Australian Government
Department of the Prime Minister and Cabinet

Australian Government
Long-term Insights Briefings

How might artificial intelligence affect the trustworthiness of public service delivery?

Contents

Summary

This report, on *How artificial intelligence might affect the trustworthiness of public service delivery*, is the first in an enduring series of public service-led Long-term Insights Briefings.

The Long-term Insights Briefings are a new initiative under the Australian Public Service (APS) Reform Agenda. They provide an opportunity for the APS to consider significant, cross-cutting and complex policy issues and how they may affect Australia and the Australian community in the medium and long term. The briefings will bring together, and help the APS to deeply understand the evidence (including the views of the Australian community, academia and industry), context, trends and implications of complex policy issues. This will build the capability and institutional knowledge of the APS for long-term thinking, and position the APS to support the public interest now and into the future, by understanding the long-term impacts of what it does.

These briefings are expected to form part of the evidence base for policy making over time. They are not intended to take the place of policy making across the public service and do not make recommendations for future action.

Artificial intelligence offers significant opportunities to improve public service delivery

This Long-term Insights Briefing explores how using artificial intelligence (AI) to deliver public services might affect the trustworthiness of public service delivery, now and in the future.

We will need to innovate to meet community expectations of public services in the future

Innovation, including adopting digital and AI technologies, can provide opportunities for efficiency gains that will allow governments to streamline everyday interactions and processes, focus resources on more complex and pressing issues, and meet the needs and expectations of the Australian people. The community's expectations around the quality of public services are growing: for a higher standard of care; for tailored and personalised services; and for greater responsiveness, convenience and efficiency when accessing services. Australia's population is ageing, increasing demand for care and support services. At the same time, an increase in the share of older Australians in the population means fewer working-age Australians to help fund public services.[1] External forces, such as climate change, are also expected to increase demand for services while decreasing the resources (people and funding) available to provide them.

Artificial intelligence could transform public service delivery, leading to a better experience and outcomes for the whole community

Opportunities exist to use AI across the spectrum of activities carried out by the APS. These include:

- automating 'backroom' administrative processes

- improving efficiency and minimising errors in data management
- improving service processing and response times, freeing up time for more creative and complex work
- AI-enabled public interfaces that offer customised services to users and facilitate communication with diverse populations.

Indeed, AI capabilities are already being used in solutions in some Commonwealth government agencies, including:

- chatbots, virtual assistants and agents in service delivery
- document and image detection and recognition for border control and for fraud detection
- data mapping to geographical areas.[2]

In the future, AI will be a critical tool for maintaining and delivering even better quality services to a growing and ageing population.

Using artificial intelligence in public service delivery is not without risk

Using AI for public service delivery is not risk free. AI systems can inherit biases present in their training data, leading to unfair or discriminatory outcomes. The collection and analysis of personal data for AI can have privacy and security risks. Complex AI systems might behave unpredictably, causing unintended outcomes at a scale and speed that are hard to control.

Although not an example of AI, the Royal Commission into the Robodebt Scheme examined an automated debt raising and recovery scheme for social security payments. In the Robodebt example, the underlying basis for debt calculation was flawed and resulted in many incorrect calculations. The scheme was then implemented quickly, and incorrectly calculated debts were raised and imposed on recipients at scale, thereby compounding the impact on the community. In her report, Commissioner Holmes AC SC concluded that '*When done well, AI and automation can enable government to provide services in a way that is "faster, cheaper, quicker and more accessible." Automated systems can provide improved consistency, accuracy and transparency of administrative decision-making. The concept of "when done well" is what government must grapple with as increasingly powerful technology becomes more ubiquitous.*' This briefing and its insights, on *How artificial intelligence might affect the trustworthiness of public service delivery*, are not a response to Robodebt. However, it offers an important framework for agencies considering how to adopt and implement AI – namely that realising the benefits of AI in public service delivery will require agencies to identify and mitigate these (and other) risks in order to ensure that AI is used in a safe and responsible manner.

Stewardship of artificial intelligence in public service delivery

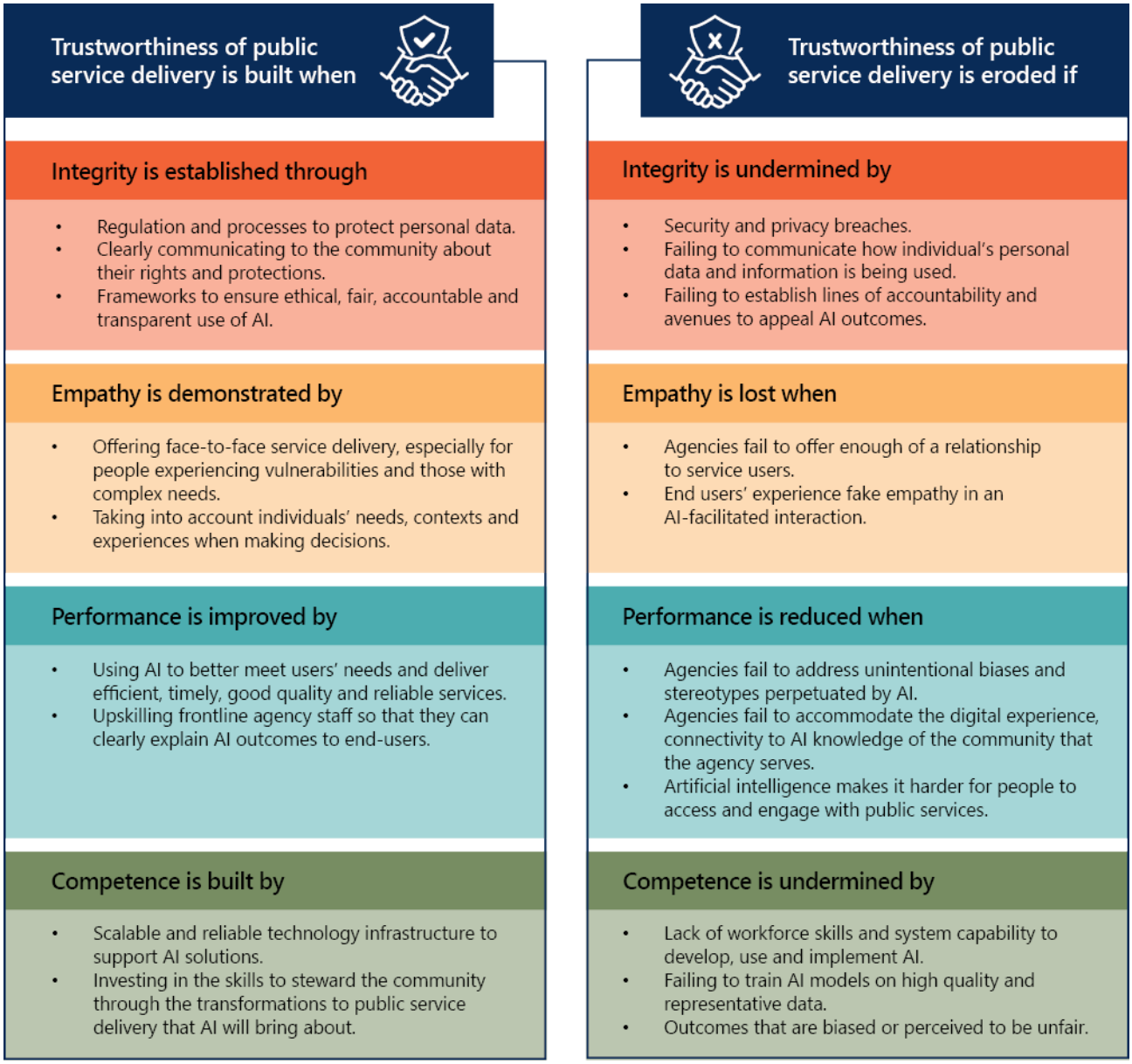
In just five years, AI and its capabilities have developed rapidly – from experimental applications to AI solutions and applications that are widely adopted and used across society – fuelled by an increase in computing power, investment and consumer demand. This rapid rate of change is expected to continue in the future.

This will pose a challenge for an APS seeking to realise the benefits of AI for public service delivery. A survey undertaken for this Long-term Insights Briefing found that people who are more familiar with and knowledgeable about AI have higher trust in government to responsibly use AI for public service delivery. However, more than half of people (57%) have zero or slight knowledge of AI, while almost two thirds (63%) have zero or slight understanding of when AI is being used.[3] Most people will gain a greater knowledge of AI and when it is being used over time, thanks to the increasing pervasiveness of AI technology in their daily lives. However, the APS will always be engaging with people who lack knowledge of and familiarity with the latest tools. Realising the benefits of AI will require the APS to steward the community through the transformations that AI will bring to how public services are designed, implemented, delivered and explained. This stewardship is necessary to ensure that AI contributes to the delivery of high quality public services and that the risks outlined above are well managed.

Framework for trustworthy use of AI in public service delivery

To build trustworthiness, government agencies need to deliver public services well, by meeting users' needs and delivering efficient, timely, good quality and reliable services. There are significant opportunities for agencies to use AI to do this. Drawing on research, survey evidence, and consultations with community representatives and AI and service delivery experts, this briefing offers insights into how agencies can design, develop and implement AI in public service delivery in ways that build trustworthiness (Figure 1).

Figure 1. Framework for trustworthy use of AI in public service delivery



Source: Summary of insights from workshops and focus groups held for this briefing, involving people from 15 organisations representing the community, 9 organisations representing academia, industry and youth, and 16 APS Agencies.

Insight 1: Artificial Intelligence must be designed and implemented with integrity

If the community does not trust AI, and the APS still uses it within a service offering, the APS may itself be seen as untrustworthy. Implementing AI in public service delivery well, in ways that demonstrate and build trustworthiness, critically depends on establishing and acting with integrity. This means that the people and organisations employing AI are accountable for the outcomes of the AI that they use, and transparent about how AI is being used; practising ethical values and principles when designing, developing and implementing AI; and ensuring personal privacy and data security.

At the same time, AI regulation and frameworks will only build trustworthiness if they are clearly communicated and explained to the community, including how they work in practice and the protections they provide.

Insight 2: Using artificial intelligence shouldn't come at the expense of empathy

AI will increase the trustworthiness of public services if it is designed and implemented in a way that demonstrates empathy. Trustworthiness is built when the APS demonstrates empathy for the people it serves. In practice, this means providing enough of a relationship with public services, where what that looks like depends on an agency's trust history, the community it serves, and the type of service it offers. A relationship with frontline staff is likely to be particularly important for people experiencing greater vulnerabilities and those with more complex needs.

Insight 3: Artificial intelligence should improve performance

AI will significantly erode trustworthiness if its use reduces the performance of public services.

Using AI in public service delivery will erode trustworthiness if end users have a poor experience with the service they are seeking. This might be the case if AI perpetuates unintentional biases and stereotypes (as a result of being trained on biased datasets or failing to include diverse perspectives in the design process); or makes it harder for people to access and engage with the services. New skills and capabilities will be needed across the APS if it is to adopt and use AI in ways that improve public services, and steward the community through the transformations to public service delivery that AI will bring about.

Insight 4: Successful service delivery depends on supporting people to engage with AI-enabled services in the long term

Maintaining trustworthiness requires the APS to deliver services to the whole community. Public services are for everyone, including those who don't want to engage with digital and AI-enabled systems or provide additional personal data, to ensure that using AI in public service delivery doesn't entrench disadvantage. Nevertheless, in the long term, opting-out will not be an option in a more connected world, where AI will be critical to address future challenges. It will be important to invest in building the AI literacy and digital connectivity of the community, particularly cohorts experiencing vulnerability and those that support them, in order to bring everyone along on the AI adoption process.

Introduction

The Long-term Insights Briefings

This is the first in an enduring series of public service-led Long-term Insights Briefings. The Long-term Insights Briefings are an initiative under Priority Two of the APS Reform Agenda: An APS that puts people and business at the centre of policy and services. They will strengthen policy development and planning in the APS by:

- Bringing together and helping the APS to understand the evidence and implications of long-term, strategic policy challenges.
- Building the capability and institutional knowledge of the APS for long-term thinking, and position the APS to support the public interest now and into the future, by understanding the long-term impacts of what the APS does.

