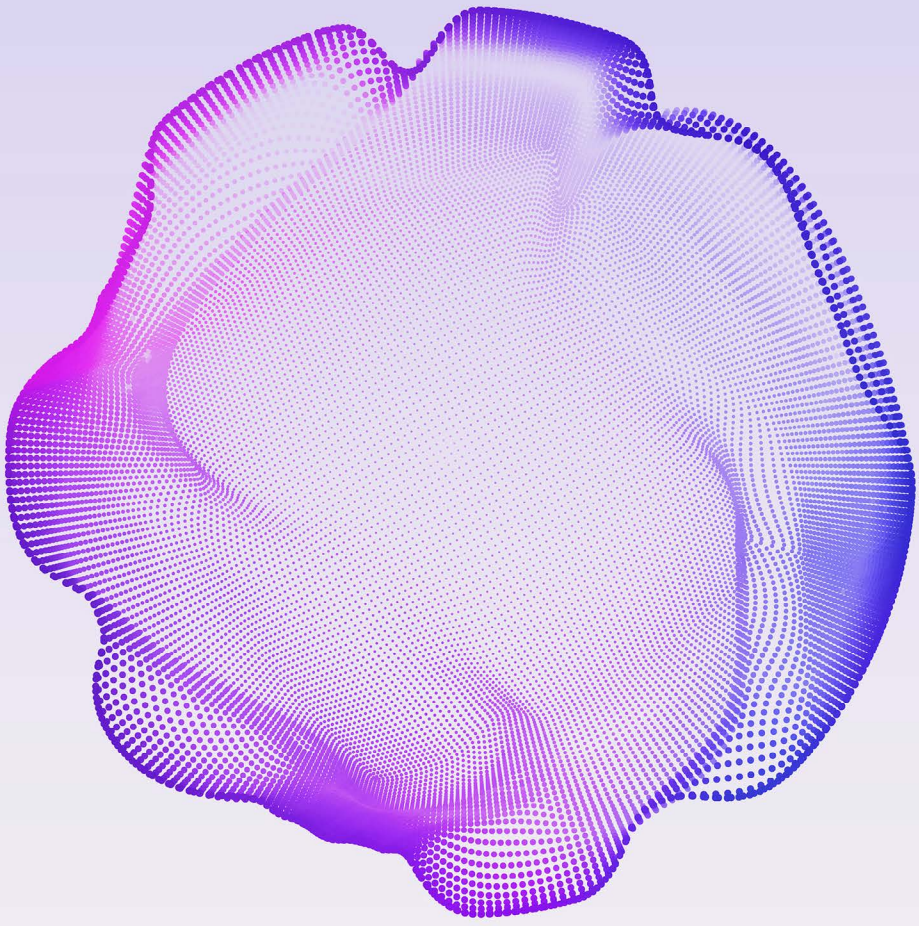
KPMG

The University Of Queensland
AUSTRALIA
CREATE CHANGE

Trust in Artificial Intelligence

A global study

2023



uq.edu.au

KPMG.com.au

Citation
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

University of Queensland Researchers
Professor Nicole Gillespie, Dr Steve Lockey, Dr Caitlin Curtis and Dr Javad Pool.

The University of Queensland team led the design, conduct, analysis and reporting of this research.

KPMG Advisors
James Mabbott, Rita Fentener van Vlissingen, Jessica Wyndham, and Richard Boele.

Acknowledgements
We are grateful for the insightful input, expertise and feedback on this research provided by Dr Ali Akbari, Dr Ian Opperman, Rossana Bianchi, Professor Shazia Sadiq, Mike Richmond, and Dr Morteza Namvar, and members of the Trust, Ethics and Governance Alliance at The University of Queensland, particularly Dr Natalie Smith, Associate Professor Martin Edwards, Dr Shannon Colville and Alex Macdade.

Funding
This research was supported by an Australian Government Research Support Package grant provided to The University of Queensland AI Collaboratory, and by the KPMG Chair in Trust grant (ID 2018001776).

