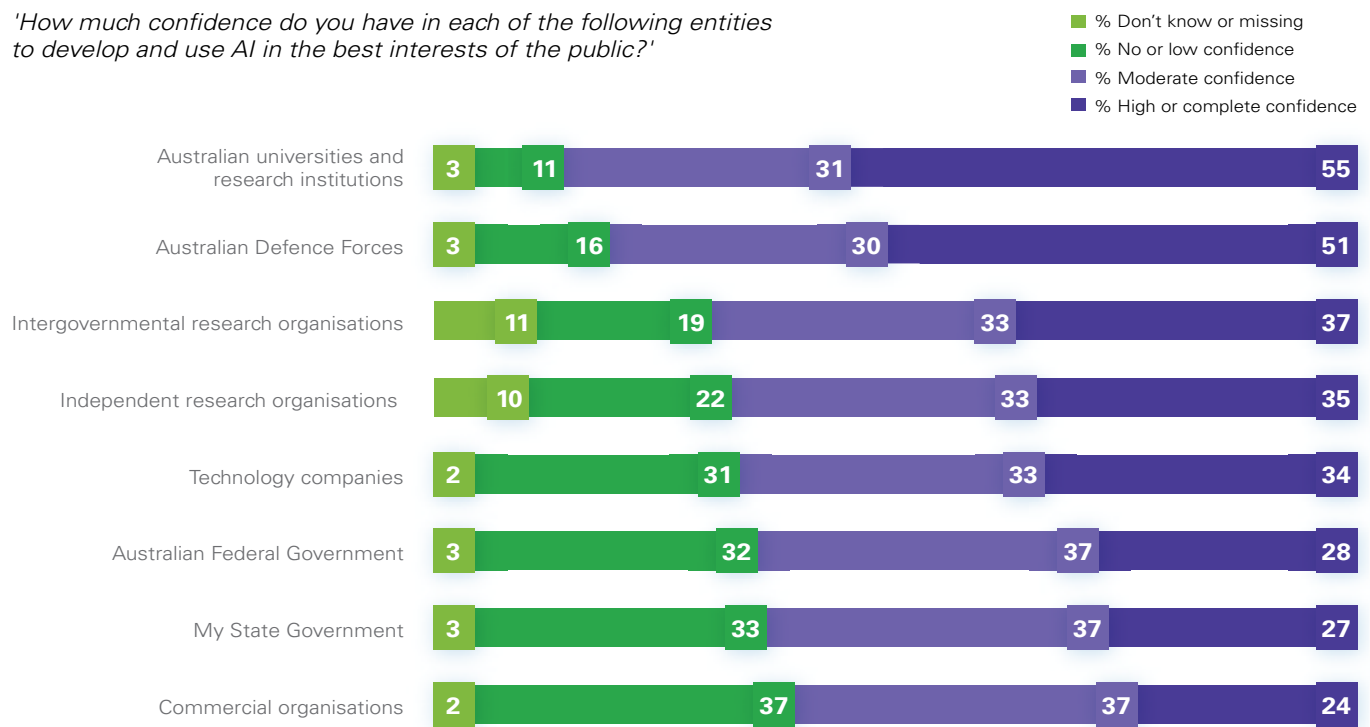
Australians are most confident in Australian research and defence organisations to develop and use, and regulate and govern AI, and least confident in commercial organisations.

The majority of Australians (51 – 55%) have high or complete confidence in Australian universities and research institutions and the Australian Defence Forces to develop and use AI in the best interests of the public. In comparison, about a third of Australians have high or complete confidence in technology companies, and a little over a quarter in federal and state governments. Australians have the least confidence in commercial organisations (see Figure 8).

It is noteworthy that around a third of Australians report no or little confidence in government (federal and state governments), technology companies and commercial organisations. The lack of confidence in technology companies and commercial organisations is striking given that the majority of the population's experience of AI is with applications developed and used by such organisations⁸. A solution may be for commercial and technology AI companies and government to collaborate with more trusted entities, such as universities and research institutions.

Figure 8. Confidence in entities to develop and use AI

'How much confidence do you have in each of the following entities to develop and use AI in the best interests of the public?'



Australians show a similar pattern regarding confidence in entities to regulate and govern AI in the best interest of the public (Figure 9). Yet Australians generally show less confidence in institutions to regulate and govern AI than to develop and use it.

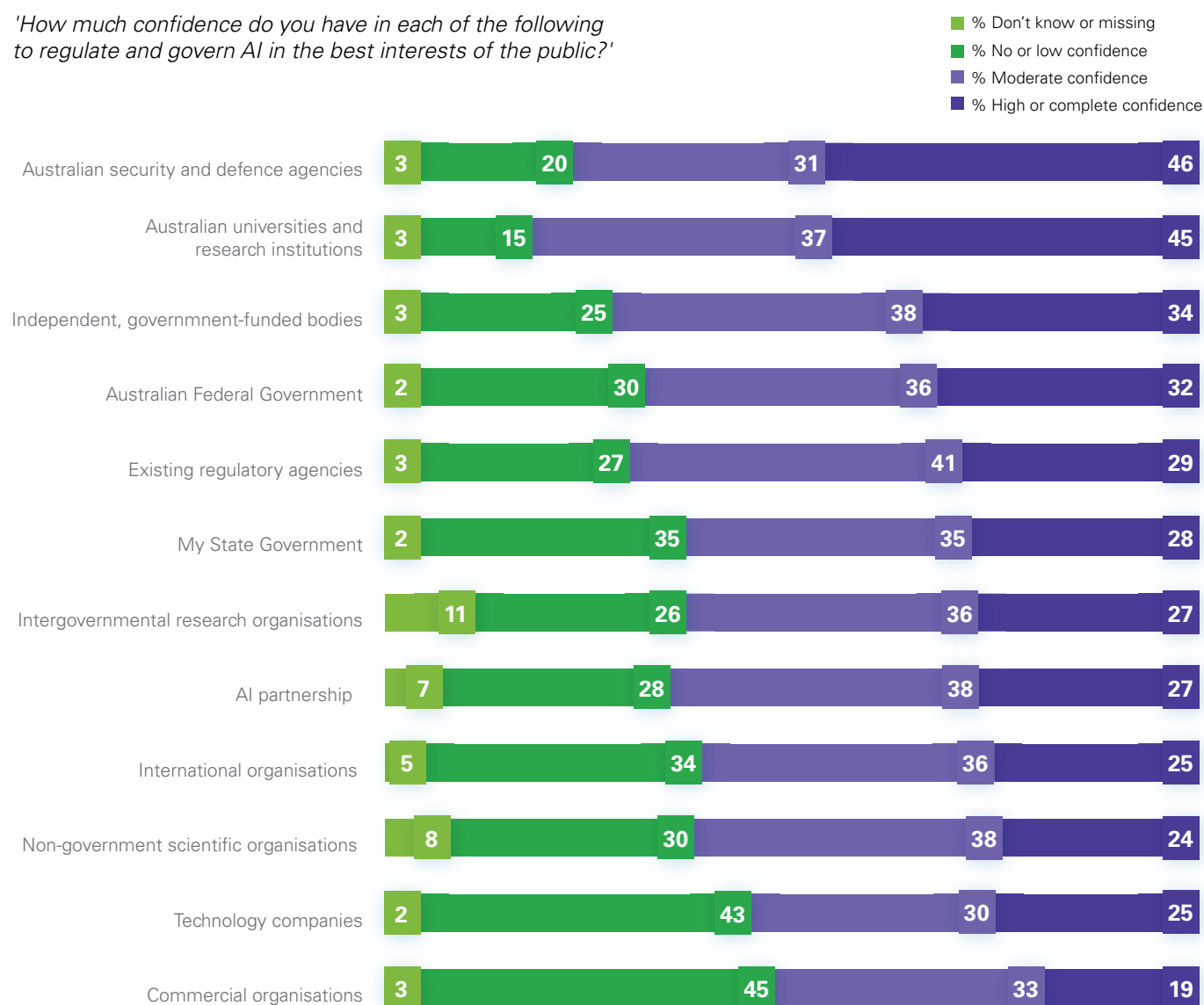
Australians report higher confidence in Australian security and defence agencies and Australian university and research institutions to regulate and govern AI than any other entity.

However, a majority also express moderate or high confidence in governments and government-funded bodies, regulators, international organisations and scientific organisations to regulate and govern AI.

Conversely, a little over two in five Australians (43 – 45%) have no or low confidence in technology companies or commercial organisations to regulate and govern AI.

Figure 9. Confidence in entities to regulate and govern AI

'How much confidence do you have in each of the following to regulate and govern AI in the best interests of the public?'

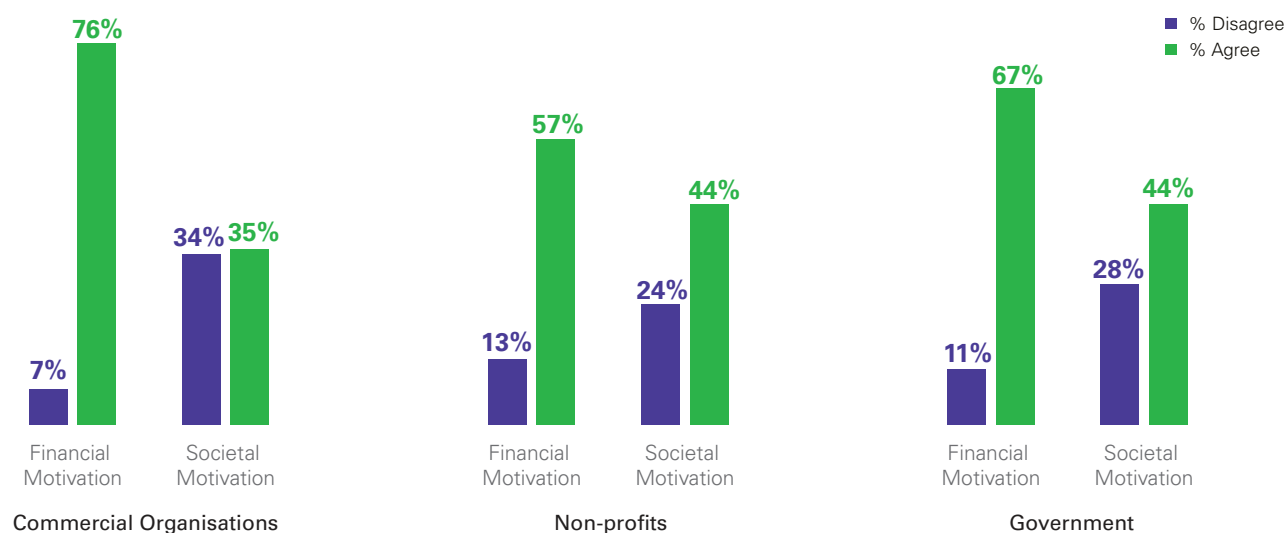


People believe organisations innovate with AI mostly for financial reasons

One reason for the lack of confidence in commercial organisations to develop and regulate AI may be that people think such organisations are motivated to innovate with AI primarily to cut labour costs and increase revenue (financial motivation) rather than to help solve societal problems and enhance societal wellbeing (societal motivation). As shown in Figure 10, about three quarters (76%) of the public agree commercial organisations innovate with AI for financial gain, whereas only a third (35%) agree they innovate for societal benefit.