The purpose of the briefings is not to make recommendations or predictions about what will happen in the future. Instead, they will provide a base to underpin future policy thinking and decision making on specific policy challenges that may affect Australia and the Australian community in the medium and long term. It is anticipated they will form part of the evidence base for policy and decision making.

Importantly, the briefings will be developed through a process of genuine engagement with the Australian community on issues affecting them, as well as with experts from the APS, academia, industry and the not-for-profit sector.

The first Long-term Insights Briefing

The first Long-term Insights Briefing explores how the APS could integrate AI into public service delivery in the future, and how this might affect the trustworthiness of public service delivery (Box 1). It complements Australian Government policy work on AI that is currently developing advice on how AI is best governed in the public service and the broader economy so that its many benefits can be realised while maintaining public trust. This includes the AI use in Government Taskforce, jointly led by the Digital Transformation Agency (DTA) and the Department of Industry, Science and Resources (DISR), and the Government's *Safe and Responsible AI in Australia* public consultation.

Rapid developments in the capabilities of AI applications such as ChatGPT[4] (and the speed and scale of adoption since its launch in November 2022) are provoking public conversations about the role that AI will have in Australian society.[5] Given the potential for AI to transform public service delivery in ways that deliver a better experience and outcomes for the whole community, it will be important for the APS to implement AI in public service delivery in ways that demonstrate and build trustworthiness. This is because implementing AI poorly – such as by failing to address known risks of the technology, or failing to understand and respond to the concerns of different cohorts in the community – could

erode the trustworthiness of the public service. This could result in the APS and the community as a whole failing to capture the benefits of AI.

Box 1: Key concepts in this Long-term Insights Briefing

This briefing explores how AI could change government decision-making systems and processes, what is needed to ensure AI leads to more effective service delivery, and the potential impacts of these changes on the trustworthiness of public services.

Artificial intelligence

There is no single agreed definition of AI. For example, the DISR discussion paper on *Safe and responsible AI in Australia*, defines AI as "an engineered system that generates predictive outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives or parameters without explicit programming. AI systems are designed to operate with varying levels of automation."[6]

The Australian Government Architecture defines AI as a catch-all term for a "family of technologies that can bring together computing power, scalability, networking, connected devices and interfaces, and data. AI systems can be programmed to perform specific tasks such as reasoning, planning, natural language processing, computer vision, audio processing, interaction, prediction and more. With machine learning, AI systems can improve on tasks over time according to a set of human-defined objectives and can operate with varying levels of autonomy."[7]

In referring to AI, this briefing is consistent with both of these definitions.

Given how rapidly AI is evolving, this Long-term Insights Briefing does not make predictions about the future potential of AI technologies. Instead, it considers the broader factors likely to influence the APS as it considers how to adopt and use AI in the next five to 10 years (and successive waves of the technology).

Trustworthiness

In contrast to the concept of 'trust', which is an attitude taken by an individual towards someone or an organisation or institution (like the APS), trustworthiness centres on the attributes of an organisation itself. In the context of this briefing, those attributes are:

- Integrity: adhering to a set of principles that the community finds acceptable, in terms of both words and actions.
- Competence: the ability to provide public services as promised.
- Performance: meeting public expectations and delivering reliable and consistent services.

- Empathy: identifying, understanding and responding to individuals' needs, contexts and experiences.

This reflects both existing frameworks of trust (including the Mayer ABI (Ability, Benevolence and Integrity) Framework[8] and the OECD Framework on Drivers of Trust in Public Institutions[9]) and views expressed by community representatives in workshops held over the course of the briefing, and by respondents to surveys.

The long-term insights in this report were informed by community engagement, input from AI and service delivery experts, and research and survey evidence. This included:

- workshops and focus groups involving people from 15 organisations representing community voices, 9 organisations representing academia, industry and youth, and 16 APS Agencies[10]
- a scenario-based workshop facilitated by the Australian National University's National Security College Futures Hub, which brought together community representatives, AI experts and APS service delivery staff to consider and respond to various scenarios of plausible futures and identify their likely impact on the trustworthiness of public service delivery
- over 5000 responses to 2 different surveys
- a Rapid Evidence Assessment of available literature and research conducted by the ANU.

Landscape

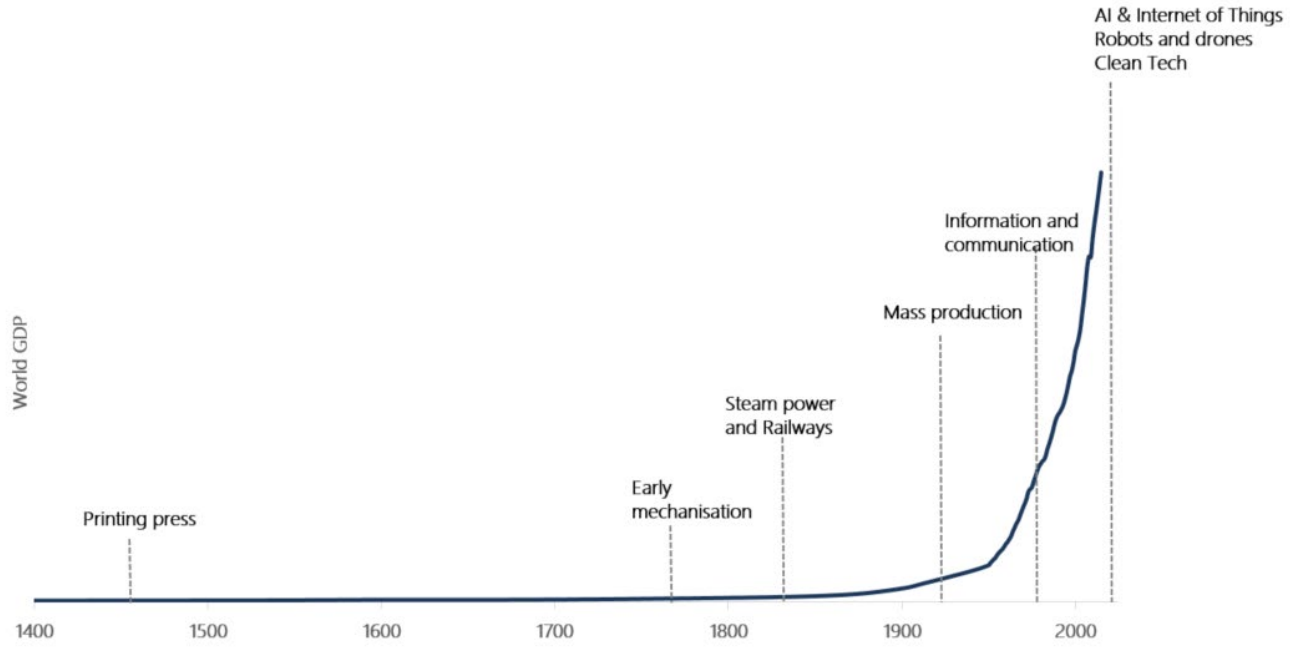
Artificial intelligence

Artificial intelligence is rapidly evolving

Technological innovations have driven world GDP growth over the past millennium, and with each wave of technological advancement, the way humans live, work and relate to each other has changed (Figure 2). This trend will continue with AI.

Indeed, the pace of change in AI has been substantial. In just five years, AI and its capabilities have undergone a rapid evolution – from experimental applications to AI solutions and applications that are widely adopted and used across society. This has been fuelled by an increase in computing power, investment and consumer demand. AI is expected to be worth $22.17 trillion to the global economy by 2030.[11] For Australia, digital innovations including AI, could cumulatively contribute $315 billion to the Australian economy by 2030.[12]

Figure 2. Technological advancements drive GDP growth and pace of change in accelerating

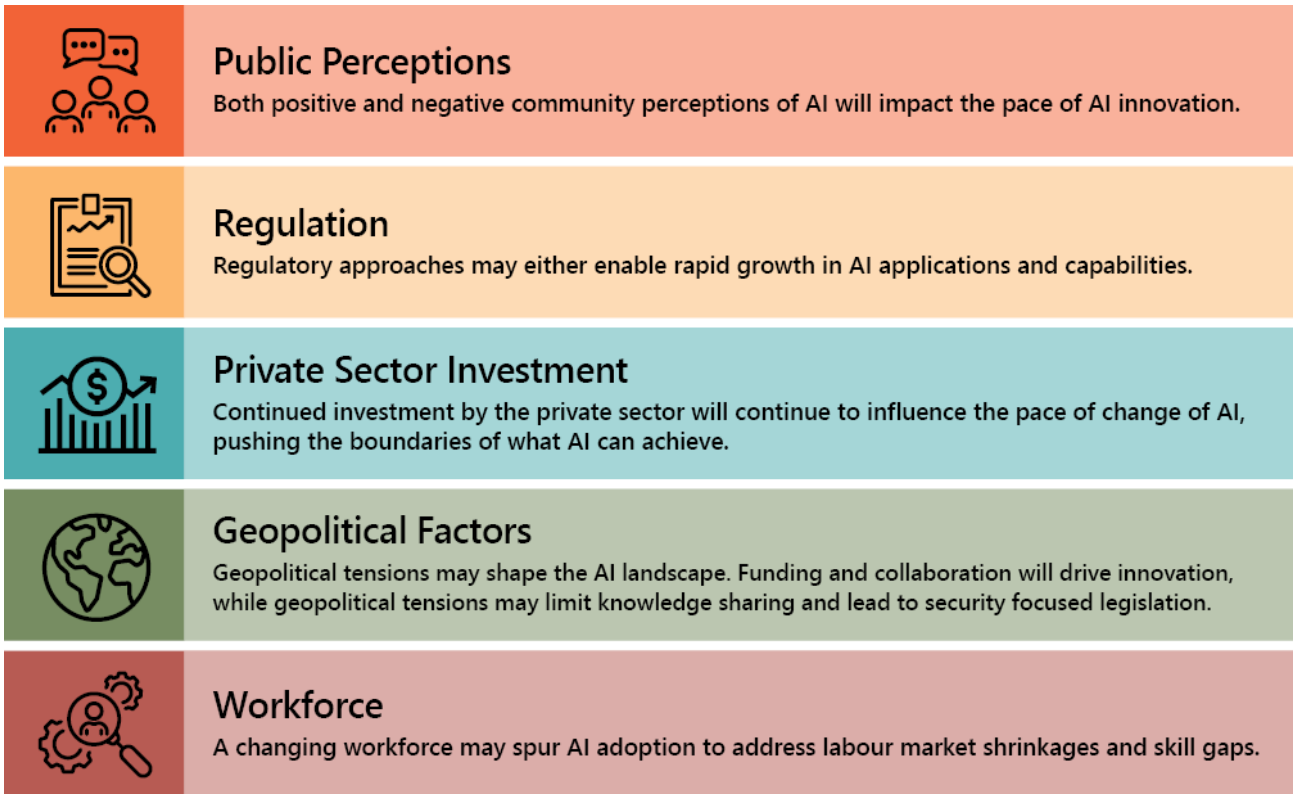


Source: Based on Our World In Data, *A long-term timeline of technology*. Output of world economy overlaid with key technological advancements.

Drivers of change

Given this rapid pace of change, experts caution against making predictions about the future potential of AI technologies.[13] Nevertheless, there are a number of factors that are likely to shape how AI is developed and used over the coming decade (Figure 3).

Figure 3. Foreseeable drivers of AI's development, adoption and use



Public perceptions: Positive perceptions of AI may drive future developments in AI systems by influencing demand, investment and adoption of AI technologies in the public and private sectors, fostering innovation and growth in the field. Conversely, negative perceptions, such as concerns about data security and privacy issues, may prompt a slower pace of AI development, implementation and use (Box 2).