Acknowledgement of Country
The University of Queensland (UQ) acknowledges the Traditional Owners and their custodianship of the lands. We pay our respects to their Ancestors and their descendants, who continue cultural and spiritual connections to Country. We recognise their valuable contributions to Australian and global society.

Contents

Executive summary

Artificial Intelligence (AI) has become a ubiquitous part of everyday life and work. AI is enabling rapid innovation that is transforming the way work is done and how services are delivered. For example, generative AI tools such as ChatGPT are having a profound impact. Given the many potential and realised benefits for people, organisations and society, investment in AI continues to grow across all sectors[1], with organisations leveraging AI capabilities to improve predictions, optimise products and services, augment innovation, enhance productivity and efficiency, and lower costs, amongst other beneficial applications.

However, the use of AI also poses risks and challenges, raising concerns about whether AI systems (inclusive of data, algorithms and applications) are worthy of trust. These concerns have been fuelled by high profile cases of AI use that were biased, discriminatory, manipulative, unlawful, or violated human rights. Realising the benefits AI offers and the return on investment in these technologies requires maintaining the public's trust: people need to be confident AI is being developed and used in a responsible and trustworthy manner. Sustained acceptance and adoption of AI in society are founded on this trust.

This research is the first to take a deep dive examination into the public's trust and attitudes towards the use of AI, and expectations of the management and governance of AI across the globe.

We surveyed over 17,000 people from 17 countries covering all global regions: Australia, Brazil, Canada, China, Estonia, Finland, France, Germany, India, Israel, Japan, the Netherlands, Singapore, South Africa, South Korea, the United Kingdom (UK), and the United States of America (USA). These countries are leaders in AI activity and readiness within their region. Each country sample is nationally representative of the population based on age, gender, and regional distribution.

We asked survey respondents about trust and attitudes towards AI systems in general, as well as AI use in the context of four application domains where AI is rapidly being deployed and likely to impact many people: in healthcare, public safety and security, human resources and consumer recommender applications.

The research provides comprehensive, timely, global insights into the public's trust and acceptance of AI systems, including who is trusted to develop, use and govern AI, the perceived benefits and risks of AI use, community expectations of the development, regulation and governance of AI, and how organisations can support trust in their AI use. It also sheds light on how people feel about the use of AI at work, current understanding and awareness of AI, and the key drivers of trust in AI systems. We also explore changes in trust and attitudes to AI over time.

Next, we summarise the key findings.

Most people are wary about trusting AI systems and have low or moderate acceptance of AI: however, trust and acceptance depend on the AI application

Across countries, three out of five people (61%) are wary about trusting AI systems, reporting either ambivalence or an unwillingness to trust. Trust is particularly low in Finland and Japan, where less than a quarter of people report trusting AI. In contrast, people in the emerging economies of Brazil, India, China and South Africa (BICS[2]) have the highest levels of trust, with the majority of people trusting AI systems.

People have more faith in AI systems to produce accurate and reliable output and provide helpful services, and are more sceptical about the safety, security and fairness of AI systems and the extent to which they uphold privacy rights.

Trust in AI systems is contextual and depends on the specific application or use case. Of the applications we examined, people are generally less trusting and accepting of AI use in human resources (i.e. for aiding hiring and promotion decisions), and more trusting of AI use in healthcare (i.e. for aiding medical diagnosis and treatment) where there is a direct benefit to them. People are generally more willing to rely on, than share information with AI systems, particularly recommender systems (i.e. for personalising news, social media, and product recommendations) and security applications (i.e. for aiding public safety and security decisions).

Many people feel ambivalent about the use of AI, reporting optimism or excitement on the one hand, while simultaneously reporting worry or fear. Overall, two-thirds of people feel optimistic about the use of AI, while about half feel worried. While optimism and excitement are dominant emotions in many countries, particularly the BICS countries, fear and worry are dominant emotions for people in Australia, Canada, France, and Japan, with people in France the most fearful, worried, and outraged about AI.