While most Australians also agree government (67%) and non-profits (57%) innovate with AI for financial reasons, 44% of the public believe these organisations also innovate with AI for societal benefit (see Figure 10).

Figure 10. Motivation to innovate with AI

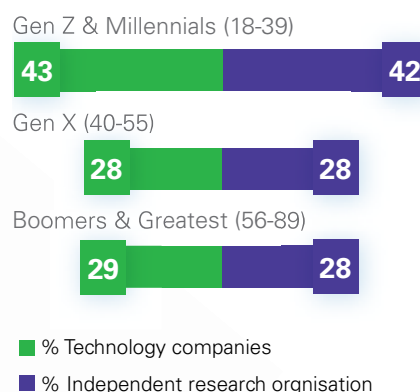


Younger people are more confident in tech companies to develop and use AI

Younger respondents are more confident in technology companies and independent research organisations to develop and use AI in the best interests of the public (see Figure 11).

Figure 11. Age influences confidence in entities to develop and use AI

% High or complete confidence





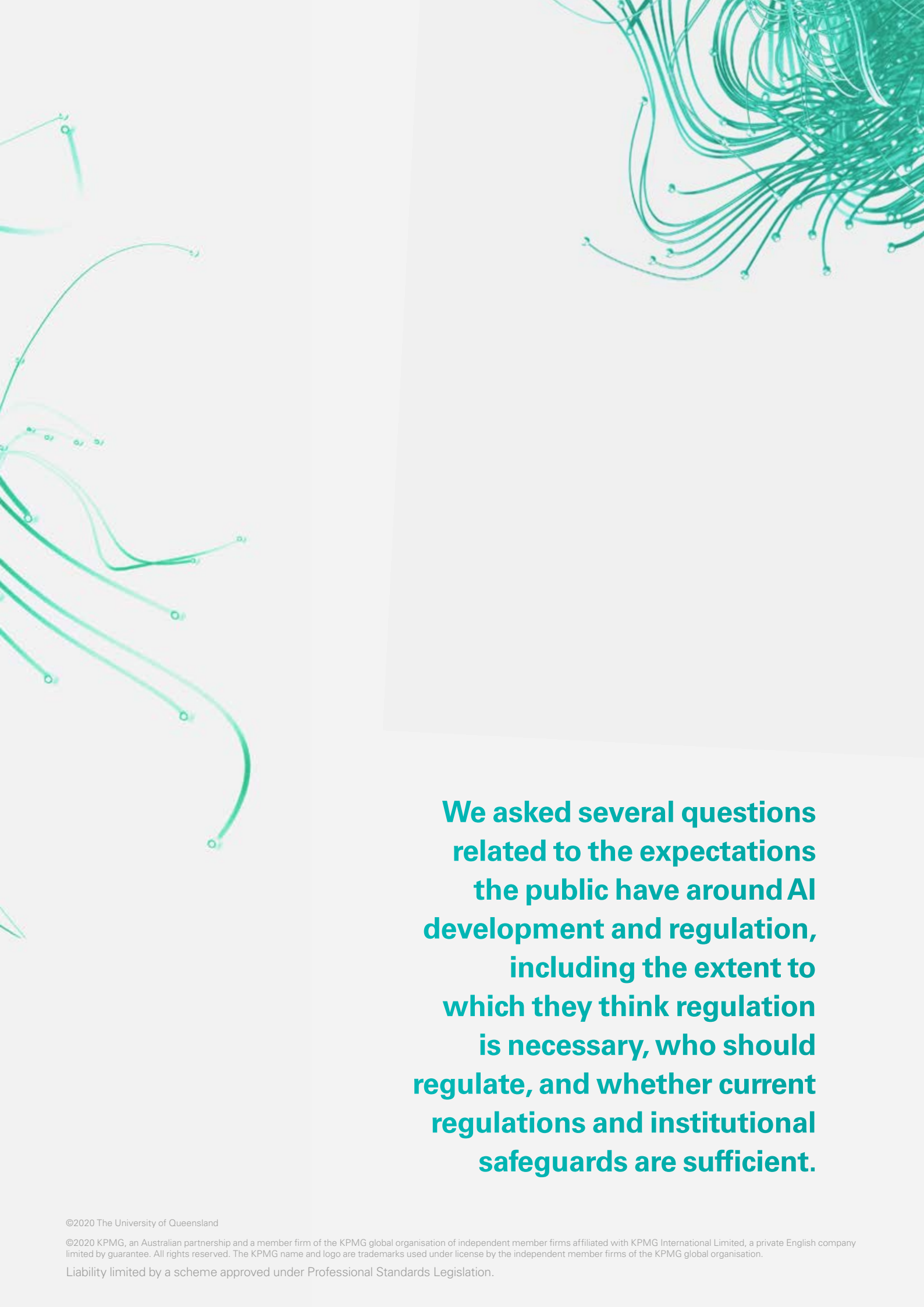
©2020 The University of Queensland

©2020 KPMG, an Australian partnership and a member firm of the KPMG global organisation of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. The KPMG name and logo are trademarks used under license by the independent member firms of the KPMG global organisation.

Liability limited by a scheme approved under Professional Standards Legislation.

The background of the slide is a complex, abstract pattern of teal-colored lines and dots. The lines are of varying thickness and are mostly curved, creating a sense of movement and connectivity. Small dots are scattered throughout, often at the ends of the lines, resembling a network or a data visualization. The overall color palette is a range of teal and green tones.

What expectations do Australians have about AI regulation?

An abstract graphic composed of numerous teal-colored lines of varying thicknesses. Some lines are straight, while others are curved or form dense, tangled clusters. The lines are scattered across the page, with a higher concentration in the top right corner and a few lines extending towards the bottom left. The overall effect is a modern, digital, and somewhat chaotic aesthetic.

**We asked several questions
related to the expectations
the public have around AI
development and regulation,
including the extent to
which they think regulation
is necessary, who should
regulate, and whether current
regulations and institutional
safeguards are sufficient.**

Regulation is clearly required and Australians expect external, independent oversight

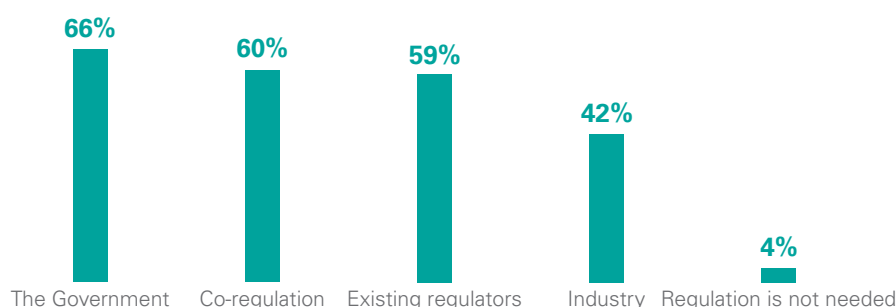
The vast majority of Australians think AI regulation is needed. This finding corroborates those of previous European⁹ and US¹⁰ surveys that also demonstrate a strong desire for regulation.

A majority of respondents (66%) think the government should regulate AI. Co-regulation between industry and government, and regulation by existing regulatory bodies are also popular options. Fewer Australians (42%) think industry self-regulation is desirable compared to the other forms of external or co-regulation (see Figure 12).

Despite the desire for government regulation, it is noteworthy that confidence in the federal government to regulate AI is not uniformly high. As discussed in the previous section, around a third of the public have high confidence, yet a similar proportion have no or low confidence. As such, while most Australians want government involvement in AI regulation, they are not all confident in the government's ability to regulate.

Figure 12. Who should regulate AI?

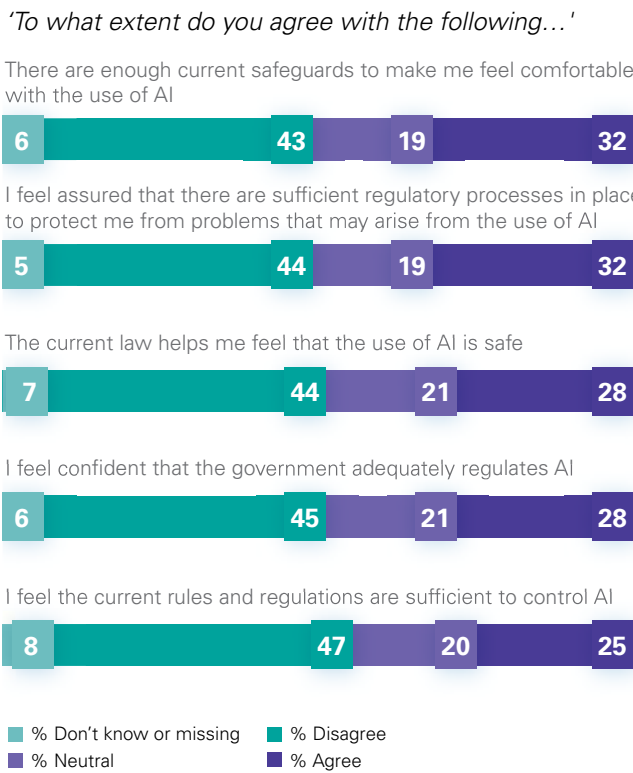
'AI should be regulated by [tick all that apply]...'



Current regulations are insufficient given the uncertainty around AI

Australians generally disagree (43 – 47%) or are ambivalent (19 – 21%) that the current safeguards around AI (rules, regulations and laws) are sufficient to make the use of AI safe or protect them from problems (see Figure 13). Similarly, the majority either disagree or are ambivalent that the government adequately regulates AI.

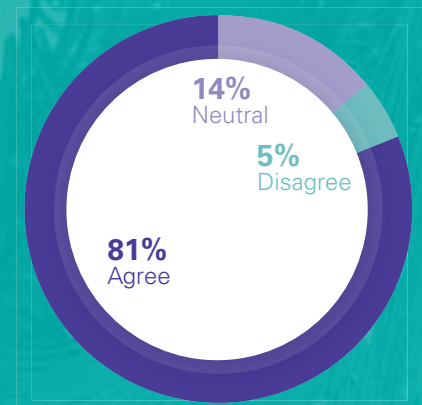
Figure 13. Perception of current regulations, laws and rules to make AI use safe.



Relatedly, the large majority (81%) believe the impact of AI is uncertain and unpredictable (see Figure 14). There are no substantive differences in either of these measures between AI in general and specific applications.