Regulation: Regulation, both domestic and international, may either enable a rapid change of pace with AI by providing confidence to investors, developers and users, or slow development, implementation and use of AI by introducing complexity, uncertainty and red tape. Innovation will be shaped by legislative requirements and regulation, and influenced by voluntary ethical standards and frameworks. In Australia, there is already considerable work underway in this space.[14] The borderless nature of the internet and digital platforms will challenge the effective enforcement of regulations and standards across jurisdictions.

Private sector investment and digital capitalism: The private sector will drive change in AI by investing in R&D, fostering innovation through start-ups and large technology companies, and deploying

bespoke AI solutions that solve real-world problems. Platform-based, data-driven, and AI-powered businesses will become ever more central to economies and societies and data will increasingly emerge as a pivotal and core valuable resource. While this may lead to advancement in AI algorithms, hardware and data infrastructure, pushing the boundaries of what AI can currently achieve, it is also likely that a handful of corporations will become dominant market players. There is potential for this market power to significantly alter economic drivers for both consumers and competitors. Trust in use of data will become paramount, with any mishandling or unauthorised use of data eroding trust including in the AI systems that rely on such information.

Geopolitical factors: Geopolitical tensions may result in restrictions on the sharing of AI-related knowledge and technologies, affecting international collaboration. Further, large cyber security incidents in other jurisdictions may spill over into Australia's digital environment, putting AI systems and tools at risk. As cyber threats become more sophisticated, the adoption and integration of AI-driven cybersecurity measures are likely to increase to protect sensitive information and critical infrastructure.

Workforce: Labour and skills gaps globally are an important consideration for developing and maintaining AI solutions. Emerging skill requirements have resulted in demand and supply issues across a range of industries. Changing workforce needs, combined with labour force shrinkage overall will increase demand for automation and AI-powered solutions to improve efficiency and fill gaps left by labour shortages. The changing demographic landscape and workforce dynamics are likely to push the private sector toward greater AI integration to ensure sustained productivity and operational continuity. This, in turn, will drive innovation advancements in AI itself.

Box 2: Opportunities and risks of AI

Public perceptions of AI are likely to reflect awareness and understanding of the opportunities and risks of AI, including:

Opportunities:

- Automation: AI-driven automation can increase efficiency by handling repetitive tasks, freeing up time for more creative and complex work.
- Healthcare advancements: AI can contribute to better diagnosis, treatment, remote health care opportunities, and drug discovery, potentially revolutionising healthcare.
- Enhanced decision making: AI's data analysis capabilities can assist in making informed decisions across various industries.
- Improved customer experience: AI powered tools like chatbots can provide personalised customer support, enhancing user satisfaction.
- Efficient resource management: AI can optimise resource allocation.

Risks:

- **Bias and fairness**: AI systems can inherit biases present in their training data, leading to unfair or discriminatory outcomes.
- **Privacy concerns:** The extensive collection and analysis of personal data for AI can raise privacy and security issues.
- **Unintended consequences:** Complex AI systems might behave unpredictably, causing unintended outcomes at scale that are hard to control.
- **Dependency and reliability:** Overreliance on AI could result in systems breaking down or causing disruptions when they fail.
- **Inaccuracies**: The outputs of AI models can be entirely erroneous, or simply misleading (known as hallucinations).
- **Job displacement**: AI could lead to a change in skills required, affecting employment opportunities for some.

Public Service Delivery

The APS will need to innovate to meet community expectations of public services in the future

Public services in the future will be shaped by changing societal norms and heightened community expectations of government around the quality of services, particularly as Australia's population ages. At the same time, external forces such as climate change are expected to increase demand for services while decreasing the resources (people and funding) available to provide them (Figure 4).

Figure 4. Foreseeable drivers of public services in the future

Drivers of change

Community expectations of public services: The Australian community's expectations of public services are changing. Ongoing digital transformation is expected to continue to raise incomes and quality of life.[15] Expectations around the quality of services have increased. Many in the community expect greater responsiveness, convenience and efficiency when accessing services. For example, online platforms and streamlined processes continue to minimise wait times and enhance accessibility.

Demographic change: Changes in Australia's population will affect demand for public services (Box 3). Australia's population is ageing, and demand for quality care and support services is expected to grow as the share of older Australians in the population increases. Public services are expected to adapt by focussing on elder care, accessible health care options and tailored social programs, while still providing existing services. Increased immigration may require agencies to provide resources that cater to the specific requirements of immigrants. At the same time, an increase in the share of older Australians in the population will mean fewer working-age Australians to help fund more government services.

Box 3: Australia's demographic changes

Population growth is projected to increase by 3.9 million people over the next 10 years to 2033.

In 2022–23, approximately 4.6 million people (17% of Australia's total population), were aged 65 and over. In 2032–33 this is projected to increase to approximately 5.9 million people (19% of Australia's total population).

The number and percentage of older Australians is expected to continue to grow. By 2062–63, it is projected that older people in Australia will make up 23% of the total population.[16]

A growing care and support economy: Social changes, including the steady expansion of women's workforce participation have led to a shift away from unpaid care provision, towards formal care arrangements, including for children and ageing parents.[17] Expectations around the standard of care (how and where care is provided) have increased.

Climate change and environmental pressures: A warming climate and more frequent and intense natural disasters associated with climate change will affect the capacity of the APS to deliver public services. Demand for services will likely experience peaks in response to climate-related events that cause acute increases in community needs. Further, the resources needed to respond to climate impacts will likely draw resources from elsewhere in the public service.

Government agencies have already used automation to deliver better outcomes for users

Automated Decision Making (ADM) refers to the application of automation in any part of the decision-making process. It includes using automated systems to either:

- make the final decision
- make interim assessments or decisions leading up to the final decision
- refer to a human decision-maker where further information is required or an adverse decision is determined
- guide a human decision-maker through relevant facts, legislation or policy
- automate aspects of the fact-finding process which may influence an interim decision or the final decision.[18]

Currently, there are a number of ways that the public service can use automation to deliver tangible benefits such as efficiency. For example, Services Australia has applied automation technologies (not Artificial Intelligence) to process payments and deliver faster outcomes to Australians, so that they receive support when they need it most. Importantly, in those instances this happened in conjunction with Services Australia staff, not instead of staff, meaning that no one will have their claim refused through automation. Refusal decisions are made through manual processing by Services Australia staff.

There are also opportunities to use AI for public service delivery

AI capabilities are already being used in some Commonwealth government agencies.[19] For example the Australian Taxation Office (ATO) uses AI models to identify business activity statement lodgements with high-risk refunds prior to the refund issuing to taxpayers. This allows targeted treatment of risks such as identity crime, fraud and incorrect reporting. ATO staff will determine whether the refund can be issued based on currently held information, or further review is required. Where staff make a decision not to issue the refund, a taxpayer has the right to seek a review of the decision.

Some other ways that AI could be used in public service delivery now or in the near future include:

- Chatbots and virtual assistants: implementing AI powered chatbots on agency websites, communication channels, and for tasks such as claims processing can help agencies to provide citizens with quick responses to common inquiries and streamline customer service delivery.
- Data analysis and insights: AI can analyse large datasets to extract valuable insights for informed decision making.
- Automated data entry: AI can automate routine tasks, reducing administrative burden and minimising errors in data management.
- Language translation: AI-powered language translation tools can aid in providing multilingual services and facilitating communication with diverse populations.
- Fraud detection: AI can be used in document and image detection and recognition for border control, and for fraud detection in benefit applications.
- Minimise documentation errors: AI can check that customers or staff have submitted appropriate and error free documentation. This reduces duplication and speeds up processing times so as to improve service delivery.
- Automate claim processing: AI can reduce the time between requests for a government service and an outcome of that request being delivered. This is particularly relevant during times of emergency (e.g. the COVID Disaster Payment and Australian Government Disaster Relief Payment).

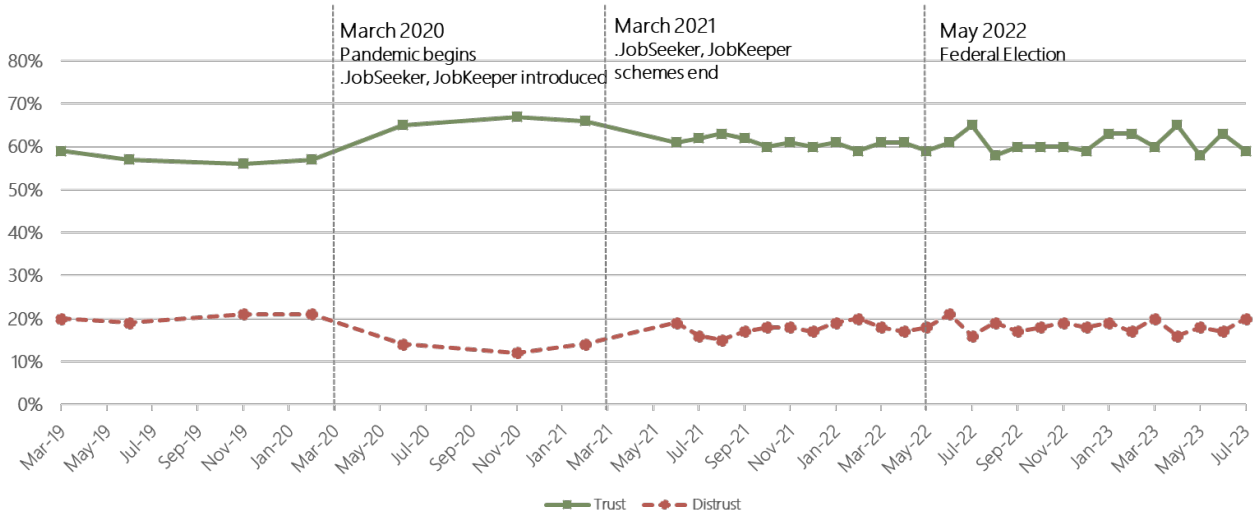
The state of trust in the public service's use of AI

Currently, most people (61%) trust Australian public services and believe they will change to meet Australians' needs in the future.[20] One way public services will change in the future is through the development and use of AI by the APS.

Drivers of trust in public services

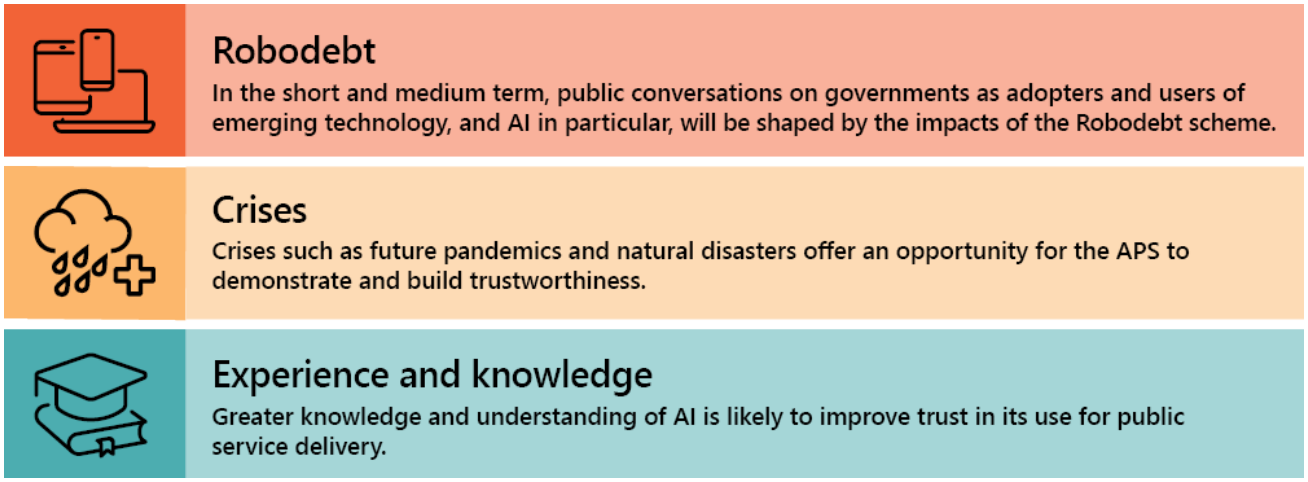
Crises: Recent events provide an insight into how trust in public services is affected by government actions (Figure 5 and Figure 6). During the COVID-19 pandemic, trust in public services increased significantly following the introduction of COVID-19 support programs like the JobSeeker and JobKeeper schemes,[21] and the Pandemic Leave Disaster Payments.