People recognise the many benefits of AI, but only half believe the benefits outweigh the risks

People's wariness and ambivalence towards AI can be partly explained by their mixed views of the benefits and risks. Most people (85%) believe AI results in a range of benefits, and think that 'process' benefits such as improved efficiency, innovation, effectiveness, resource utilisation and reduced costs, are greater than the 'people' benefits of enhancing decision-making and improving outcomes for people.

However, on average, only one in two people believe the benefits of AI outweigh the risks. People in the western countries and Japan are particularly unconvinced that the benefits outweigh the risks. In contrast, the majority of people in the BICS countries and Singapore believe the benefits outweigh the risks.

People perceive the risks of AI in a similar way across countries, with cybersecurity rated as the top risk globally

While there are differences in how the AI benefit-risk ratio is viewed, there is considerable consistency across countries in the way the risks of AI are perceived.

Just under three-quarters (73%) of people across the globe report feeling concerned about the potential risks of AI. These risks include cybersecurity and privacy breaches, manipulation and harmful use, loss of jobs and deskilling, system failure, the erosion of human rights, and inaccurate or biased outcomes.

In all countries, people rated cybersecurity risks as their top one or two concerns, and bias as the lowest concern. Job loss due to automation is also a top concern in India and South Africa, and system failure ranks as a top concern in Japan, potentially reflecting their relative heavy dependence on smart technology.

These findings reinforce the critical importance of protecting people's data and privacy to secure and preserve trust, and supporting global approaches and international standards for managing and mitigating AI risks across countries.

There is strong global endorsement for the principles of trustworthy AI: trust is contingent on upholding and assuring these principles are in place

Our findings reveal strong global public support for the principles and related practices organisations deploying AI systems are expected to uphold in order to be trusted. Each of the Trustworthy AI principles originally proposed by the European Commission[3] are viewed as highly important for trust across all 17 countries, with data privacy, security and governance viewed as most important in all countries. This demonstrates that people expect organisations deploying AI systems to uphold high standards of:

- data privacy, security and governance
- technical performance, accuracy and robustness
- fairness, non-discrimination and diversity
- human agency and oversight
- transparency and explainability
- accountability and contestability
- risk and impact mitigation
- AI literacy support

People expect these principles to be in place for each of the AI use applications we examined (e.g., Human Resources, Healthcare, Security, Recommender, and AI systems in general), suggesting their universal application. This strong public endorsement provides a blueprint for developing and using AI in a way that supports trust across the globe.

Organisations can directly build trust and consumer willingness to use AI systems by supporting and implementing assurance mechanisms that help people feel confident these principles are being upheld. Three out of four people would be more willing to trust an AI system when assurance mechanisms are in place that signal ethical and responsible use, such as monitoring system accuracy and reliability, independent AI ethics reviews, AI ethics certifications, adhering to standards, and AI codes of conduct. These mechanisms are particularly important given the current reliance on industry regulation and governance in many jurisdictions.

People are most confident in universities and defence organisations to develop, use and govern AI and least confident in government and commercial organisations

People have the most confidence in their national universities and research institutions, as well as their defence organisations, to develop, use and govern AI in the best interest of the public (76–82% confident). In contrast, they have the least confidence in governments and commercial organisations to do this. A third of people lack confidence in government and commercial organisations to develop, use and regulate AI. This is problematic given the increasing scope with which governments and commercial organisations are using AI, and the public's expectation that these entities will responsibly govern and regulate its use. An implication is that government and business can partner with more trusted entities in the use and governance of AI.

There are significant differences across countries in people's trust of their government to use and govern AI, with about half of people lacking confidence in their government in South Africa, Japan, the UK and the USA, whereas the majority in China, India and Singapore have high confidence in their government. This pattern mirrors people's general trust in their governments: we found a strong association between people's general trust in government, commercial organisations and other institutions and their confidence in these entities to use and govern AI. These findings suggest that taking action to strengthen trust in institutions generally is an important foundation for trust in specific AI activities.