Figure 14. Perceptions of AI uncertainty

Uncertainty – 'To what extent do you agree with the following: There are many unknowns about AI [sample item]'



Australians generally believe the benefits of AI outweigh the risks, but this varies across specific AI systems

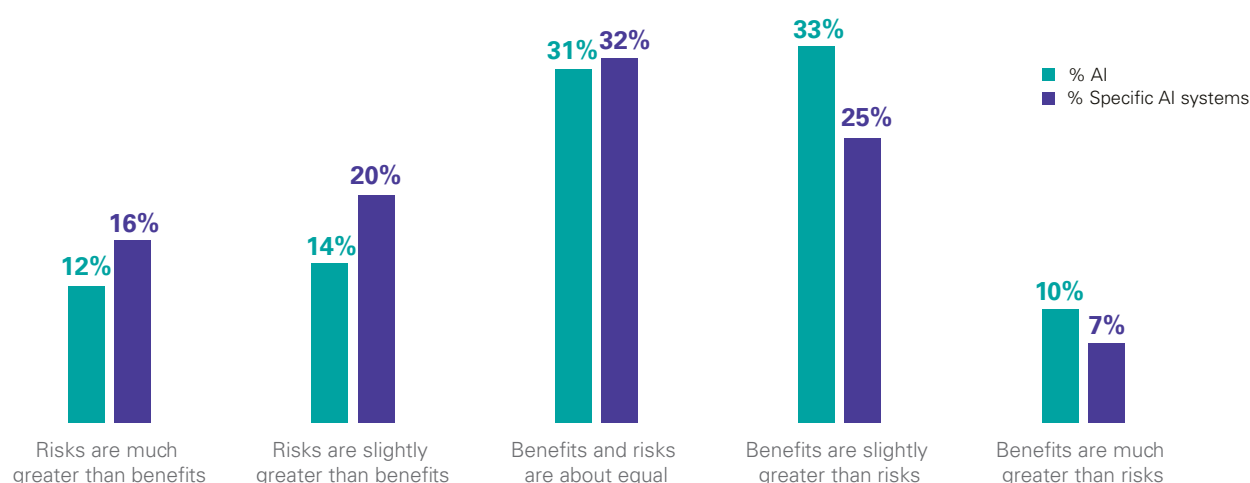
While Australians are not necessarily confident in current safeguards and believe that there is a high degree of uncertainty surrounding AI, they still generally perceive the benefits to outweigh the risks.

Figure 15 demonstrates that 43% of Australians perceive the benefits of AI outweigh the risks, 26% believe that the risks outweigh the benefits, and around a third believe that the benefits and risks of AI are about equal. Of those who believe the benefits outweigh the risks, more believe that the benefits are slightly greater (33%) than much greater (10%).

The pattern is somewhat different for specific AI systems, however. A similar proportion of Australians perceive the risks to outweigh the benefits (36%) than the benefits to outweigh the risks (32%). There is a significant difference between perceptions of risks and benefits of AI for healthcare diagnosis and treatment (42% perceive the benefits to outweigh the risks) compared to AI use for HR analytics (42% perceive the risks to outweigh the benefits).

Figure 15. Perceived balance of risks and benefits of AI and specific AI systems

'Overall, which best represents your view on the benefits and risks of AI systems?'

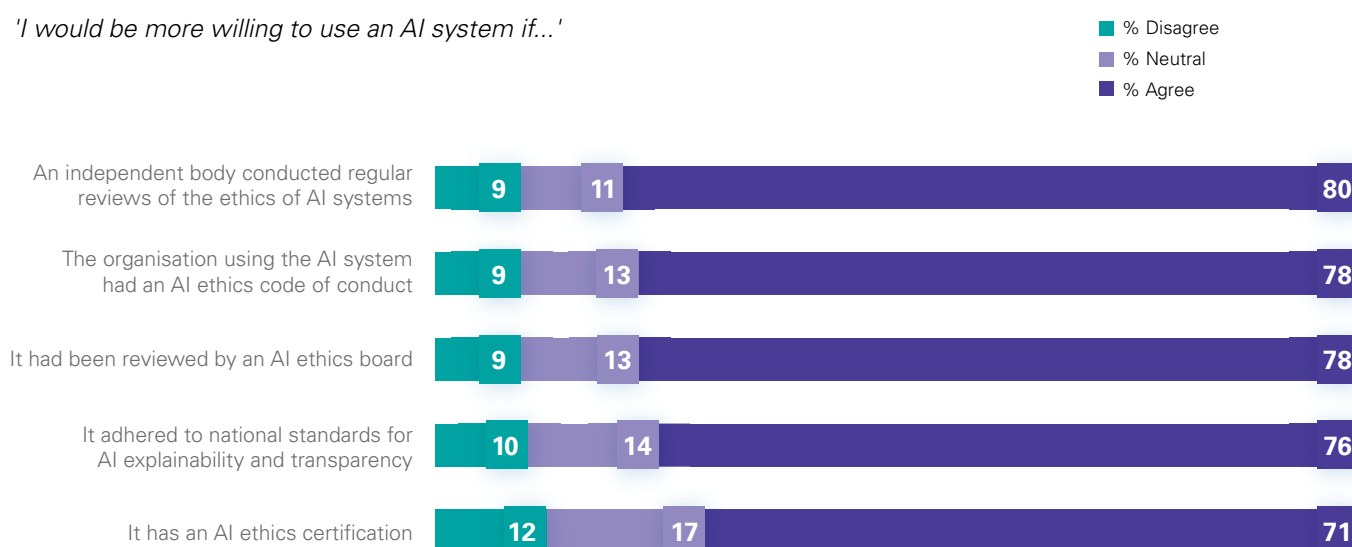


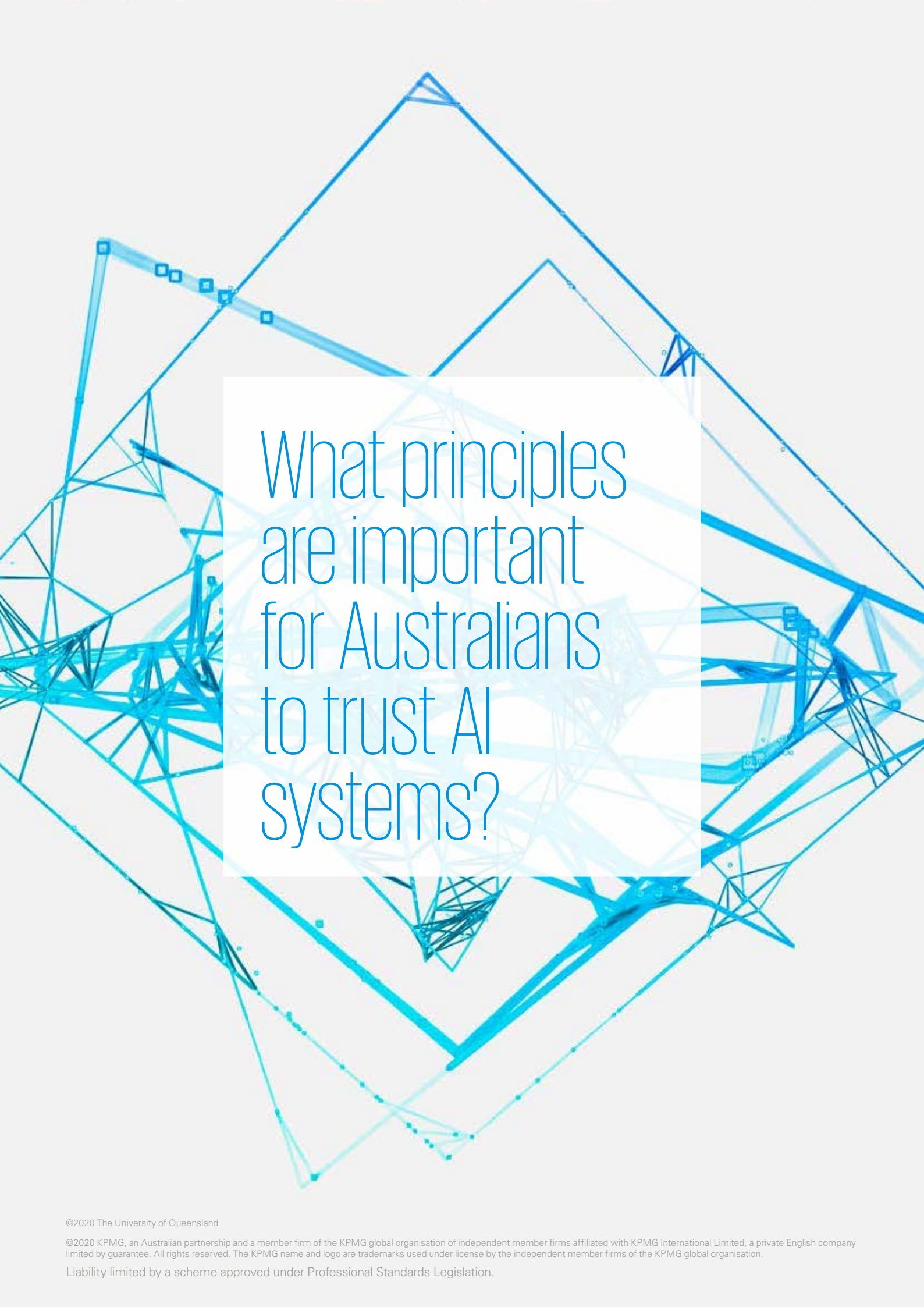
Assurance mechanisms enhance trust in AI systems

The vast majority of Australians (71-80%) agree that assurance mechanisms that support the ethics of AI systems would make them more willing to use AI systems. These mechanisms include independent checks, codes of conduct, national standards on AI explainability and transparency, and AI ethics certification (see Figure 16). The introductions of such mechanisms are likely to increase perceptions of safeguards and reduce uncertainty.

Figure 16. AI Assurance mechanisms

'I would be more willing to use an AI system if...'





What principles
are important
for Australians
to trust AI
systems?



**Eight AI design and governance
principles and practices are highly
important for trust.**

A proliferation of reports and guidance documents on the development and deployment of trustworthy AI have been produced over the past few years¹¹, with public sector institutions and private organisations producing policy documents outlining key practices and principles.

One goal of this survey was to determine what practices and principles are important for Australians to trust in AI. To answer this question, we asked about the importance of 40 practices associated with the eight principles for trustworthy AI shown below. These principles were adapted primarily from the 2019 European Commission Principles for Trustworthy AI¹², as well as the Australian AI ethical principles¹³. Specifically, we asked how important each of these practices are for people to trust in AI systems.

Principles and Practices for Trustworthy AI



Technical robustness and safety

The performance and accuracy of AI system output is assessed before and regularly during deployment to ensure it operates as intended. The robustness of output is tested in a range of situations, and only data of appropriate quality is used to develop AI.



Data privacy, security and governance

Safety and privacy measures are designed into the AI system. Data used for AI is kept secure, used only for the specific purpose to which it is agreed, and is not shared with other apps or third parties without permission. Robust security measures are in place to identify and prevent adversarial attacks.



Human agency and oversight

There is appropriate human oversight and control of AI systems and their impact on stakeholders by people with required expertise and resources to do so. AI systems are regularly reviewed to ensure they are operating in a trustworthy and ethical manner.



Transparency and explainability

The purpose of the AI system, how it functions and arrives at its solutions, and how data is used and managed is transparently explained and reasonably understandable to a variety of stakeholders. Developers keep an audit trail of the method and datasets used to develop AI.



Fairness and non-discrimination

The outcomes of AI systems are assessed regularly to ensure they are fair, free of unfair bias, and designed to be inclusive to a diversity of users. AI is developed with the participation and input of a diverse range of people.