Figure 5. Trends in trust and distrust in public services



Source: Department of the Prime Minister and Cabinet (2022), Trust in Australian public services: Annual Report 2022, Trust in Australian public services: 2022 Annual Report | PM&C (pmc.gov.au).

Figure 6. Foreseeable drivers of trust in public services in the future



On the other hand, the community's response to the recent Robodebt scheme has demonstrated how poor policy design can undermine trust in government and public services,[22] specifically the social security system:

> *"…Scandals such as Robodebt undermine my confidence in the public service and government." Man, 35–44 (Have Your Say survey)*

While Robodebt calculations were completed via algorithmic decision making (ADM) rather than AI, for most people this distinction will be irrelevant. The community's views on public agencies as trustworthy adopters and users of digital technologies, including AI, are likely to be shaped by the impacts of the Robodebt Scheme, and the government's response to the findings of the Royal Commission report, for the foreseeable future.

Community's views on the use of AI in public service delivery

Australians views and trust in AI systems are a starting point for understanding how using AI in public service delivery might affect trustworthiness.

Overall, trust in AI systems is low in Australia

A recent global study of trust in AI by the University of Queensland and KPMG (UQ/KPMG) found that trust in AI systems is low in Australia.[23] Less than half of Australians (44%) believe the benefits of AI outweigh the risks. The study also found that only 35% of Australians agree that institutional safeguards, like regulations and legislative frameworks, are adequate.[24] However, the study did find that trust in Australia of AI systems has increased over time as the public gains more awareness of AI and its use in common applications.

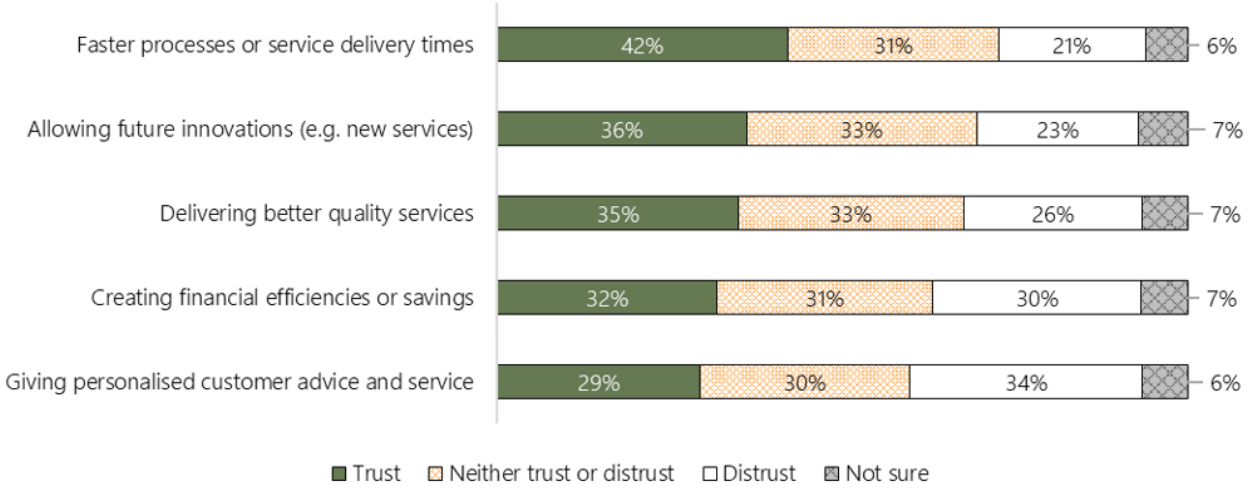
Knowledge and understanding of AI is low in Australia

A recent survey of *Trust in Australian Democracy* found that more than half of people (57%) reported having zero or slight knowledge of AI, while almost two thirds (63%) reported having zero or slight understanding of when AI is being used.[25]

Trust in government to responsibly use AI in public service delivery varies by the purpose for which AI is being used

When choosing between a list of purposes for using AI, people reported highest trust in government to responsibly use AI to deliver services faster (42% of respondents). Over a third indicated that they trust government to responsibly use AI to allow future innovations like creating new services. While personalisation is often discussed as one of the benefits that AI may offer for public service delivery, highest distrust was reported in government to responsibly use AI to provide personalised customer advice or services (34% distrust) (Figure 7). More than a third of respondents reported that they neither trust nor distrust government, or were unsure.

Figure 7. Trust in government to responsibly use AI in public service delivery varies by the purpose for which AI is being used



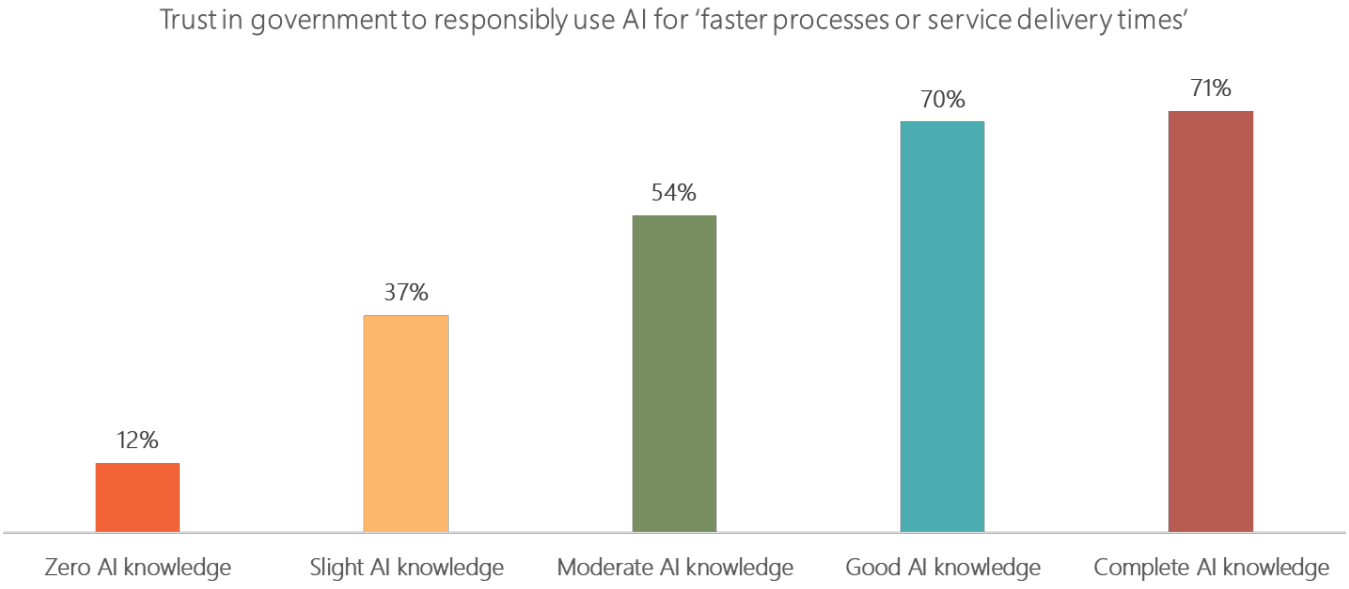
Notes: Trust results show percentage of people who said they 'trust' or 'strongly trust' government agencies to responsibly use AI for the outlined purpose; Distrust results show percentage of people who said they 'distrust' or 'strongly distrust' government agencies to responsibly use AI for the outlined purpose

Source: Australian Public Service Commission, Survey of Trust in Australian Democracy *(forthcoming)*

Higher knowledge of AI is associated with higher trust in government to responsibly use AI in public service delivery

Individuals who reported having higher knowledge of AI reported higher trust in government to responsibly use AI for any – or indeed, all – of the possible uses. For example, 70% of people who reported knowing 'very well' about AI trust government to responsibly use AI to deliver faster services, compared with 54% who reported a moderate knowledge of AI, and 37% who reported having a slight knowledge of AI (Figure 8).

Figure 8. Higher knowledge of AI is associated with higher trust in government to responsibly use AI in public service delivery



Notes: Results show percentage of people who said that they 'trust' or 'strongly trust' government agencies to responsibly use AI for providing 'faster processes or service delivery times". Results are cut by respondents' self-reported knowledge of AI from the options 'Not at all', 'Slightly', 'Moderately well', 'Very well', and 'Completely'.

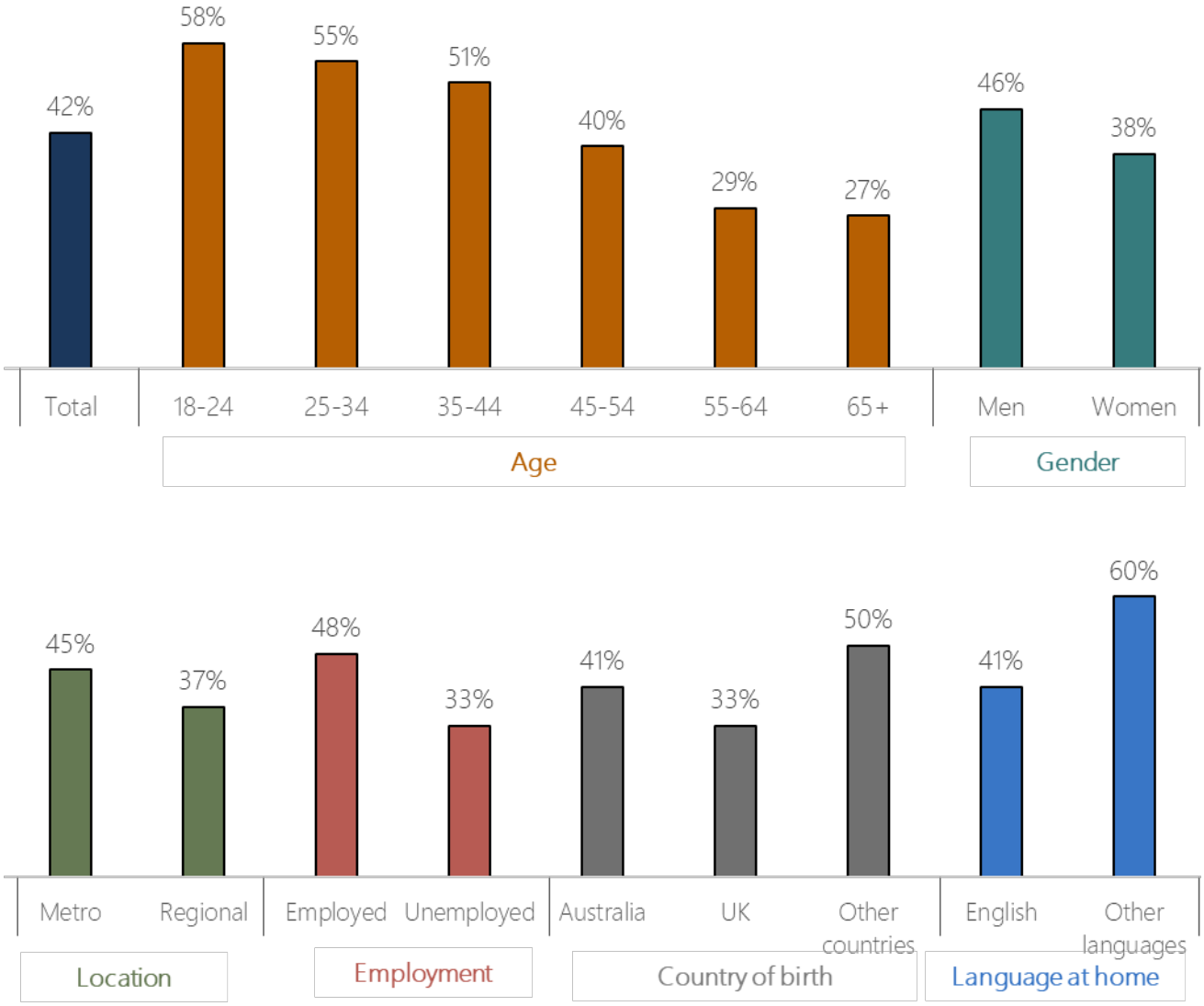
Source: Australian Public Service Commission, Survey of Trust in Australian Democracy *(forthcoming)*

Trust in government to responsibly use AI varies across cohorts

Trust in government's use of AI in public service delivery is shaped by social, cultural, economic and generational factors (Figure 9). For example:

- Younger people reported higher trust in government's use of AI than older people.
- Women reported lower trust in government's use of AI than men.
- People in regional Australia reported lower trust in government's use of AI than those in metro areas.
- People born in Australia reported lower trust in government's use of AI than those born overseas (with the exception of those born in the United Kingdom).
- People who speak English at home reported lower trust in government's use of AI than those who speak a language other than English at home.