People expect AI to be regulated with some form of external, independent oversight, but view current regulations and safeguards as inadequate

The large majority of people (71%) expect AI to be regulated. With the exception of India, the majority in all other countries see regulation as necessary. This finding corroborates prior surveys[4] indicating strong desire for regulation of AI and is not surprising given most people (61%) believe the long-term impact of AI on society is uncertain and unpredictable.

People are broadly supportive of multiple forms of regulation, including regulation by government and existing regulators, a dedicated independent AI regulator, and co-regulation and industry regulation, with general agreement of the need for some form of external, independent oversight.

Despite the strong expectations of AI regulation, only two in five people believe current regulations, laws and safeguards are sufficient to make AI use safe. This aligns with previous surveys[5] showing public dissatisfaction with the regulation of AI, and is problematic given the strong relationship between current safeguards and trust in AI demonstrated by our modelling. This highlights the importance of strengthening and communicating the regulatory and legal framework governing AI and data privacy.

There are, however, substantial country differences, with people in India and China most likely to believe appropriate safeguards are in place (74–80% agree) followed by Brazil and Singapore (52–53%). In contrast, people in Japan and South Korea are the least convinced (13–17% agree) as are the majority of people in western countries. These differences in the perceived adequacy of regulations may partly explain the higher trust and acceptance of AI among people in the BICS countries.

Most people are comfortable with the use of AI to augment work and inform managerial decision-making, but want humans to retain control

Most people are comfortable with the use of AI at work to augment and automate tasks, but are less comfortable when AI is focused on them as employees, for example for HR and people management (e.g. to monitor and evaluate employees, and support recruitment). On average, half the people are willing to trust AI at work, for example by relying on the output it provides. People in Australia, Canada, France and Germany are the least comfortable with the use of AI at work, while those in the BICS countries and Singapore are the most comfortable.

Most people view AI use in managerial decision-making as acceptable, and actually prefer AI involvement to sole human decision-making. However, the preferred option is either a 25%-75% or 50%-50% AI-human collaboration, with humans retaining more or equal control. This indicates a clear preference for AI to be used as a decision aid, and a lack of support for fully automated AI decision-making at work.

While about half believe AI will enhance their competence and autonomy at work, less than one in three people believe AI will create more jobs than it will eliminate. However, most managers believe the opposite – that AI will create jobs. This reflects a broader trend of managers being more comfortable, trusting and supportive of AI use at work than other employees, with manual workers the least comfortable and trusting of AI at work. Given managers are typically the drivers of AI adoption in organisations, these differing views may cause tensions in the implementation of AI at work.

A minority of people in western countries, Japan and South Korea report that their employing organisation invests in AI adoption, recognises efforts to integrate AI, or supports the responsible use of AI. This stands in contrast to a majority of people in the BICS countries and Singapore.

People want to learn more about AI but currently have low understanding

While 82% of people are aware of AI, one in two people report feeling they do not understand AI or when and how it is used. Understanding of AI is highest in China, India, South Korea, and Singapore. Two out of five people are unaware that AI enables common applications they use. For example, even though 87% of people use social media, 45% do not know AI is used in social media.

People who better understand AI are more likely to trust and accept it, and perceive greater benefits of AI use. This suggests understanding AI sets a foundation for trust. Most people across all countries (82%) want to know more about AI. Considered together, these findings suggest a strong need and appetite for public education on AI.

Younger generations, the university educated and managers are more trusting, accepting and generally hold more positive attitudes towards AI

Younger generations, the university educated, and managers show a consistent and distinctly more positive orientation towards AI across the findings, compared to older generations, those without a university education and non-managers. They are more trusting and accepting of AI systems, including their use at work, and are more likely to feel positive about AI and report using it.