Accountability and contestability

There is clear accountability and responsibility if something goes wrong with an AI system. Any impacted user or stakeholder is able to challenge the outcomes of an AI system via a fair and accessible human review process.



AI literacy

People are supported in understanding AI systems, including when it is appropriate to use them, and the ethical considerations of their use.



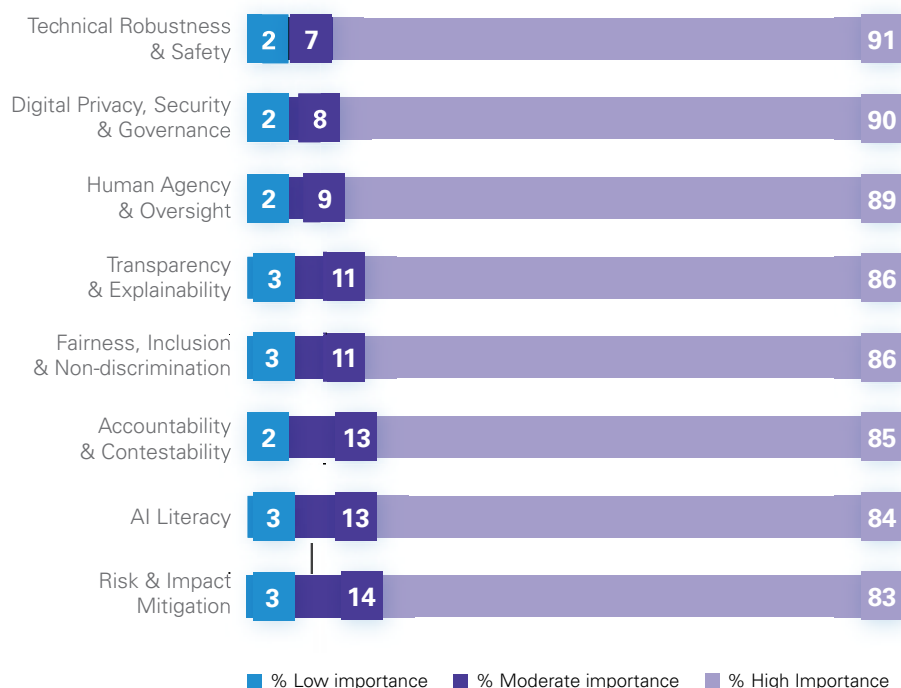
Risk and impact mitigation

The risks, unintended consequences and potential for harm from an AI system are fully assessed and mitigated prior to and during its deployment.

Results indicate that every one of the practices underlying these eight principles are highly important for the public's trust in AI systems (see Figure 17). This provides clear public endorsement for these principles and practices and a blueprint for developing and using AI in a way that supports trust.

Figure 17. Principles for Trustworthy AI

'How important are the following [...] for you to trust AI systems?'




Low importance = 'Not at all important', 'Very low importance' or 'Low importance'

Moderate importance = 'Moderately important'

High importance = 'High importance', 'Very high importance' or 'Extremely important'



How do Australians feel about AI at work?



AI is becoming more common in the workplace. To understand how Australians feel about the use of AI at work, we asked questions about AI use at work, comfort with AI use to support different work functions, the impact of AI on jobs, and expectations of employers that automate jobs.

Australians vary in their use of AI at work

Most Australians (62%, see Figure 18) report that little or none of their work involves AI. Approximately a third indicate that AI is not involved in their work at all, and a little under a third report that it is used in a small proportion of their work (i.e. in 20% of their work or less). In contrast, about two in five Australians report that AI is used in more than 20% of their work.

Given many Australians have a low understanding and awareness of AI use, these figures may reflect that AI is not being used at work, or that people are not aware of its use at work.

Figure 18. Use of AI at work

'What percentage of your work involves some form of AI?'

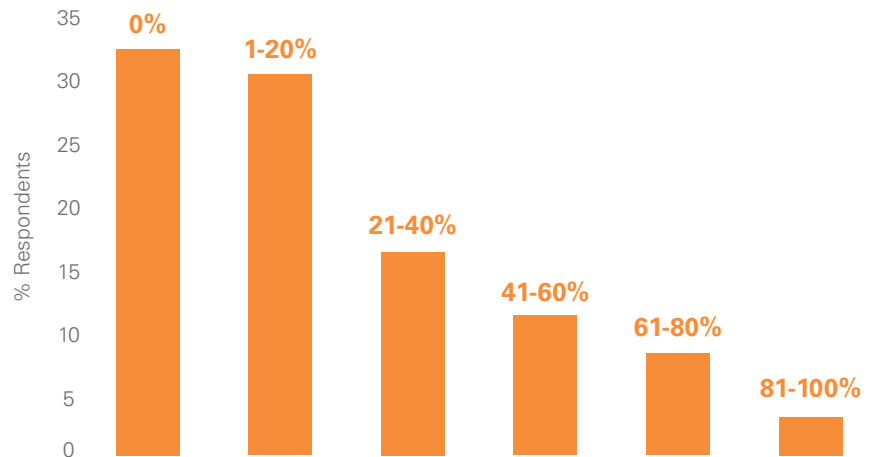
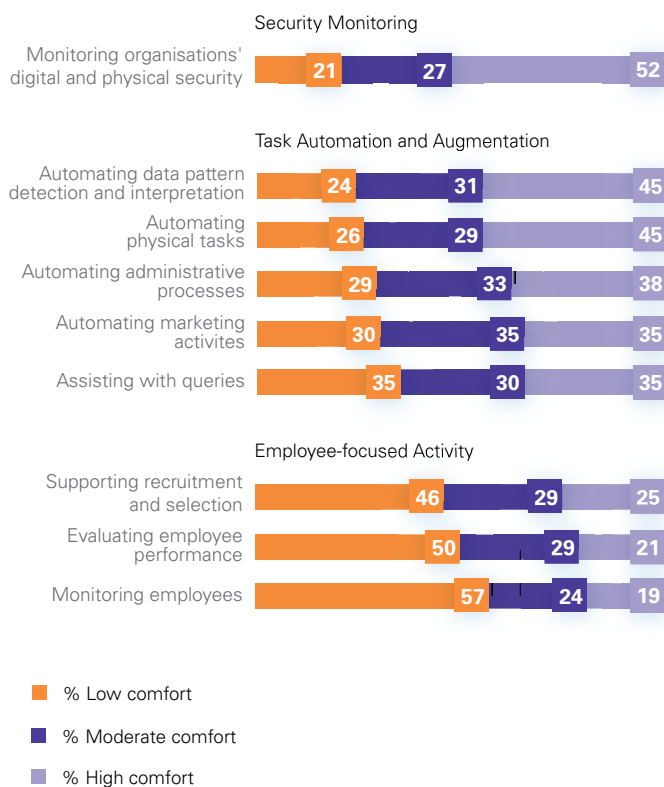


Figure 19. Comfort with the use of AI at work

'How comfortable are you with AI being used in the following ways at work?'



Australians are comfortable with AI use at work when it is not focused on them

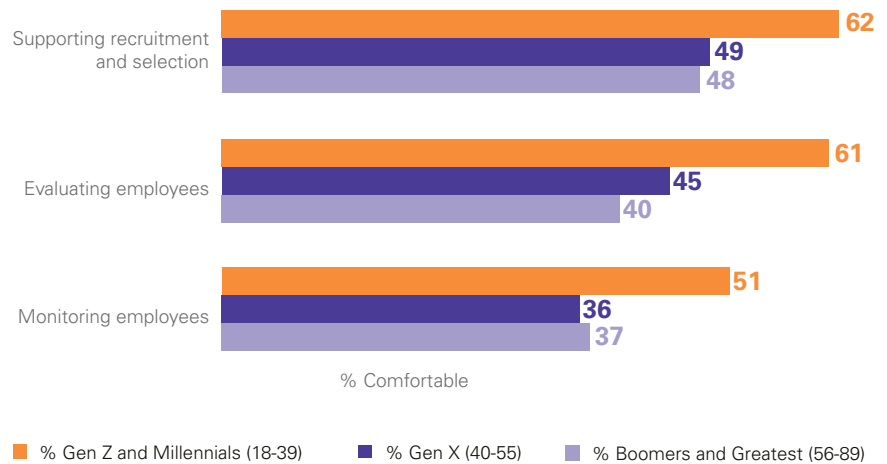
As demonstrated in Figure 19, most Australians are either highly or moderately comfortable with AI use in monitoring organisational security, and task automation and augmentation, such as supporting marketing and assisting with queries. However, they are considerably less comfortable with AI use when it is focused on themselves as employees – such as to monitor employees, evaluate performance and support recruitment and selection decisions.

On average, people who report using AI in their work are more likely to feel comfortable with its use across various functions than those who do not report using AI in their work. This most likely reflects their greater familiarity with the use of AI at work.

Younger people are more comfortable with employee-focused AI at work

Younger people are more likely to be comfortable with employee-focused AI activities at work than older respondents are. For instance, over half of Generation Z and Millennials are at least moderately comfortable with the use of AI to monitor employees, compared to just a third of older respondents (see Figure 20).

Figure 20. Age differences in comfort with employee-focused AI at work

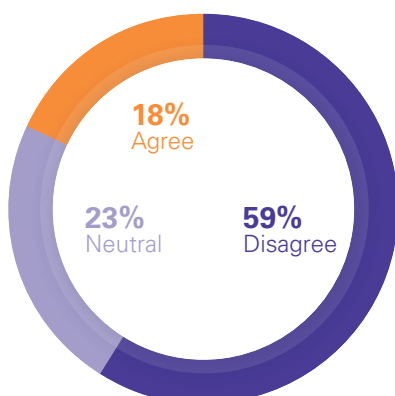


Most Australians don't believe AI will create more jobs than it will eliminate

Most Australians disagree (59%) that AI will create more jobs than it will eliminate (see Figure 21). The concern that AI will eliminate jobs is also expressed in prior national¹⁴ and transnational¹⁵ surveys.

Figure 21. Perceived impact of AI on jobs

'In general, to what extent do you agree that AI will create more jobs than it will eliminate?'

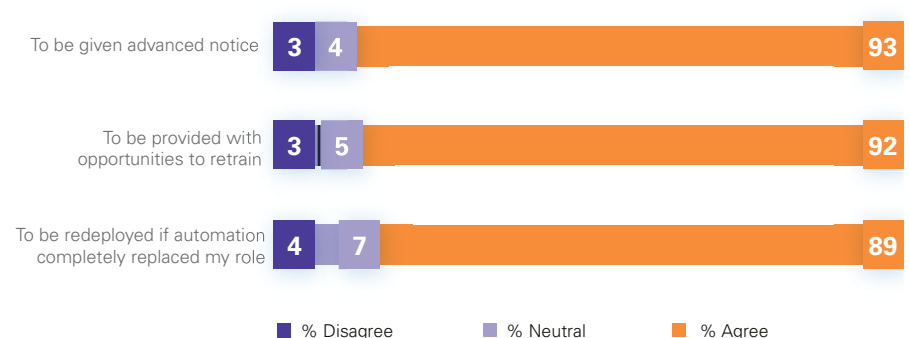


If jobs are automated, Australians expect support

We asked three questions around expectations if one's job was to be automated. Almost all Australians agree that they would expect to be given advanced notice, to be provided with opportunities to retrain and to be redeployed if their role was to be fully automated (see Figure 22).