Figure 9. Trust in government to responsibly use AI varies across cohorts



Source: Australian Public Service Commission, Survey of Trust in Australian Democracy *(forthcoming)*

To some extent, this may reflect different cohorts' experiences with public services and (self-reported) knowledge of AI. For example, women and people in regional areas trust public services significantly less than men and people in metro areas, respectively, even though their service satisfaction levels are similar.[26] Women and people in regional areas also reported having lower knowledge and understanding of AI and its applications.

In contrast, younger people (under 45) have higher trust in government's use of AI in public service delivery, despite tending to have lower trust in public services than people aged 65 and over[27] – a trend seen in most OECD countries.[28] This likely reflects their higher reported knowledge and understanding of AI and its applications.

Artificial intelligence and the trustworthiness of public service delivery

Artificial intelligence holds significant potential to improve the delivery of public services in Australia. Opportunities exist across the spectrum of activities carried out by the APS: from automating backroom administrative processes; to improving service processing and response times; to AI-enabled public interfaces that offer customised services to users.

Realising these benefits will require the public service to develop, adopt and use AI in ways that are trustworthy. If done well, with AI solutions that are designed and implemented in ways that demonstrate integrity, competence and empathy towards users, and deliver a clear improvement in the performance of public services, using AI could see the APS realise substantial efficiency gains and deliver a better experience and outcomes for the whole community. If done poorly, the APS will not only fail to capture the benefits of AI, there is a risk of eroding trustworthiness such that future adoption or use of AI is not accepted by the community. This could also affect service delivery innovation by the APS more broadly, compromising the ability of the APS to serve the Australian community.

Insight 1: Artificial Intelligence must be designed and implemented with integrity

Feedback received from the community suggests that improving the performance of public service delivery through the use of AI offers a significant opportunity to build agency trustworthiness

The performance of public services – how well the system meets user's expectations – is a critical attribute of a trustworthy agency. Survey respondents reported that meeting users' needs and delivering efficient, timely, good quality and reliable services, is important for them to trust public services.[29] To respondents, trust meant:

> *"…a reputation for quality, timely service delivery, engaging with empathy and following through on what [they] say [they] will do. Public services need to be reliable, consistent, contemporary and meet the needs of all Australians." Woman, 45–54 (Have Your Say survey)*

> *"The capacity of government to commit to and deliver high-quality public services that are sustainable and provide inclusive and genuine benefits." Man, 18–24, (Have Your Say survey)*

The community expressed a higher degree of trust in government to responsibly use AI to improve some aspects of the performance of public services than others. For example, 42% of survey respondents indicated that they trust government to responsibly use AI to deliver services faster.

However, trust was lower in government to use AI to provide personalised services. Similarly, community representatives indicated that they saw opportunities for governments to use AI to provide information faster in transaction type services, where people want to get what they want without needing to talk to a human. Community representatives also saw opportunities to simplify bureaucratic processes, optimise resource allocations, and reduce wait times by leveraging AI as a 'sidekick' for public servants.

AI systems need to be well-designed to improve public service performance

There was a strong view that public service design mattered – for AI-enabled public services and the APS as a whole:

> *[Trust in the context of public services means]…"That the services provided have been well designed with the end users' needs in mind" Woman, 65–74 (Have Your Say survey)*

Many community representatives argued that managing the risks and benefits of AI in public services can only realistically be achieved with design input from those who use those services. To achieve the potential benefits of AI-enabled service delivery, community representatives suggested that the developers of AI algorithms and systems would require a detailed understanding of lived experiences of the broader community, and the experience of specific groups, for example cohorts who experience vulnerability, that rely on service delivery in particular. Equally, community representatives recognised the risk of harms arising from the lack of diverse perspectives in AI design processes, noting that this could lead to unintentional biases and stereotypes being perpetuated.

Implementing AI in public service delivery well – in ways that demonstrate and build trustworthiness – critically depends on establishing and acting with integrity.

In the APS, integrity means:

> *…doing the right thing – both in 'what' we do and in 'how' we do it. Integrity is about demonstrating sound ethics and values through our work and our behaviour, and earning trust in our ability to act in the best interest of the Australian community.*[30]

Simply put, acting with integrity means adhering to a set of principles that the community finds acceptable, in terms of both words and actions.

The ANU Rapid Evidence Assessment of the literature suggests that there is significant overlap between factors that are important for the integrity of the public service and factors that ensure that AI systems are designed, developed and implemented with integrity:

- **Ethics and values:** AI systems should be designed and developed with ethical principles in mind, including the principle of to do no harm. In the context of public service delivery, this means mitigating the risks of bias and discrimination through managing data collection, storage, pre-processing, and algorithm design to ensure fairness in outcomes.

- **Accountability:** The people and organisations employing AI should be accountable for the outcomes and decisions of AI systems. In the context of public service delivery, this means understanding who is responsible for the AI's behaviour (developer, agency etc.), and having avenues for appealing outcomes.
- **Transparency:** Transparency in AI refers to the openness and comprehensibility of the AI systems' operations and decision-making processes. Transparency aids in accountability. In the context of public service delivery, services that involve AI should provide understandable, audience-appropriate explanations of decision-making processes and outcomes. Given that end users will have varying knowledge of AI – as outlined above, more than half of Australians (57%) report having zero or slight knowledge of AI – this means conveying an effective mental model of the AI system's decision process to an end user, even if they don't fully understand the internal workings. This may involve simplifying the true decision process to capture the most relevant factors and derive generalisable insights for general users, while ensuring that decisions can be fully explained in an accessible format where required, including for review and accountability purposes.[31]

Acting with integrity takes on a particular importance in the case of the APS' adoption and use of AI (Box 4). In part, this reflects some properties of AI and the need to address associated risks and concerns. Many types of AI are black box technologies – for example, generative AI like ChatGPT built on large language models and multimodal foundation models – making it difficult or even impossible for most people to understand how the model arrived at its outputs.[32] Addressing the risks involved in the use of generative AI tools in a government context, including the risks of erroneous, misleading or inappropriate outputs in response to prompts, means being clear when those AI tools are being used by government to inform activities. It also means ensuring that the bounds on the role of AI tools in decisions and outcomes are clearly communicated.[33] This includes that generative AI tools are not the final decision-maker on government services where there would be an adverse impact on individuals or organisations (unless or until those tools are sufficiently mature and transparent).

Box 4: Integrity is important in the context of AI

Like all technologies, AI can be used for positive or harmful purposes.[34] However, the nature of AI models, combined with their capacity to learn from vast and diverse datasets, can make predicting their behaviour challenging.

- AI is unique because of the scope of what is possible. It can take actions at a speed and scale that would otherwise be impossible (with implications for both benefits and harms).[35]
- There may be potentially unforeseen patterns in the data that make it hard to know precisely how AI will generate its decisions and outputs. The extent of this risk significantly depends on the model used, and its training data. Different models have different complexities, and if the training data is incomplete, biased or unrepresentative, the AI's behaviour can reflect those shortcomings, making it challenging to predict how it will behave in various situations.

- It is impossible to guarantee that the AI knows, and has learned, what was originally intended by the developer. The challenge of certifiability in AI lies in developing consistent and widely accepted criteria to assess and ensure the reliability, safety and ethical compliance of AI systems across various use cases. Given the diversity of AI models, applications, and data sources, creating a unified framework for certifying AI's performance and behaviour becomes complex, requiring careful consideration of technical, ethical and societal factors.

However, there is a range of tools that agencies can use to address these risks from a technical and/or procedural angle, to help agencies to develop, use and implement AI systems in trustworthy ways. This includes *Australia's AI Ethics Principles*;[36] the *OECD Principles on AI* and its associated guidance on tools for implementing trustworthy AI systems;[37] the Commonwealth Ombudsman's Automated Decision-Making Better Practice Guide[38] and the DTA guidance on public sector adoption of AI as part of its Australian Government Architecture, including the *Interim guidance for agencies on government use of generative Artificial Intelligence platforms*.[39] Agencies also have in-house tools and guidance, such as the ATO *data ethics principles*.[40]

Integrity was by far the most important dimension of trustworthiness for participants of workshops held with AI and APS experts and the community. Community representatives highlighted that AI systems and tools themselves need to be implemented in ways that demonstrate accountability, transparency and adherence to ethical values, in other words, with integrity.

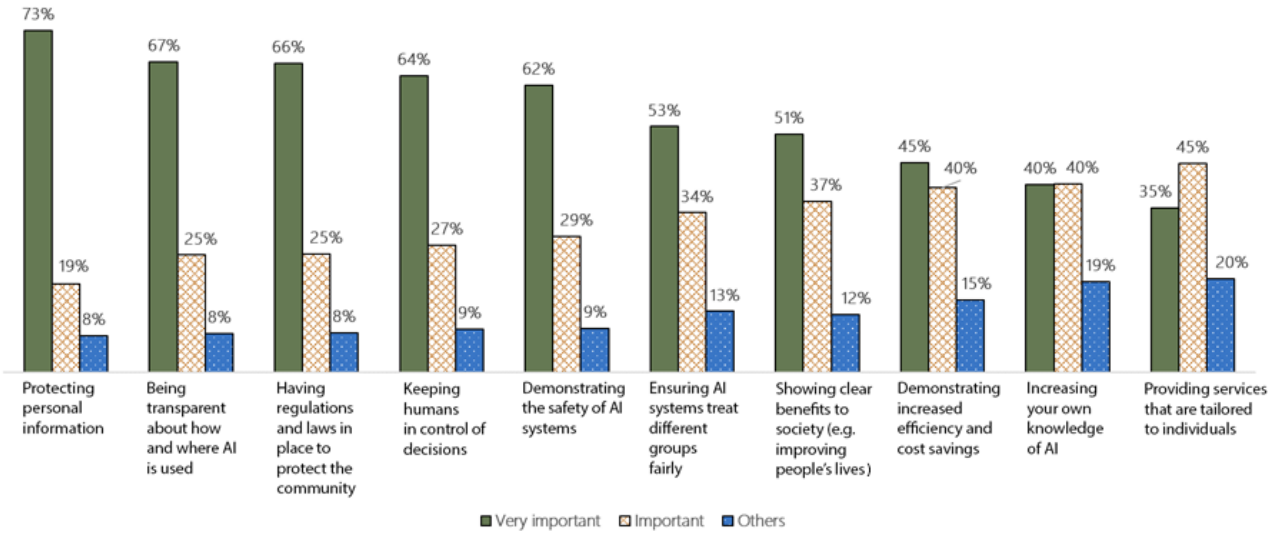
> *"I trust public service delivery when final decisions are transparent and adequately explained. AI can feel like a black box for people, it's not clear how it actually came to the conclusion." Workshop participant*

For example, there was a view that the increased scale and speed of an AI-enabled system would require increased oversight to ensure the integrity of the system, the integrity of the data, and the value to the public of the outcomes. Similarly, the *Survey of Trust in Australian Democracy* found that 3 in 4 Australians consider protecting personal information as very important; and around 2 in 3 consider that transparency in how/when AI is used and laws/regulations protecting community is also very important (Figure 10).