They have greater knowledge of AI and are better able to identify when AI is used, and have greater interest in learning about AI. They perceive more benefits of AI, but remain the same as other groups in their perceptions of the risks of AI. They are more likely to believe AI will create jobs, but also more aware that AI can perform key aspects of their work. They are more confident in entities to develop, use and govern AI, and more likely to believe that current safeguards are sufficient to make AI use safe. It is noteworthy that we see very few meaningful differences across gender in trust and attitudes towards AI.

There are stark differences in trust and attitudes across countries: people in the emerging economies of Brazil, India, China, and South Africa are more trusting and accepting of AI and have more positive attitudes towards AI

A key insight from the survey is the stark differences in trust, attitudes and use of AI between people in the emerging economies of Brazil, India, China and South Africa and those in other countries.

People in the emerging economies are more trusting and accepting of AI and hold more positive feelings and attitudes towards AI than people in other countries. These differences held even when controlling for the effects of age and education. Singapore followed this positive orientation on several indicators, particularly comfort, trust and familiarity with the use of AI at work, adequacy of current AI regulation and governance, confidence in companies to use and govern AI, and the belief that AI will create jobs.

Our data suggests that this high trust is not blind to the risks. People in BICS countries and Singapore did not perceive the risks of AI, or the uncertain impact of AI on society, any lower than people in other countries. Nor did they differ from other countries on the importance of the principles and practices required to ensure AI is trustworthy. Rather, a key differentiator is that most people in the BICS countries and Singapore believe the benefits of AI outweigh the risks, whereas a minority of people in western countries, such as Australia, Canada, France, Netherlands, the UK and USA, hold this view.

The higher trust and more positive attitudes in the BICS countries is likely due to the greater benefits afforded by technological advances and deployment in emerging economies and the increasingly important economic role of AI technologies in these countries.

This may encourage a growth mindset that motivates acceptance and use of technology as a means to accelerate economic progress, prosperity, and quality of life. An implication is that these countries may be uniquely positioned to rapidly accelerate innovation and technological advantage through AI. It is notable, however, that on international rankings these countries rank low on governance and regulation frameworks to ensure the ethical and responsible use of AI compared to western countries.[6]

AI awareness, understanding and trust in AI has increased over time, but institutional safeguards continue to lag

We had the opportunity to examine how trust and select attitudes to AI compared with our 2020 Trust in AI survey data, which was based on representative sampling from five western countries (Australia, Canada, Germany, the UK and the USA).[7] Comparisons were made between data from these five countries in 2020 and 2022 using equivalent measures over time.

This comparison suggests that trust in AI systems has increased in these countries over time, as has awareness of AI and understanding of AI use in common applications. However, there has been no increase in the perceived adequacy of institutional safeguards, such as regulation and laws to protect people from problems, despite most people in these countries perceiving such institutional safeguards as insufficient in 2020. Similarly, there was no increase in people's confidence in government and business to develop, use or regulate AI, despite low levels of confidence in these entities. There was, however, an increase in the view that AI regulation is needed in two countries - the UK and USA.

These findings suggest the institutional safeguards governing AI are not keeping pace with expectations and technological uptake. In some jurisdictions, these findings may reflect a lack of communication and awareness of regulatory change.

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

Our analysis demonstrated that trust strongly influences AI acceptance, and hence is critical to the sustained societal adoption and support of AI.[8]

Our modelling identified four distinct pathways to trust, which represent key drivers that influence people's trust in AI systems:

1. an *institutional pathway* reflecting beliefs about the adequacy of current safeguards, regulations and laws to make AI use safe, and confidence in government and commercial organisations to develop, use and govern AI
2. a *motivational pathway* reflecting the perceived benefits of AI use
3. an *uncertainty reduction pathway* reflecting the need to address concerns about the risks associated with AI
4. a *knowledge pathway* reflecting people's understanding of AI use and efficacy in using technology

Of these drivers, the institutional pathway had the strongest influence on trust, followed by the motivational pathway.