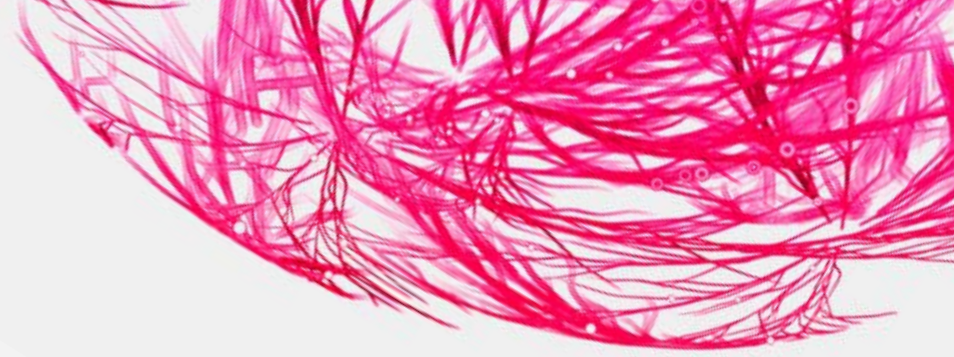
These findings highlight the clear expectations that employees have of their employers in the context of automation. Employers are expected to give advanced notice of the change and provide substantive support to employees impacted by automation.

Figure 22. Employee expectations should their work be automated





How do Australians view key AI challenges?



The pervasive use of AI in society is leading to a range of challenges. We asked respondents to rate the extent to which a series of AI societal challenges need to be carefully managed, and the likelihood of these challenges affecting large numbers of Australians in the next ten years.

Australians expect all AI challenges to be carefully managed

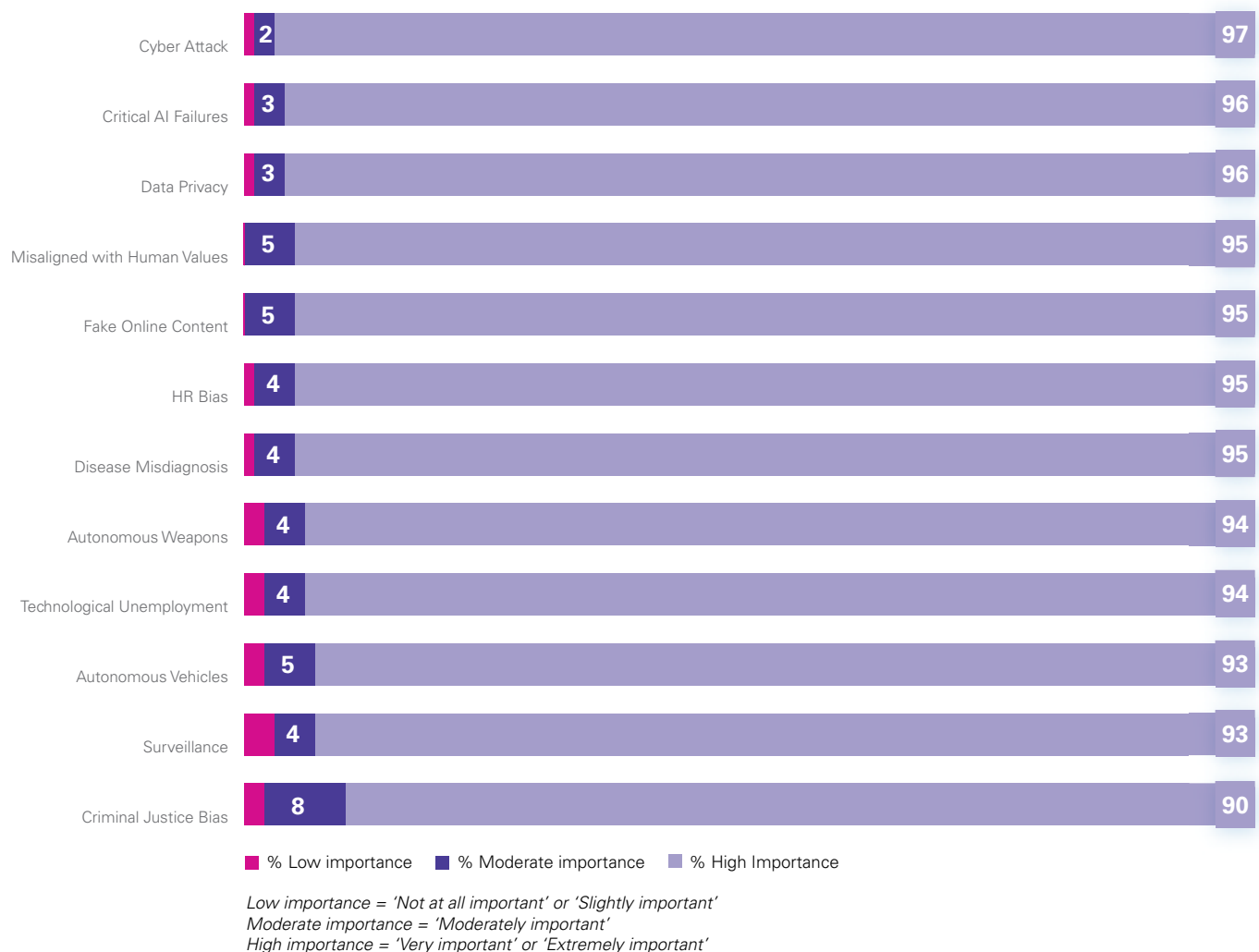
All twelve of the AI challenges we presented need to be carefully managed by governments and organisations. Figure 23 shows that almost all respondents (90%, or more) rate the careful management of AI challenges as very or extremely important.

While there are small generational differences in opinions about the management of AI challenges, nearly all Australians rate the careful management of these challenges as highly important.

These findings are not unique to Australia and align with those found in a previous US survey, where Americans also note that these AI challenges need to be carefully managed¹⁶.

Figure 23. Importance of carefully managing AI challenges

'How important is it for companies and governments to carefully manage this challenge?'



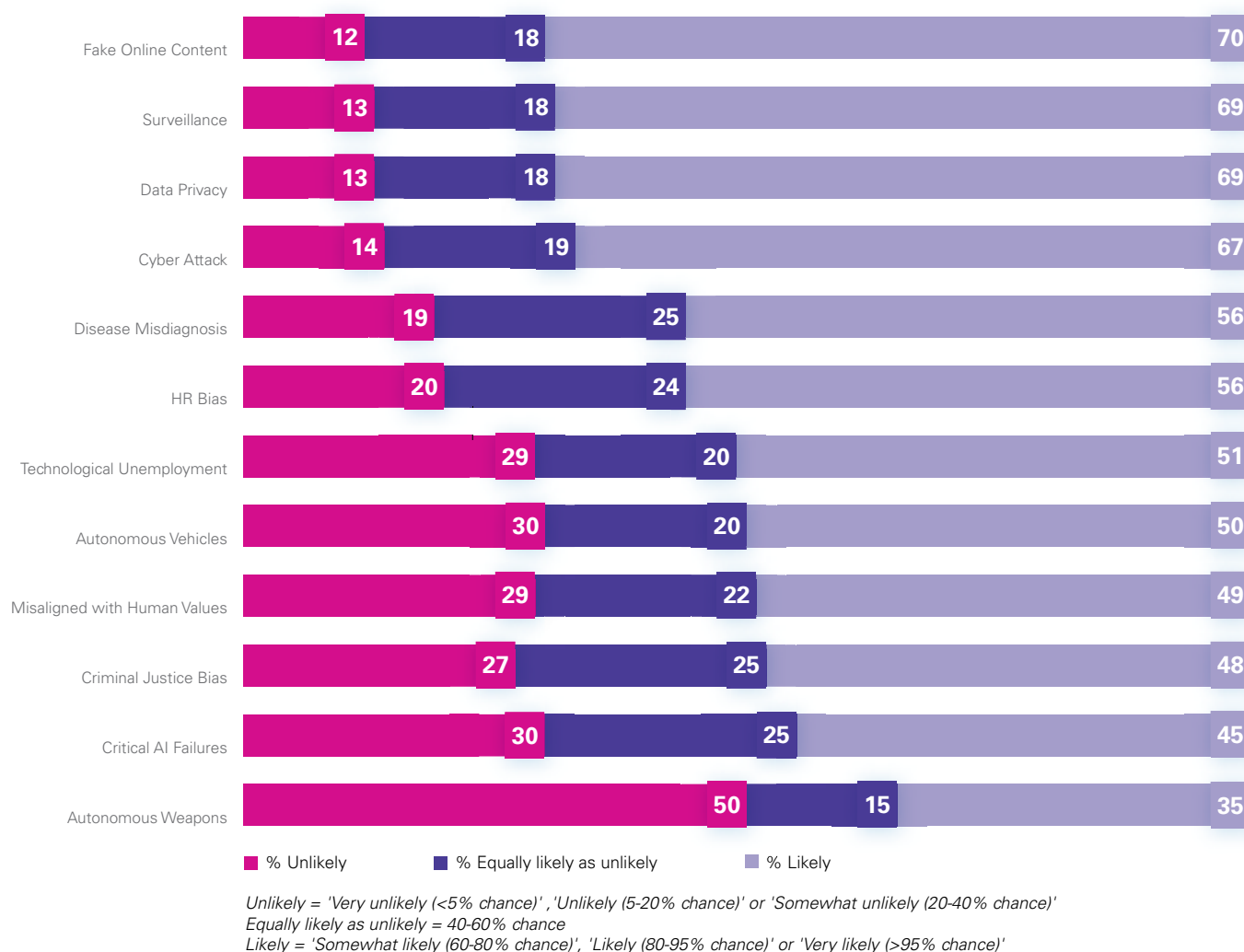
Australians think data challenges are most likely to impact people in the near future

Figure 24 indicates that most respondents (67 – 70%) think data challenges such as fake online content, surveillance, data privacy and cyber-attacks are most likely to impact large numbers of Australians over the next 10 years. The only challenge which people perceive to be more unlikely (50%) than likely (35%) to impact large numbers is the use of lethal autonomous weapons.

Our findings broadly corroborate those of a previous US survey¹⁷ in which US respondents also believed that data challenges would be the most likely to impact large numbers of Americans in the near future.



Figure 24. Likelihood of AI challenges impacting large numbers of Australians

'In the next 10 years, how likely do you think it is that this challenge will impact large numbers of the people in Australia?'





How well do Australians understand AI?