> *"Trust in the public service means that there is integrity within the services delivered. That people in the public service are acting with integrity to create and deliver services that improve all Australians and do not embed biases." Woman, 25–34 (Have Your Say survey)*

> *"Acting and making decisions in the best interest of the Australian public. Being able to believe that decisions made are considered and are made in good faith." Man, 55–64 (Have Your Say survey)*

Figure 10. Importance of factors in trusting government agencies to responsibly use AI in public service delivery



Notes: Results show importance given to each of the listed factors in trusting government agencies to use AI. Others include results for people who selected "Somewhat important", "Not important at all" and "Not sure".

Source: Australian Public Service Commission, Survey of Trust in Australian Democracy *(forthcoming)*

Personal privacy and data security are very important for the community

AI-enabled public services will rely heavily on data from the community, yet for many in Australia, use of personal data by AI is a significant concern. For example, a recent report by the Office of the Australian Information Commissioner (OAIC) finds that 43% of people in Australia say that AI using their personal information (in both businesses and government agencies) is one of the biggest privacy risks they face.[41]

For many workshop participants and survey respondents, the trustworthiness of a system that used personal data centred on knowing when data was being collected, knowing how data was going to be used, and being able to opt out of collection at any time (Future scenario 1). Participants emphasised that the trustworthiness of a system (or agency) would be drastically reduced if it could not guarantee that their personal data would not be shared with private organisations, or that it would be stored outside of Australia's borders. There was also some concern that there was no way to guarantee the integrity of future governments and how they may use people's personal information.

Future scenario 1: High personalisation and intervention

This scenario asked workshop participants to consider a future where AI integration with public services was entirely co-designed with community. In this scenario, a comprehensively co-designed model incurred a delay in implementation, but allowed for services to be delivered through tiered options. The top-tier required extensive collection of personal data and provided

a highly-personalised level, which included the option for automated intervention around health, finances and service offerings delivered through a single online platform. The baseline tier provided a greater level of personal privacy and agency, but it came without personalisation and a lower suite of service offerings, some of which could only be accessed via physical shopfronts in major cities. This scenario sought to:

- Understand better the desire within community for a co-designed system of AI-enabled public service delivery, and a willingness to invest in such a process, provoking discussion around potential levels of community engagement and ongoing participation in processes of oversight and integrity assurance.
- Have participants describe their level of willingness to provide high levels of personal data for higher levels of personalisation of service delivery within an AI-enabled system.
- Harvest participant responses around the concept of differing levels of public service delivery, which is based on the amount of data individuals choose to share.
- Have participants consider the possible permanency of data in an AI model once it has been provided, to provoke discussion around the potential for discrimination based on levels of privacy, geographic location and desire for human-in-the-loop service delivery.

We learnt that:

- How comfortable people are with sharing their data with government fundamentally depends on their individual circumstances.
- A highly personalised, fully integrated public service may bring benefits in certain situations, such as only having to tell your story to the government once. But, it may also come with risks in that a single point of access for services also becomes a single point of failure
- People value the ability to be forgotten, the ability to correct data, and the ability to take context into account in decisions.
- A sensitivity to creating unique experiences and a loss of a shared understanding and awareness of what public services are and how they engage with and support the Australian public.

There were also concerns about sharing of biometric data for identification purposes. For some, this was due to the risk that their identifying data might be used to impersonate them if breached or faked by AI applications. However, a majority of participants were concerned that it would be used to predict their behaviour. This finding reflects OAIC survey results, which indicates that the majority of Australians are not comfortable with biometric analysis, such as using AI to make assumptions or predictions about the characteristics of an individual from their biometric data.[42]

Data sovereignty is a priority for First Nations peoples

Participants in workshops highlighted the data concerns of First Nations people. They noted that many First Nations people are sensitive to the inherent right to self-determination and governance over their peoples, country (including lands, water and sky) and resources. There is a strong desire that data be used in a way that supports and enhances the overall wellbeing of Indigenous people. In practice, this may mean that Indigenous people need to be the decision-makers around how data about them is used or deployed, to build trustworthiness in AI systems and tools within the public sector (Box 5).

Box 5: Indigenous Data Sovereignty

Indigenous Data Sovereignty emphasises the rights of Indigenous communities to control, manage, and benefit from their own data. It acknowledges the historical marginalisation of Indigenous people and their data, and strives to empower these communities to make informed decisions about how their data is collected, used and shared. When Indigenous communities have ownership over their data and are able to co-design custom AI applications, it may help to foster a sense of respect and collaboration, ensuring that AI respects cultural sensitivities, local knowledge and community values. This approach ultimately enhances the credibility and ethical standing of AI systems, promoting a more inclusive and equitable adoption of AI in the public sector.[43]

Insight 1.1: Artificial intelligence regulation and frameworks will build trustworthiness if they are clearly communicated and explained to the community

In practice, the perceived as well as actual integrity of AI systems is important for trustworthiness. Participants in workshops described an expectation that there will be regulations and frameworks in place that ensure fair, accountable and transparent use of AI. Almost a quarter of respondents to the *Have Your Say* survey suggested that AI use be limited until guardrails like governance and ethics frameworks are established, risks are properly understood, and mitigation strategies and controls are in place.

> *"No public services until governance frameworks are established. Then no service or decision affecting an individual, no law enforcement, until frameworks are proven effective through use in low risk settings." Woman, 45–54 (Have Your Say survey)*

In fact, there are a large number of current Australian Government initiatives that are relevant to the development, application or deployment of AI (including frameworks that are specific to the public sector) although most take the form of self-regulation and voluntary standards approaches.[44] Work is also ongoing to identify potential gaps in the existing domestic governance landscape and whether additional AI governance mechanisms are required to support the safe and responsible development and adoption of AI.[45]

But those frameworks will only build trustworthiness if they are clearly communicated and explained to the community. There was not a strong awareness of initiatives, voluntary or otherwise, let alone understanding of how they work in practice and the protections provided. Similarly, there was low awareness of the roles and responsibilities of different regulators (both sector-specific and economy-wide). Accessible communication will be key, particularly in a crowded and contested regulatory space.

Insight 2: Using artificial intelligence shouldn't come at the expense of empathy

"AI cannot make human-centred decisions and cannot understand wants or be empathetic" (Workshop participant)

Emerging technologies such as AI may fundamentally shape not just how public services are designed, but how people interact with government. While this could lead to significant improvements in service performance (as outlined above) there is a view that it could come at the expense of empathy for individuals.

> *"I would expect cost savings and time reductions but a depersonalising of interactions and a bureaucratic adherence to process." Man, 65–74 (Have Your Say survey)*

> *"Presently, AI is not well suited to provide tailored customer service to the public and will lack the nuance and empathy that the public are likely to expect when they try to engage with a service." Man, 25–34 (Have Your Say survey)*

AI-enabled public service delivery offers opportunities to improve performance by streamlining access to services for many in the community. While people are comfortable with using AI to improve process and operational efficiency, when it comes to applications where a decision maker has a lot of discretion in decision-making, some workshop participants and survey respondents questioned whether human complexity could be translated into an AI system.

> *"So in some ways I'm happy enough to sit by and do my tax online and have the little question box…So for a pretty simple response that's yes/no/go here, that's all fine. Of course it's helpful. But when you actually get into parts of the human condition that are intrinsically based in behaviour and complex behaviours…how do you actually translate that into an AI space?" Workshop participant*

> *"I highly value a great user experience through digital channels including on a self-service basis. However, I still also value the ability to talk to someone to escalate an issue or obstacle including delays or inability to find what I am looking for via digital channels…" Man, 35–44 (Have Your Say survey)*

A loss of human interaction in moments of need could significantly erode trust. Research literature indicates that a lack of interpersonal interaction with public service actors and decision-makers is a significant driver of distrust in AI, while community representatives noted that AI might create barriers to getting in touch with a real person in moments of need (Future Scenario 2).

Future scenario 2: Climate's influence on the system

In one scenario, workshop participants were asked to consider a future where an increase in climate driven events required a redistribution of public resources. Most routine and transactional services were automated to allow for greater focus on rapid support for people impacted by crisis such as floods, bushfires and cyclones. Human-to-human services were largely reserved for high-trauma and high-impact incidents. To receive automated services people needed to identify themselves using their biodata, those refusing to provide biodata would need to apply for and wait for in-person services to become available. This scenario helped us to understand:

- The willingness of people to access the majority of public services via AI-enabled digital platforms without engaging humans as part of the process, and gather insights on whether people are willing to trade potentially slower human-based service delivery for potentially faster AI-based service delivery.
- The willingness of participants to use biometric data as a means for personal identification, and gain a response to a proposition that those who do not wish to share biometric data will not receive as prompt a service as those who share biometric data as a means for identification.

We learnt that people:

- Understand the need to prioritise a response to those in crisis, but do not think that people who preferred to interact in person should receive a slower service.
- Are concerned about the level of service provided after a crisis and how local communities would be resourced for the longer term work of rebuilding resilience in a community after an adverse event.
- Are very concerned about the use of personal data (and biometric data specifically), even in a crisis; and the impacts on those who would opt out of sharing their personal data altogether.
- Feel sympathy for those who might want to opt out of sharing personal data because of their experiences and concerns that their past history might discriminate against them.
- See a risk of a greater social divide for those who are experiencing complexity and vulnerability, and need ongoing public services through human engagement.

Trustworthiness is built when the public service demonstrates concern and empathy for the people it serves

Traditionally, public service delivery has included face-to-face interactions, which provide an opportunity for public servants to recognise and respond to an individual's circumstances with empathy in the moment. Community representatives and experts in workshops highlighted the importance of those interactions, noting that many people desire to talk to, be seen by, and receive help from a real person who can empathise and offer a service suited to their unique circumstances.

> *"In order to promote trustworthy public services, the APS must ensure that its services are human-centred and responsive to human need and the nuance of diverse experiences. This includes by increasing accessibility, clarity, responsiveness, and transparency. It also requires clear and tangible demonstrations of ethical decision-making, fairness and commitment to the public interest." Written submission*

More generally, respondents to the *Have Your Say* survey highlighted the importance for trustworthiness of a public service that is seen to act in the best interests of Australians:

> *"Trust means people are putting their faith in the public service to do right by them." Man, 25–34 (Have Your Say survey)*

It is likely that in many instances, AI-augmented processes could help frontline staff to engage in stronger and more in-depth relationships with users than is currently the case. Again, workshop participants highlighted that design input from those who use public services could help ensure that AI-enabled services remain human centred, and even help agencies to become more humane and compassionate:

> *"Culturally and trauma-informed AI doesn't exist, but if we could get there in the future, that could be really special and make it easier" [for agencies and for the people that engage with them]. Workshop participant*

On the other hand, it was felt that AI is not capable of empathy. While AI's human-like capability can compel users to perceive an AI-system as a person,[46] this was seen to pose a significant risk to trustworthiness, with focus group participants noting that fake empathy from AI could completely destroy trust. Equally, there is a risk that decisions and outcomes that are seen to be lacking empathy are attributed to the use of AI – whether or not this is justified – reducing trust in government's ability to use AI responsibly in service delivery.

Insight 2.1: People want enough of a relationship with public services – what that looks like depends on an agency's trust history, the community it serves, and the type of service it offers

For some in the community, human interaction and a relationship with public services is as important as the service itself for demonstrating trustworthiness. For many people, interacting with services can

be confusing or overwhelming. Direct contact with a person can reduce some of that discomfort and make it easier for them to access and engage with public services. In contrast, other groups may not want or need to access public services in person — for some people, convenience may be much more important:

> *[Potential benefits of AI use in public service] "Shorter wait times, access to services when you need them (24/7, 7 days a week)…" Woman, 35–44 (Have Your Say survey)*

Considering the community as a whole, what matters is ensuring enough of a relationship. For a given agency, this will depend on several factors:

An agency's trust history…

Community representatives noted that how communities have been treated in the past and in times of crisis will determine future interactions and perceptions of trustworthiness. For example, responses to the *Have Your Say* survey highlighted how the Robodebt Scheme had undermined trust in public service agencies:

> *"Robodebt killed much of my trust in public services. Trust is gradually being redeemed through compensation and action, but it's too little, too late, for too many." Woman, 45–54 (Have Your Say survey)*

…the community an agency serves…

Community representatives emphasised that the attributes and actions that are seen to build an agency's trustworthiness are different for different cohorts of the community, and that cultural differences come into play. In particular, empathy dimensions of trustworthiness are likely to be particularly important for those experiencing greater vulnerabilities and those with more complex needs. Participants in workshops also emphasised the importance of person to person relationships – which participants suggested are the opposite of AI – for building trustworthiness with First Nations peoples, noting that First Nations peoples want to be deeply listened to and heard by governments.