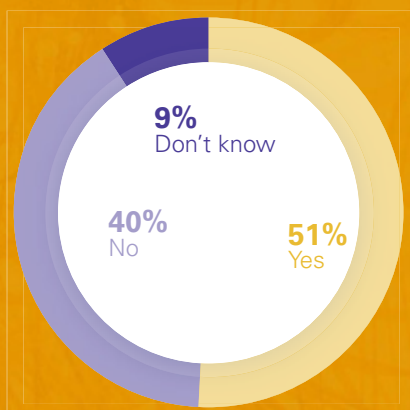
To identify how well Australians understand AI, we asked about AI awareness, knowledge about AI and interest to learn more. Our findings reveal that, in general, Australians have low awareness and understanding of AI and low knowledge of its use in common everyday applications. We further show in the next section of the report that awareness and understanding of AI influences trust in AI systems. Most Australians want to know more about AI.

Just over half of Australians are aware of AI

Only 51 % of the public had heard, read or seen something about AI in the previous 12 months (see Figure 25). This finding is similar to results from recent European¹⁸ and American¹⁹ national surveys that indicate that public awareness of AI is low.

Figure 25. Awareness of AI

'In the past 12 months, have you heard, read or seen anything about AI?'

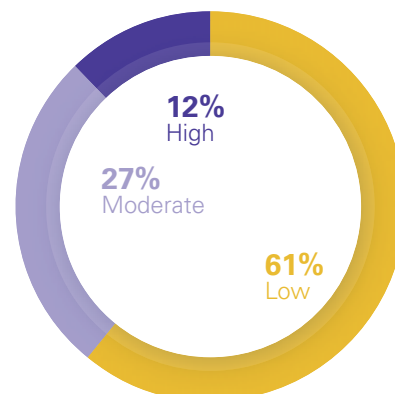


Australians report a low understanding of AI

The majority of Australians (61 %, see Figure 26) also report low subjective knowledge of AI, indicating that they feel they know little about AI, or when and how it is being used. Only a small proportion of Australians (12%) report high subjective knowledge of AI.

Figure 26. Subjective knowledge of AI

'To what extent do you...'
a) *feel you know a lot about AI?*
b) *feel informed about how AI is used?*
c) *think you understand when AI is being used?'*



Low = 'Not at all' or 'Slightly'
Moderate = 'Moderately'
High = 'Considerably' or 'A great deal'

Australians report a low understanding of when AI is used

Given the low understanding of AI, it is not surprising that Australians are often unaware that AI is used in common everyday technologies. When asked if the common technologies shown in Figure 27 use AI, less than 50% correctly answered yes. That is, people could not correctly identify if the technology used AI better than a chance guess.

In particular, the majority of Australians (53 – 73%) are unaware that AI is being used in common applications such as accommodation and ridesharing applications (Airbnb and Uber), social media platforms (Facebook and Instagram) and product recommendation apps (Netflix and Amazon).

In contrast, most Australians are aware that AI is used in embodied forms, which is where the AI system has a physical or vocal representation. For example, most Australians know that AI is used in autonomous vehicles, virtual assistants, social bots and smart speakers.

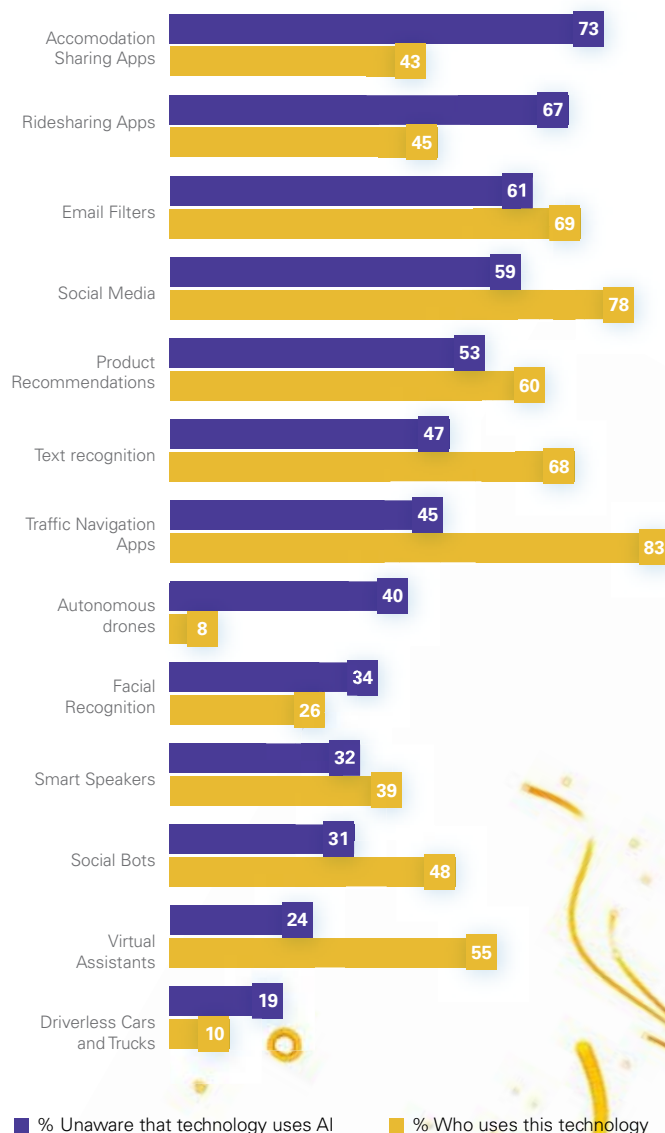
Surprisingly, use of a technology does not necessarily translate into an increased understanding of whether that technology uses AI. As shown in Figure 27, this is particularly the case when AI is used in an embedded form (without physical or vocal manifestation), such as social media, email filters and traffic navigation apps. For example, even though 78% of Australians report using social media, 59% of them are unaware that social media apps use AI.

These findings broadly mirror results from other international surveys reporting low public understanding of AI. For example, in a recent survey of the American public²⁰, a lower percentage correctly understand that AI is used in virtual assistants, smart speakers, driverless vehicles and autonomous drones compared to our Australian sample.

We further probed people's awareness of the emerging use of AI in four domains. We asked Australians whether they were aware or had experience of AI use in healthcare diagnostics, predictive policing, human resource analytics, or personal financial investment (see Appendix for more details). Four out of five Australians (81%) indicated they had no awareness of such systems.

Figure 27. Use of AI technologies and understanding of these technologies use AI

'For each technology below, please indicate if you have used it and if it uses AI?'



Most Australians want to know more about AI

While Australians generally lack knowledge and awareness of AI, the large majority (86%) are interested in learning more about AI. Only 14% stated they are not at all interested in learning more about AI.

Some population segments have more awareness and knowledge of AI

Men, younger people, those with a university education, and people with computer science experience are more likely to be aware of AI (see Figure 28), have higher subjective knowledge about AI (see Figure 29) and are more likely to understand when AI is being used in common applications. Higher income earners are also more likely to report higher subjective knowledge of AI²¹.

Figure 28. AI awareness by population segment

'In the past 12 months, have you heard, read or seen anything about AI?'

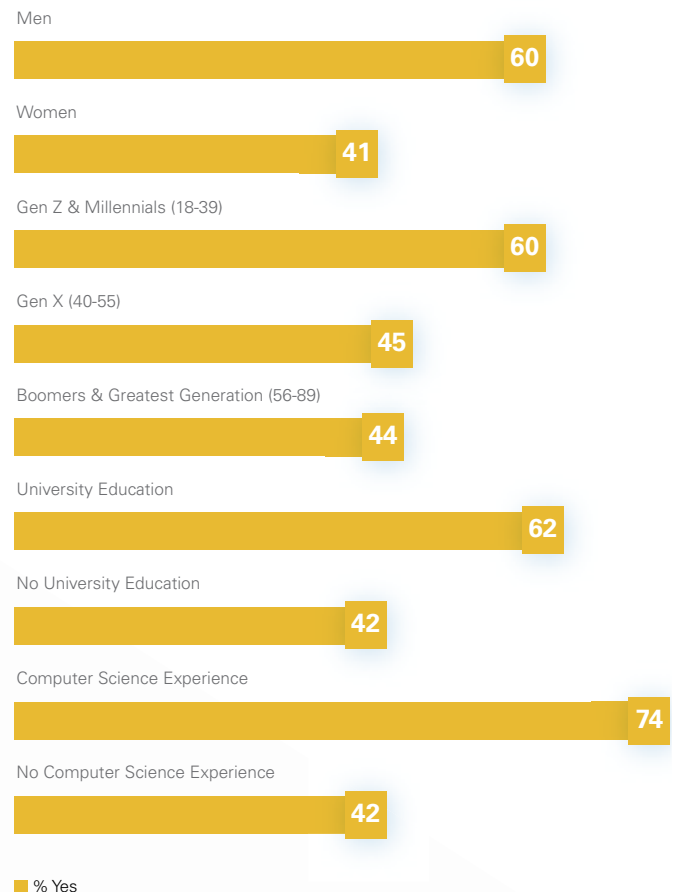


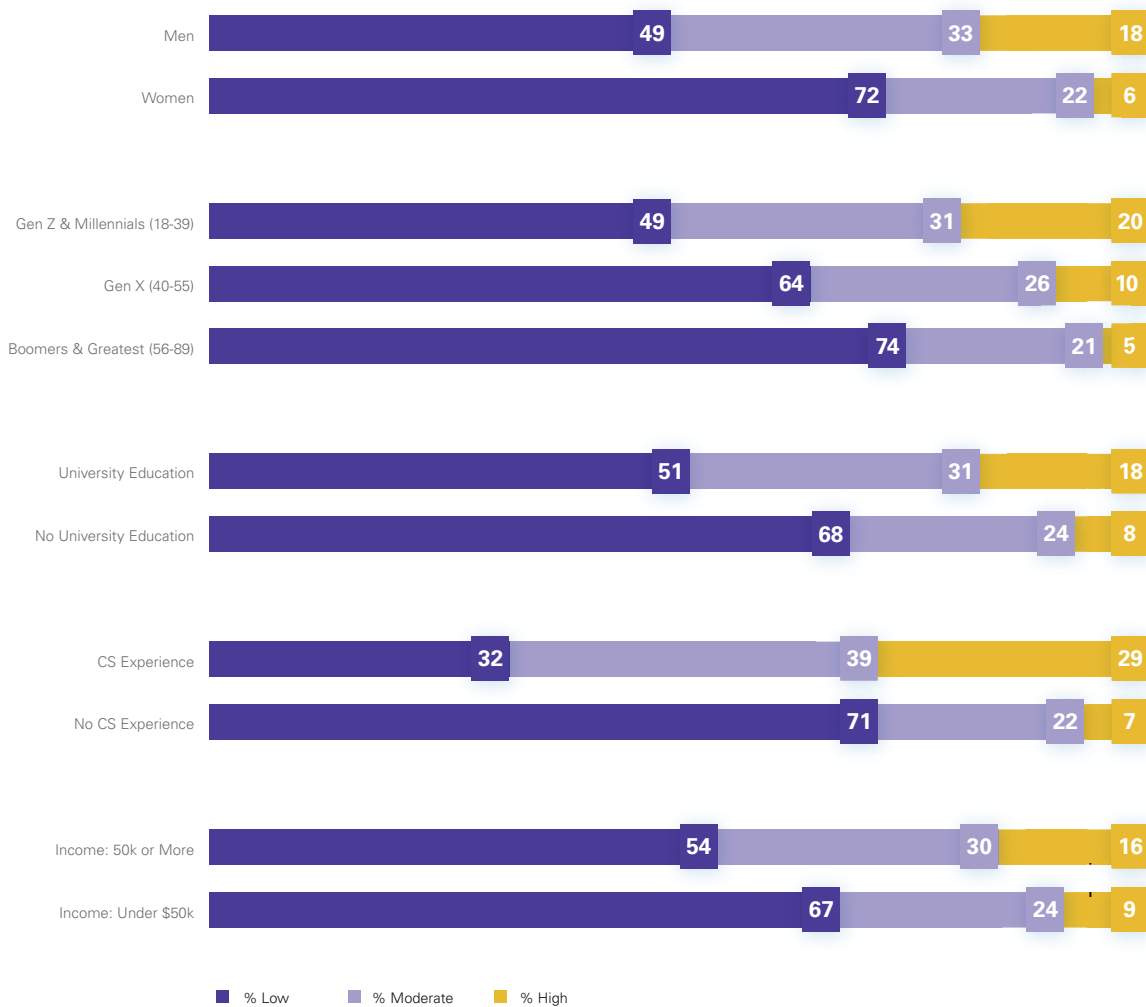
Figure 29. Subjective knowledge by population segment

'To what extent do you...

a) feel you know a lot about AI?


b) feel informed about how AI is used?

c) think you understand when AI is being used?'