> *"All of the things we are talking about with AI/tech developments are the complete opposite of what First Nations communities want. We know they want relationships, they want to be listened to deeply, they want to spend time with government, and have people live among their communities to understand their perspective and way of life." Workshop participant.*

…and the type of service an agency offers

Participants in workshops suggested that there were opportunities to use AI for transaction type interactions with the public service, as people just want to get what they need, making empathy much less important.

Insight 3: Artificial intelligence should improve the performance of public services

Poorly designed and implemented AI could reduce the performance of public services

Beyond demonstrating integrity and empathy in the design and delivery of AI-enabled public services, there are other factors that are important for trustworthiness – in particular, that use of AI does improve, or at least maintains the performance of public services. Performance may decline when:

- Individuals or agencies lack competence in developing and using AI.
- Agencies fail to incorporate diverse experiences and perspectives in AI design.
- Agencies fail to establish lines of accountability for AI outcomes. In particular, where there is a potential for adverse impacts on an individual, the reasoning underpinning the AI models used must be transparent and explainable, and appropriate human supervision be exercised, including having a human involved in an adverse decision.[47] The individual must also be able to functionally challenge the outcome.
- Agencies fail to accommodate the nature of the population and its digital experience and connectivity. For example, people in very remote areas are more likely to have mobile-only internet access (32.6% of people in very remote areas, compared with 10.5% nationally), which can hinder their ability to effectively access some government services.[48]

Data quality is critical for ensuring that AI improves public service delivery

The quality of an AI model's outputs is driven by the quality of its data, making it important that Agencies create, manage, use and maintain high-quality, accurate and representative datasets, and practice robust data governance practices.[49] The performance of public service delivery will decline if the data used to train and deploy AI models is poor. As noted above, AI learns from patterns and biases present in data. If the data used to train AI is incomplete, biased or unrepresentative, the AI's output can reflect those shortcomings.

Workshop participants also raised concerns around how data shared today might affect access to services tomorrow. For example, if individuals who had previously shared data chose to opt out of future sharing, would a system that had learned about them then continue to view them as the person they were prior to opting out? Similarly, they noted that if data collection systems were not comprehensive, which might be the case in the context of health, mental health and other serious life issues, then individuals might receive services or interventions that were not appropriate for their current circumstances.

New skills will be needed across the APS to steward the community through the transformations to public service delivery that will take place with AI

Digital technologies and data are transforming how the APS operates and delivers services. The APS will need a range of skills and capabilities if it is to adopt and use AI in ways that improve public

services, and steward the community through the transformations to public service delivery that AI will bring about. Various cohorts of the APS will require a level of upskilling and reskilling as their jobs evolve and change through digital transformation. For individuals, these include skills like critical thinking and creativity to use AI effectively. At an agency level, this includes leadership for how AI will be used as well as frameworks and processes for managing risks, data and establishing responsibilities for decisions about AI.

Communication skills will also be critical, to communicate effectively with both technical and non-technical audiences about AI. The recent OAIC *Australian Community Attitudes to Privacy Survey* found that 71% of Australians consider it essential that people are told that AI is being used.[50] This means that the communication skills and knowledge of agencies' frontline staff about how AI is being used will be important for ensuring end users receive and experience a better level of service. In particular, frontline staff will need to be able to explain the output of AI systems to others in a clear, understandable and audience-appropriate way.

Insight 3.1: Trustworthiness will be eroded if artificial intelligence makes it harder for people to access and engage with public services

AI-enabled public service delivery offers opportunities to streamline access to services for many in the community. However, for others, providing information, querying decisions and navigating through to human support may become much harder. Given that users of public services are often at a vulnerable point in their lives, the ease of their interaction with a service may determine if they ultimately access the support and assistance they are entitled to, or even seek support again.

> *"Rural people only ask for help once – when they really need it – and they don't ask again". Workshop participant*

Insight 3.2: Experiencing bias and discrimination significantly erodes trustworthiness

The risks of bias and discrimination with AI are well-known, including the implications of using biased datasets to train AI, and failing to include diverse perspectives in the design process. Workshop participants argued that trustworthiness would be lost if using AI perpetuated unintentional biases and stereotypes. They also noted that some people are unlikely to access a service again, having experienced a biased process or outcome.

However, experts in workshops noted that biases in AI can be detected and corrected, while the biases that exist in the current human-based system are sometimes impossible to detect and/or correct. Given this, failing to address biases and discrimination as they arise will significantly erode trustworthiness.

Insight 4: Successful service delivery depends on supporting people to engage with AI-enabled services in the long term

A trustworthy APS is united in serving all Australians, enabling the Australian government to provide security, drive productivity and jobs in the economy, improve citizens' experience of government, and deliver fair and equitable support where it is most needed.[51]

Insight 4.1: Seizing the opportunities of artificial intelligence should not undermine the premise of public services

> *"The question isn't what would I like AI to do for me. The key question is what do I want my government service to be? I want my government service to be responsive, understanding and able to provide solutions so that I can live a dignified life, no matter my age, or my ability." Workshop participant*

In the future, AI might allow for highly personalised public services, even offering a service before an individual knows they need it. However, this could require a system where individuals effectively opt in to more personalised services by sharing greater amounts of personal data (Future scenario 3).

Future scenario 3: AI making public services healthy

Workshop participants were asked to consider a scenario where health services were delivered through revolutionised technology to track and manage health at the level of the individual. AI embedded within third party provided wearable and smart devices was used to identify early signs of physical and mental illness, facilitating faster interventions and quicker recovery, reducing illness in society and lowering the cost of public health. Differing levels of services were delivered based on lived circumstances. This resulted in those who provided higher levels of data for earlier interventions creating less of a burden on the public health system and the national budget. Regional and remote populations were not able to participate in the scheme due to a lack of supporting digital infrastructure, resulting in data from cities being used to predict resourcing and development elsewhere across the nation. This scenario sought to:

- Present a situation where public service delivery differs based on access to digital infrastructure, driving discussion around disadvantage and biases in data collection and usage.
- Provoke discussion regarding how society might navigate a future where choices around engaging with technological solutions contribute to an individual's impact on the cost and delivery of public services for others.
- Test participant interest in highly personalised service delivery, to understand potential desire for personalised interventions (in this circumstance, for health benefits); and to provide an opportunity for participants to consider their willingness to provide high levels of highly personal data and biographic data to government agencies.

- Present a scenario where private industry was integrated into the delivery of public services, driving discussion around the difference between sharing data with government and private industry.

We learnt that people:

- Are concerned that societal divides and inequalities may increase, based on demographic factors and comfort with sharing data; and that already marginalised groups could fall through the cracks.
- Are sensitive to the existence of disadvantage and discrimination in the way public services are delivered, regardless of whether it is happening to them or others.
- Are concerned that cohorts who lack trust in public services could be penalised further through reduced access to services; and that providing access to services based on acceptance of technological solutions is coercive.
- Believe that a reliance on impersonalised, automated and data-enabled services would further erode trust among First Nations communities (where trust is already low), and drive greater disengagement with government.

A future where providing more personal data increases access to personalised services could entrench disadvantage and discrimination

Workshop participants expressed significant concerns about the impacts of a system where sharing more personal data or having better access to digital infrastructure determined the level and quality of services received. Participants suggested this would be viewed as a way to coerce people into providing their information – and penalise those who chose not to – as public service agencies sought to improve efficiency. There was also a view that the more privileged in society have the freedom to share more of their data (as they are less vulnerable to negative consequences and more likely to have the smart devices needed to collect data).

Importantly, workshop participants often identified a risk that AI-enabled public services would reinforce inequality, discrimination and disadvantage, even if they (or the community they represent) would be unaffected. In other words, people would lose trust in a system that was seen to disadvantage others (Future scenario 4).

Future scenario 4: Public private partnerships

A fourth scenario asked workshop participants to consider a potential future where developing large foundational models is difficult for risk averse organisations such as governments in liberal democracies. Falling behind the private sector in service delivery standards, government (in this future scenario) engages with large international corporations to access powerful foundational AI models in order to enhance experiences and make services efficient. Private companies would

utilise their powerful AI models to collect data and match users to public services and obligations, while the government covers the cost of public service delivery with public funding. In this future, different service providers use and store personal data in different ways, which impacts on the effectiveness of the services users are able to access. This scenario sought to:

- Have participants consider a future where private industry-owned AI models are integrated into the delivery of public services. This tested sensitivities of personal data being in the hands of private industry, and test sensitivities regarding the potential for personal data to travel outside of Australia's jurisdiction.
- Drive discussion around the differing levels of sensitivity in providing personal data across differing societal cohorts.
- Have participants consider the fairness of varying levels of public service delivery and discuss the influence on Australian society when people share differing experiences in what kind of services they receive as a citizen or resident of the one country.

We learnt that people:

- Want agency and control about the level of personal data they provide.
- Are concerned about equity, access to services, and privacy, and that being unwilling to share data may lead to some groups being excluded from the benefits offered by this scenario.
- Are concerned about the government sharing data with private companies; that these AI solutions could be developed in-house, or at the very least, that corporations are held to the same standards as government.
- Are sensitive around the security of personal data, how it may be used by future governments and if personal data would be stored outside of Australia's jurisdiction.

Insight 4.2: Some people will opt out of engaging with digital – including AI enabled – systems in the short to medium term.

The results of the *Survey of Trust in Australian Democracy* indicated that people who are more familiar with and knowledgeable about AI have higher trust in government to responsibly use AI for public service delivery. This suggests that for some in the community (e.g. older people, people in regional areas), the lower trust in government's adoption and use of AI will dissipate over time as people become more familiar with and knowledgeable about AI.

> *"While we are starting from a low trust base, familiarity with technology and people getting used to the technology can have an overall positive impact on trust" Workshop participant*

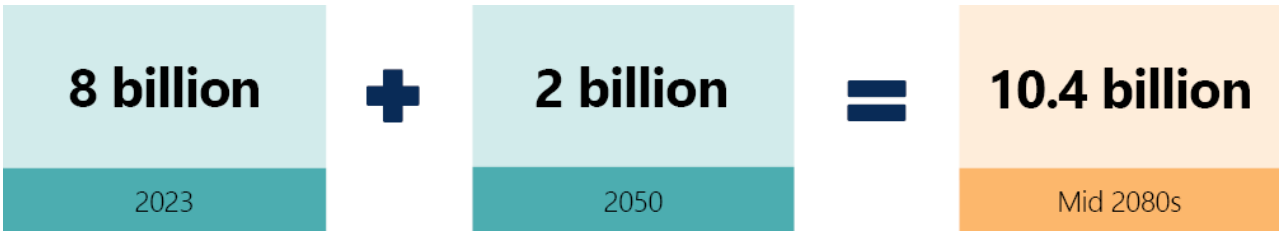
However, some – perhaps a significant share of people – will choose to opt out of engaging with digital systems, although participants in workshops did not agree on how important this could be.

Community representatives suggested that the share of the population who choose to opt out of engaging with public services (and with the government in general) would grow due to their concerns around the technology as well as the collection and security of personal data. For example, trust and cyber-safety are key concerns for many First Nations people, and can affect the extent to which they engage with digital technologies and government online services. In contrast, experts believe that there will be a small number of citizens who disconnect and disengage rather than share their data with AI technology. Experts suggested that these people would be difficult to win back from a trust perspective.

Insight 4.3: The success of service delivery in a more connected world will depend on bringing people along on the AI adoption process

In the longer term it will be necessary to invest in bringing everyone along on the AI adoption process.