What are the key drivers of trust and acceptance of AI?



To identify the most important drivers of trust and acceptance of AI systems examined in this report, we used an advanced statistical technique called path analysis. We explain the path model in Figure 30, together with notes on interpreting the model.

Trust is central to AI acceptance

The path model shows that trust is the central driver of AI acceptance ($B = .44$). This finding empirically supports why trustworthy AI matters: if people perceive AI systems to be trustworthy and are willing to trust them, this leads to the acceptance necessary to realise the potentially vast societal benefits that AI can produce.

Trust acts as the central vehicle through which other drivers impact AI acceptance. Each of the four drivers on the left-hand side of the model influences trust, which in turn influences acceptance²². Given the key role of trust in driving acceptance, it is important to understand what drives trust in AI systems.

The strongest driver of trust is believing current regulations and laws are sufficient to ensure AI use is safe

As shown in the path model, believing current safeguards are sufficient is the strongest driver of trust. The relative importance of current safeguards ($B = .50$) is more than twice that of the next strongest driver, AI uncertainty ($B = -.20$).

This demonstrates the importance of developing adequate regulatory and legal systems that protect people from problems that may arise from AI use, and make them feel safe to use AI systems. Given most Australians either disagree or are ambivalent that current AI safeguards are adequate, ensuring AI is governed by an appropriate regulatory and legal framework is a critical first step towards enhancing trust and acceptance of AI.

The more people believe the impact of AI is uncertain, the less they trust AI

The importance of regulations and laws as a driver of trust is not surprising given three out of four Australians believe the impact of AI on society is uncertain and unpredictable (see Figure 14). The path model shows that the extent to which people believe the impact of AI on society is uncertain and unpredictable, the less they trust in and accept AI. This is the second strongest driver of trust.

The perceived impact of AI on jobs, and familiarity with AI, influence trust

People's beliefs about the impact of AI on jobs is the third strongest driver of trust ($B = .17$). People who believe AI will create more jobs than it will eliminate are more likely to trust in AI systems. Familiarity with AI was the fourth driver of trust ($B = .13$). This shows that people who feel that they understand how and when AI is used, and have experience using AI systems, are more likely to trust AI.

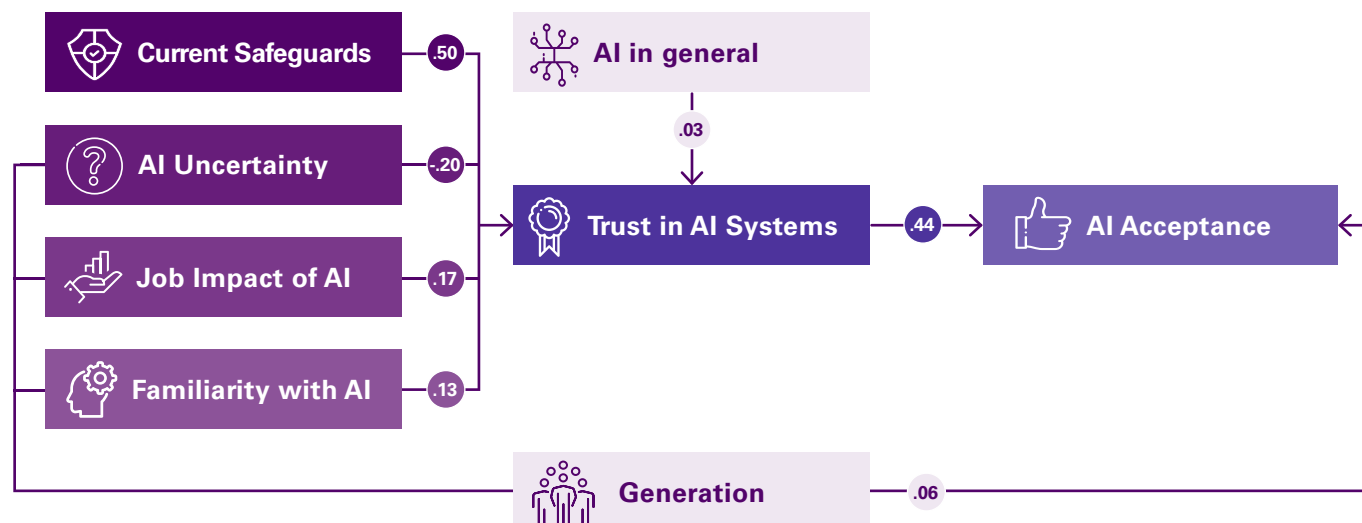
We also found that two other factors had a smaller impact on trust and acceptance. People are generally more trusting of AI systems in general rather than of specific applications of AI (see Appendix), but this effect is small ($B = .03$). Younger people in the Generation Z and Millennial age category also tend to be more accepting of AI than people born into Generation X, Baby Boomer and the Greatest Generation ($B = .06$)²³.

How to read the path model

When reading the path model, follow the arrows from left to right. The values on the arrows indicate their relative importance in driving trust and acceptance: the larger the number, the stronger the effect. The negative value of AI uncertainty indicates that when uncertainty increases, trust and acceptance decrease. All other relationships are positive, which means, for example the more people believe current safeguards are sufficient, the more they will trust AI systems, and the more they trust AI systems, the more they accept AI. Only significant relationships are shown.

Trust is central to the acceptance of AI systems and is influenced by four key drivers. This model lays out a pathway to building trust and acceptance of AI.

Figure 30. A model of the key drivers of trust and acceptance of AI systems



The belief that current regulations, laws and safeguards are sufficient to protect people and ensure AI use is safe. This is the strongest predictor of trust in AI systems.



The belief that the societal impact of AI is unpredictable and there are many unknowns about AI.



The belief that AI will create more jobs than it will eliminate.



The extent to which people feel they understand AI, know when AI is used in common applications, and have used common AI applications.



The extent to which people trust AI systems and perceive them to be trustworthy.



Other factors also had a small impact on trust and acceptance.



People are slightly more trusting of AI systems in general than of specific applications of AI.

Gen Z and Millennials are more accepting of AI than older generations.



The extent to which people accept and approve of AI.

Conclusion and implications

Together, the findings of this national survey of the Australian community highlight important insights on the public's trust and attitudes towards AI and lay out a pathway for building trust and acceptance of AI systems. The findings also provide a clear overview of the current and future challenges to building and preserving trust and acceptance of AI systems.

A key insight from the survey is that the Australian public is generally ambivalent in their trust towards AI systems. Given trust is a central factor influencing the acceptance and adoption of AI²⁴, this ambivalence in trust is likely to impair societal uptake and the ability of Australia to realise the societal and economic benefits of AI, if left unaddressed. The following insights lay out a roadmap for enhancing public trust in AI.

Live up to Australians' expectations of trustworthy AI

- Our findings reveal that the public have very clear expectations of the principles and practices they expect AI systems to uphold in order to be trusted. They expect organisations to maintain high standards of AI systems in terms of:
 - performance and accuracy
 - data privacy, security and governance
 - transparency and explainability
 - accountability
 - risk and impact mitigation
 - fairness
 - human oversight.

These principles and practices reflect those identified in numerous recent government reports on trustworthy, ethical AI²⁵, and our findings provide clear public endorsement for them, as well as underscoring their importance for public trust.

- The public clearly expect AI systems will be monitored and evaluated on an ongoing basis. Organisations should undertake regular in-house and independent ethics reviews of their AI systems to ensure they operate according to these principles and practices.
- Our survey revealed that most people believe organisations innovate with AI for commercial reasons (e.g. cost saving or profit maximisation) rather than to benefit society more broadly. This imbalance is most pronounced for commercial organisations, followed by government and then non-profit organisations. This highlights the opportunity for all organisations to better use AI systems for the benefit of citizens, customers and employees, as well as better demonstrate how their use of AI supports societal health and wellbeing.
- In the event their jobs are automated, employees clearly expect to be given fair notice and provided with opportunities to retrain or be redeployed. Many Australians believe AI will eliminate more jobs than it creates making this a real threat. Living up to employees' expectations in the event of automation will require strategic long-range workforce planning and retraining opportunities that are available to employees of all ages.
- Our findings further reveal that while most Australians are comfortable with AI use for the purposes of protecting organisational security and task automation and augmentation, they are less comfortable with AI use for employee-focused activities, such as evaluating and monitoring performance, and recruitment and selection.
- Taken together, these findings highlight that organisations looking to accelerate the use and uptake of AI need to build trust with customers, employees and the public more broadly – it is not enough to focus on only one stakeholder group.
- Organisations also need to consider that different cohorts in the workplace and community have different views about AI, with younger people and the university educated being more trusting and accepting of AI. A one-size-fits-all approach is therefore unlikely to work.



Strengthen the regulatory framework for governing AI

- While most Australians believe the benefits of AI are either greater than or equal to the risks, the majority also view the societal impacts of AI as uncertain and unpredictable. Furthermore, most Australians believe the challenges associated with AI such as fake online content, surveillance, data privacy, cyber security, bias, technological unemployment and autonomous vehicles, are likely to impact a large number of Australians. The Australian public are near unanimous in their expectation that the government and the companies deploying AI carefully manage these challenges.
- It is understandable, therefore, that the large majority (96%) of the Australian community expect AI to be regulated. However, many view the current regulatory and legal framework as insufficient to make AI use safe and protect people from the risks.
- Given this pattern and the finding from this survey that the perceived adequacy of current regulations and laws is the single most important driver of public trust in AI systems, a clear pathway to enhancing trust in AI is to strengthen the regulatory and legal framework governing AI.
- The public clearly want appropriate regulation of AI that is fit-for-purpose to manage the risks and uncertainties associated with AI. Our results further show that the public expect the government and existing regulators to take the lead in regulating and governing AI systems, rather than leaving it to industry only. The majority of the public have at least moderate confidence in the government to do so in the public's best interest.
- Given the public has the most confidence in Australian universities, research and defence organisations to develop and use, as well as regulate and govern AI systems, there is an opportunity for business and government to partner with these organisations around AI initiatives.
- Our findings further indicate that organisations can directly build trust and willingness to use AI systems by adopting assurance mechanisms that support the ethical deployment of AI systems. These include actions such as establishing independent AI ethics reviews, adopting codes of conduct and national standards, and obtaining AI ethics certification.