The global human population has reached 8.0 billion people. It is expected to increase by nearly 2 billion people in the next 30 years, and could peak at nearly 10.4 billion in the mid-2080s.[52] Over 50% of people live in cities today. By 2050, it's projected that more than two-thirds of the world's population will live in urban areas, with 7 billion of the expected 9.7 billion people occupying cities globally.



Population densities are directly related to the development of smart cities. As the population expands, Governments need to find a way to support people effectively particularly in highly dense areas. Australia's major cities contribute nearly 80% of the national GDP.[53]

Data and digital technologies including AI can be part of the solution by delivering efficiencies, cutting red tape, providing better value for money and engaging citizens. Digital revolution brings opportunities for ground-breaking innovations in urban design, policymaking and infrastructure. In order for this to happen, citizens will need to be connected.

What is a smart city?

The OECD defines smart cities as "initiatives or approaches that effectively leverage digitisation to boost citizen wellbeing and deliver more efficient, sustainable and inclusive urban services and environments as part of a collaborative, multi-stakeholder process."[54]

Most people will gain a greater knowledge of AI and when it is being used over time, thanks to the increasing pervasiveness of AI technology in their daily lives. And in time, AI will be integrated into

many services offered by both the public and private sectors. Nevertheless, the public service will always be engaging with people who lack knowledge of and familiarity with the latest tools. Realising the benefits of AI will require the public service to steward the community through the transformations that AI will bring to how public services are designed, implemented and delivered. This stewardship is necessary to ensure that AI contributes to the delivery of high quality public services.

Supporting citizens to *opt in* to a more connected world would be supported by investment in:

- infrastructure and capital equipment at the individual level
- the AI and digital literacy of intermediaries to help individuals navigate the system
- community-based assets and organisations to support and mentor people, particularly those who are inclined to step away from services when they most need them, because of the rapid pace of change.

This investment in the fairness and equity of delivery of public services will improve familiarity with and knowledge of AI innovations among the community, and with it perceptions of trustworthiness of public service delivery.

[1] Commonwealth of Australia (2023), Intergenerational Report, Australia's future to 2063, accessed 29 August 2023
[2] Digital Transformation Agency, Adoption of Artificial Intelligence in the Public Sector, accessed 24 August 2023
[3] Australian Public Service Commission, Survey of Trust in Australian Democracy *(forthcoming)*
[4] As outlined in the National Science and Technology Council's *Rapid Response Information Report: Generative AI - language models (LLMs) and multimodal foundation models (MFMs)*, ChatGPT is an early example of the kinds of applications and services that will emerge from Generative AI built on LLMs and MFMs. Generative AI takes its name from its capacity to generate novel content, as varied as text, image, music and computing code, in response to a user prompt.
[5] Bell, G, Burgess, J, Thomas, J, and Sadiq, S (2023, March 24) *Rapid Response Information Report: Generative AI - language models (LLMs) and multimodal foundation models (MFMs)*, Australian Council of Learned Academies.
[6] Department of Industry, Science and Resources (2023), Supporting responsible AI: discussion paper consultation, accessed 29 August 2023
[7] Digital Transformation Agency, Adoption of Artificial Intelligence in the Public Sector, accessed 24 August 2023
[8] Mayer, R, Davis, J and Schoorman, F (1995) 'An integrative model of organizational trust', *The Academy of Management Review*, 20(3): 709–734, doi: 10.2307/258792
[9] Brezzi, M, González, S, Nguyen, D and Prats, M (2021) 'An updated OECD framework on drivers of trust in public institutions to meet current and future challenges', *OECD Working Papers on Public Governance*, No. 48, OECD Publishing, Paris, https://doi.org/10.1787/b6c5478c-en.
[10] 15 organisations representing the community: YWCA Australia, Office for Women SA, Federation of Ethnic Communities Councils of Australia, Community First Development, National Women's Safety Alliance, Equality Rights Alliance, Harmony Alliance, the National Rural Women's Coalition, Women with Disabilities Australia, National Older Women's Network, National Employment Services Association, St Vincent de Paul Society, InDigital, DVA (representing veterans voices), and the Australian Healthcare and Hospital Association. Nine organisations representing industry, academia and youth: Australian Medical Association, Amazon Web Services, ANU (School of Computing, College of Arts and Social Sciences, College of Science, College of Law), ANU (Youth), The Gradient Institute, National AI Centre, Australian Information Industry Association, IBM, and the University of Technology Sydney Human Technology Institute. 16 APS agencies: Department of Industry, Sciences and Resources, Digital Transformation Agency, CSIRO/Data 61, Office of the Chief Scientist, National Disability Insurance Agency, Australian Taxation Office, Department of Home Affairs, Department of Veteran Affairs, Services Australia, Department of the Treasury, Australian Public Service Commission, Department of Employment and Workplace Relations, Department of Foreign Affairs and Trade, Productivity Commission, and Australian Border Force.
[11] Commonwealth Scientific and Industrial Research Organisation (2023), Australia announces world first responsible AI Network to uplift industry, accessed 29 August 2023
[12] Department of Industry, Science and Resources, AI technologies | List of Critical Technologies in the National Interest, accessed 29 August 2023
[13] Bell, G, Burgess, J, Thomas, J, and Sadiq, S (2023, March 24) *Rapid Response Information Report: Generative AI - language models (LLMs) and multimodal foundation models (MFMs)*, Australian Council of Learned Academies.
[14] Department of Industry, Science and Resources (2023), Supporting responsible AI: discussion paper consultation, accessed 29 August 2023
[15] Commonwealth of Australia (2023), Intergenerational Report, Australia's future to 2063, p.10, accessed 29 August 2023
[16] Commonwealth of Australia (2023), Intergenerational Report, Australia's future to 2063, p. 240 accessed 29 August 2023
[17] Commonwealth of Australia (2023), Intergenerational Report, Australia's future to 2063, accessed 29 August 2023

[18] Department of Industry, Science and Resources (2023), Supporting responsible AI: discussion paper consultation, accessed 21 September 2023
[19] Digital Transformation Agency, Adoption of Artificial Intelligence in the Public Sector, accessed 24 August 2023
[20] Department of the Prime Minister and Cabinet (2022), Trust in Australian public services: Annual Report 2022, Trust in Australian public services: 2022 Annual Report | PM&C (pmc.gov.au).
[21] Department of the Prime Minister and Cabinet (2022), Trust in Australian public services: Annual Report 2022, Trust in Australian public services: 2022 Annual Report | PM&C (pmc.gov.au).
[22] Commonwealth of Australia (2023), Report of the Royal Commission into the Robodebt Scheme Report | Royal Commission into the Robodebt Scheme, p. 373
[23] Relative to BICS countries and Singapore. 34% of Australians are willing to trust AI. In Gillespie, N., Lockey, S., Curtis, C., Pool, J., & Akbari, A. (2023). Trust in Artificial Intelligence: A Global Study. The University of Queensland and KPMG Australia. doi:10.14264/00d3c94
[24] Gillespie, N, Lockey, S, Curtis, C, Pool, J and Akbari, A (2023) *Trust in Artificial Intelligence: A Global Study*, The University of Queensland and KPMG Australia. doi:10.14264/00d3c94
[25] Australian Public Service Commission, Survey of Trust in Australian Democracy *(forthcoming)*
[26] Department of the Prime Minister and Cabinet (2022), Trust in Australian public services: Annual Report 2022, Trust in Australian public services: 2022 Annual Report | PM&C (pmc.gov.au).
[27] Department of the Prime Minister and Cabinet (2022), Trust in Australian public services: Annual Report 2022, Trust in Australian public services: 2022 Annual Report | PM&C (pmc.gov.au).
[28] OECD (2022), *Building Trust to Reinforce Democracy: Main Findings from the 2021 OECD Survey on Drivers of Trust in Public Institutions*, Building Trust in Public Institutions, OECD Publishing, Paris, https://doi.org/10.1787/b407f99c-en.
[29] 40% of respondents to the *Have Your Say* survey indicated that trust to them meant service performance — more than any other option.
[30] Australian Public Service Commission (2021), Fact sheet: Defining Integrity | Australian Public Service Commission, accessed 29 August 2023
[31] Reid, A, O'Callaghan, S and Lu, Y (2023). Implementing Australia's AI Ethics Principles: A selection of Responsible AI practices and resources. Gradient Institute and CSIRO.
[32] Bell, G, Burgess, J, Thomas, J, and Sadiq, S (2023, March 24) *Rapid Response Information Report: Generative AI - language models (LLMs) and multimodal foundation models (MFMs)*, Australian Council of Learned Academies.
[33] Digital Transformation Agency, Interim guidance for agencies on government use of generative Artificial Intelligence platforms, accessed 21 September 2023.
[34] Department of Industry, Science and Resources (2023), Supporting responsible AI: discussion paper consultation, accessed 29 August 2023
[35] Department of Industry, Science and Resources (2023), Supporting responsible AI: discussion paper consultation, accessed 29 August 2023
[36] Alistair Reid, Simon O'Callaghan, and Yaya Lu. 2023. Implementing Australia's AI Ethics Principles: A selection of Responsible AI practices and resources. Gradient Institute and CSIRO.
[37] OECD (2023), OECD AI Policy Observatory, AI-Principles Overview - OECD.AI, accessed 20 September 2023. OECD (2021) 'Tools for trustworthy AI: A framework to compare implementation tools for trustworthy AI systems, *OECD Digital Economy Papers*, No. 312, OECD Publishing, Paris, https://doi.org/10.1787/008232ec-en.
[38] Commonwealth Ombudsmand (2007), Automated Decision Making, accessed 22 September 2023.
[39] Digital Transformation Agency, Adoption of Artificial Intelligence in the Public Sector, accessed 24 August 2023.
[40] Australian Taxation Office's data ethics principles, accessed 20 September 2023
[41] Office of the Australian Information Commissioner (2023) *Australian Community Attitudes to Privacy Survey 2023*, 8 August 2023 Australian Community Attitudes to Privacy Survey 2023 | OAIC

[42] Office of the Australian Information Commissioner (2023) *Australian Community Attitudes to Privacy Survey 2023*, 8 August 2023 Australian Community Attitudes to Privacy Survey 2023 | OAIC
[43] Walter, M and Kukutai, T (2018), *Artificial Intelligence and Indigenous Data Sovereignty*, input paper for the Horizon Scanning Project "The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve Our Wellbeing" on behalf of the Australian Council of Learned Academies, www.acola.org.
[44] Department of Industry, Science and Resources (2023), Supporting responsible AI: discussion paper consultation, accessed 29 August 2023
[45] Department of Industry, Science and Resources (2023), Supporting responsible AI: discussion paper consultation, accessed 29 August 2023
[46] Australian National University, Rapid Evidence Assessment, Finding 3.3
[47] For example, see the Australian Taxation Office's data ethics principles, accessed 20 September 2023
[48] Thomas, J, McCosker, A, Parkinson, S, Hegarty, K, Featherstone, D, Kennedy, J, Holcombe-James, I, Ormond-Parker, L and Ganley, L (2023) *Measuring Australia's Digital Divide: Australian Digital Inclusion Index: 2023*. Melbourne: ARC Centre of Excellence for Automated Decision-Making and Society, RMIT University, Swinburne University of Technology, and Telstra. Australian Digital Inclusion Index
[49] Reid, A, O'Callaghan, S and Lu, Y (2023). Implementing Australia's AI Ethics Principles: A selection of Responsible AI practices and resources. Gradient Institute and CSIRO.
[50] Office of the Australian Information Commissioner (2023) *Australian Community Attitudes to Privacy Survey 2023*, 8 August 2023 Australian Community Attitudes to Privacy Survey 2023 | OAIC
[51] Commonwealth of Australia (2019), Our Public Service Our Future, Independent Review of the Australian Public Service, p. 13, accessed 29 August 2023
[52] United Nations, Our growing population, accessed 29 August 2023
[53] Infrastructure Australia, State of Australian cities (2010), accessed 29 August 2023
[54] OECD (2019), Enhancing the contribution of digitalisation to the smart cities of the future, OECD Publishing, Paris, accessed 29 August 2023