Strengthen Australia's AI literacy

- A key finding is that the Australian community generally has low awareness and understanding of AI and its use in everyday life. While younger people, men, the university educated, and those with computer science experience tend to understand AI better, even these groups report low to moderate AI understanding.
- At the same time, an overwhelming majority of the community are interested in learning more about AI (86%), and report that supporting people to understand AI, is important for them to trust AI systems (97%). This last insight is further supported by our path model, which identified familiarity and understanding of AI as a key driver of trust and acceptance of AI.
- Collectively these insights paint a clear picture of the need to increase the AI literacy of the Australian public. Educating the community about what AI is and when and how it is being used is important for a range of reasons. First, despite the current low awareness and understanding, the community have strong views on the regulation, use and design of AI. Increasing public literacy will assist in ensuring these views are well informed. Second, the more informed citizens, consumers and employees are about AI, the better able they will be to seize the benefits of such systems, while identifying and appropriately managing the associated risks (e.g. of data sharing and privacy). Third, AI literacy is fundamental to the public's ability to contribute to effective public policy and debate on the stewardship of AI in society, and facilitate meaningful consultation with the public on AI design.
- Some countries have already invested in freely available AI public literacy courses²⁶. We recommend that enhancing Australia's AI literacy be a responsibility shared by government (e.g. formal programs within schools and for citizens), and organisations using or developing AI (e.g. by investing in employee and customer AI literacy programs and tools).

Appendix

AI Systems in general versus Specific AI use cases.

We asked a subset of questions to explore whether Australians trust and attitudes towards AI differ depending on whether we asked about AI systems in general versus specific use cases of AI systems (where the purpose of using the AI system is clear).

To do this, we randomly allocated respondents so the subset of questions asked either about AI in general or one of four specific use cases of AI systems: healthcare diagnostics, human resource analytics, predictive policing or personal investment.

Before answering questions, respondents were provided with a brief description of what the system does, how it works and how it is used. These descriptions were developed by the research team with input from domain experts and are based on a range of in use systems. Short descriptions are provided here.

In reporting the findings, when we find meaningful differences between the specific AI use systems, we highlight these in the report. When we find no substantive differences between the four specific use systems, we amalgamate these results into a single category of specific AI systems and compared them to responses about AI systems in general. Relatedly, when we find no meaningful differences between AI in general and specific AI systems, we amalgamate the results into a single category of AI systems.



Healthcare AI

An AI system used to improve the diagnosis and treatment of disease. The system compares patient health data to existing databases to produce recommendations. Doctors use the system to inform decisions about how to diagnose and treat patients.



Human Resource AI

An AI system used to improve the prediction and evaluation of performance by collecting and comparing employee data and job performance over time. Managers use the system to inform decisions about hiring and promotion.



Policing AI

A predictive policing system that analyses crime data to make predictions about the most likely crime locations and types of offenders. The police use the system to inform resourcing decisions about where to place officers to prevent crime.



Investment AI

An AI system that recommends shares to trade on the stock market by analysing millions of traditional and non-traditional data sources in real time, and factoring in the investment preferences of the user. People use the system to inform investment decisions.

Endnotes

- 1 International Data Corporation. (2019, September 4). Worldwide Spending on Artificial Intelligence Systems Will Be Nearly \$98 Billion in 2023, According to New IDC Spending Guide. Retrieved from www.idc.com
- 2 AI HLEG. (2019). Ethics Guidelines for Trustworthy AI. European Commission. Retrieved from <https://ec.europa.eu/>
Walsh, T., Levy, N., Bell, G., Elliott, A., Maclaurin, J., Mareels, I.M.Y., Wood, F.M., (2019) The effective and ethical development of artificial intelligence: An opportunity to improve our wellbeing. Report for the Australian Council of Learned Academies, www.acola.org
- 3 <https://www.chiefscientist.gov.au/news-and-media/ai-can-we-get-it-right-please>
<https://www.weforum.org/agenda/2018/05/alan-finkel-turing-certificate-ai-trust-robot/>
- 4 Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures*, 90, 46-60.
- 5 Schwab, K. (2016). The Fourth Industrial Revolution. World Economic Forum.
- 6 OECD (2019), Artificial Intelligence in Society, OECD Publishing, Paris. <https://doi.org/10.1787/eedfee77-en>
- 7 Biddle, N., Edwards, B., Gray, M., Hiscox, M., McEachern, S., & Solis, K. Data trust and privacy in the COVID-19 period. Retrieved from <https://csrc.cass.anu.edu.au/>
- 8 In line with prior surveys, we separate out technology companies (e.g. Google, Facebook, Apple) from other commercial organisations using AI, such as banks and retailers.
- 9 Eurobarometer. (2017). Attitudes towards the impact of digitisation and automation on daily life (Report no. 460). Retrieved from <https://ec.europa.eu/>
- 10 Zhang, B., & Dafoe, A. (2019). Artificial intelligence: American attitudes and trends. Retrieved from SSRN 3312874.
- 11 Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389-399.
- 12 AI HLEG. (2019). Ethics Guidelines for Trustworthy AI. European Commission. Retrieved from <https://ec.europa.eu/>
- 13 Australia's AI Ethics Principles. Retrieved from <https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework/ai-ethics-principles>
- 14 Zhang, B., & Dafoe, A. (2019). Artificial intelligence: American attitudes and trends. Retrieved from SSRN 3312874.
- 15 Eurobarometer. (2017). Attitudes towards the impact of digitisation and automation on daily life (Report no. 460). Retrieved from <https://ec.europa.eu/>
Edelman AI. (2019). 2019 Edelman AI Survey. Retrieved from <https://www.edelman.com/>
- 16 Zhang, B., & Dafoe, A. (2019). Artificial intelligence: American attitudes and trends. Retrieved from SSRN 3312874; Descriptions of the AI challenges were adapted from this US survey.
- 17 Zhang, B., & Dafoe, A. (2019). Artificial intelligence: American attitudes and trends. Retrieved from SSRN 3312874.
- 18 Eurobarometer. (2017). Attitudes towards the impact of digitisation and automation on daily life (Report no. 460). Retrieved from <https://ec.europa.eu/>
- 19 Zhang, B., & Dafoe, A. (2019). Artificial intelligence: American attitudes and trends. Retrieved from SSRN 3312874.
- 20 Zhang, B., & Dafoe, A. (2019). Artificial intelligence: American attitudes and trends. Retrieved from SSRN 3312874.
- 21 Recent research undertaken by Essential Research on behalf of the Australian Human Rights Commission examined Australians' awareness of the federal government's use of automated decision-making. The findings of that survey also indicates greater awareness in men, the university-educated and higher income earners.
- 22 Current safeguards only influence AI acceptance through trust, meaning that when current safeguards are perceived as sufficient, trust in AI systems increases, which in turn promotes AI acceptance. AI uncertainty, the job impact of AI and familiarity with AI influence acceptance directly, as well as through their impact on trust.
- 23 University education was also included in the model. However, as it did not have a significant impact on either trust or acceptance, it is not displayed for ease of presentation.
- 24 AI HLEG. (2019). Ethics Guidelines for Trustworthy AI. European Commission. Retrieved from <https://ec.europa.eu/>
Futureye. (2019). Humanising AI: Creating Social Licence for Artificial Intelligence to Flourish. Retrieved from <https://futureye.com/>
Siau, K., & Wang, W. (2018). Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal*, 31(2), 47-53.
Walsh, T., Levy, N., Bell, G., Elliott, A., Maclaurin, J., Mareels, I.M.Y., Wood, F.M., (2019) The effective and ethical development of artificial intelligence: An opportunity to improve our wellbeing. Report for the Australian Council of Learned Academies, www.acola.org.
- 25 AI HLEG. (2019). Ethics Guidelines for Trustworthy AI. European Commission. Retrieved from <https://ec.europa.eu/>
Australia's AI Ethics Principles Retrieved from <https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework/ai-ethics-principles>
The Assessment List for Trustworthy AI (2020). European Commission. Retrieved from <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines/2>
For a review see Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389-399.
Fjeld, J., Achten, N., Hillgoss, H., Nagy, A., & Srikanth, M. (2020). Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI. Berkman Klein Center Research Publication No. 2020-1. doi.org/10.2139/ssrn.3518482
- 26 An example is Finland's Elements of AI course run by the University of Helsinki (see <https://www.elementsofai.com/>)



Key contacts

University of Queensland

Nicole Gillespie
KPMG Chair in Organisational Trust
Professor of Management,
The University of Queensland
T: +61 7 3346 8076
E: n.gillespie1@uq.edu.au

Dr Steve Lockey
Research Fellow (Trust)
The University of Queensland
T: +61 7 3443 1206
E: s.lockey@uq.edu.au

Dr Caitlin Curtis
Research Fellow (Trust)
The University of Queensland
T: +61 7 3346 8083
E: c.curtis@uq.edu.au

KPMG

James Mabbott
National Leader, KPMG Innovate
KPMG Australia
T: +61 2 9335 8527
E: jmabbott@kpmg.com.au

Richard Boele
Global Leader, Business
& Human Rights Services
KPMG Australia
T: +61 2 9346 585
E: rboele@kpmg.com.au

Ali Akbari
Artificial Intelligence
Capability Lead
KPMG Australia
T: +61 2 9335 7740
E: aakbari@kpmg.com.au

Rossana Bianchi
Associate Director, Strategy,
Growth & Digital
KPMG Australia
T: +61 2 9335 7036
E: rbianchi2@kpmg.com.au

Rita Fentener van Vlissingen
Associate Director, Human
Rights & Social Impact
KPMG Australia
T: +61 2 9346 6366
E: ritafentener@kpmg.com.au

KPMG.com.au



©2020 The University of Queensland

The information contained in this document is of a general nature and is not intended to address the objectives, financial situation or needs of any particular individual or entity. It is provided for information purposes only and does not constitute, nor should it be regarded in any manner whatsoever, as advice and is not intended to influence a person in making a decision, including, if applicable, in relation to any financial product or an interest in a financial product. Although we endeavour to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act on such information without appropriate professional advice after a thorough examination of the particular situation.

To the extent permissible by law, KPMG and its associated entities shall not be liable for any errors, omissions, defects or misrepresentations in the information or for any loss or damage suffered by persons who use or rely on such information (including for reasons of negligence, negligent misstatement or otherwise).

©2020 KPMG, an Australian partnership and a member firm of the KPMG global organisation of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved.

The KPMG name and logo are trademarks used under license by the independent member firms of the KPMG global organisation.

Liability limited by a scheme approved under Professional Standards Legislation.

October 2020. 573005795MC