These findings highlight the importance of developing adequate governance and regulatory mechanisms that safeguard people from the risks associated with AI use and public confidence in entities to enact these safeguards, as well as ensuring AI is designed and used in a human-centric way to benefit people and support their understanding.

Pathways to strengthen public trust and acceptance

Collectively, the survey insights provide evidence-based pathways for strengthening the trustworthy and responsible use of AI systems, and the trusted adoption of AI in society. These insights are relevant for informing responsible AI strategy, practice and policy within business, government, and NGOs at a national level, as well as informing AI guidelines, standards and policy at the international and pan-governmental level.

There are a range of resources available to support organisations to embed the principles and practices of trustworthy AI into their everyday operations and put in place mechanisms that support stakeholder trust in the use of AI.[9] While proactively investing in these trust foundations can be time and resource intensive, our research suggests it is critical for sustained acceptance and adoption of smart technologies over time, and hence a return on investment.

Introduction

AI is rapidly becoming a ubiquitous part of everyday life and continuing to transform the way we live and work.[10] All sectors of the global economy are now embracing AI, with AI applications expanding and diversifying into domains ranging from transport, crop and service optimisation, the diagnosis and treatment of diseases, and the protection of physical, financial, and cyber security, for example by fining distracted drivers, detecting credit card fraud, identifying children at risk, and enabling facial recognition.

While the benefits and promise of AI for society and business are undeniable, so too are the risks and challenges. These include the risk of codifying and reinforcing unfair biases, infringing on human rights such as privacy, spreading fake online content, deskilling and technological unemployment, and the risks stemming from mass surveillance technologies, critical AI failures and autonomous weapons. Even in cases where AI is developed to help people (e.g. to protect cybersecurity), there is the risk it can be used maliciously (e.g. for cyberattacks). These issues are causing public concern and raising questions about the trustworthiness and governance of AI systems.[11]

The public's trust in AI technologies is vital for continual acceptance. If AI systems do not prove worthy of trust, their widespread acceptance and adoption will be hindered, and the potential societal and economic benefits will not be fully realised.[12]

Despite the central importance of trust, to date little is known about how trust in AI is experienced by people in different countries across the globe, or what influences this trust.[13] In 2020, we conducted the first deep dive survey examining trust in AI systems across five western countries – Australia, Canada, Germany, the UK and the USA (Gillespie, Lockey & Curtis, 2021). The current study extends this focus on trust in AI by examining the perspectives of people representing 17 countries drawn from all global regions: the original five western countries in addition to Brazil, China, Estonia, Finland, France, India, Israel, Japan, Netherlands, Singapore, South Africa, and South Korea.

Our research aims to understand and quantify people's trust in and attitudes towards AI, benchmark these attitudes over time, and explore similarities and differences across countries. Taking a global perspective is important given AI systems are not bounded by physical borders and are rapidly being deployed and used across the globe.

Our report is structured to provide evidence-based insights on the following questions about the public's trust and acceptance of AI:

- To what extent do people trust AI systems?
- How do people perceive the benefits and risks of AI?
- Who is trusted to develop, use and govern AI?
- What expectations do people have about the development, governance and regulation of AI?
- How do people feel about the use of AI at work?
- How well do people understand AI?
- What are the key drivers of trust and acceptance of AI?
- How have trust and attitudes towards AI changed over time?

In the next section, we provide an overview of the research methodology used. In the concluding section, we draw on the survey insights to identify evidence-based pathways for strengthening the trusted and responsible use of AI systems, and discuss the implications for industry, government, universities, and non-government organisations (NGOs).

How we conducted the research

We collected data in each country using representative research panels.[14] This approach is common in survey research to recruit people who are representative of a national population. Panel members were invited to complete the survey online, with data collected between September and October 2022.

The total sample included 17,193 respondents from 17 countries. We chose the countries based on three criteria: 1) representation across all nine global regions[15]; 2) leadership in AI activity and readiness[16], and 3) diversity on the Responsible AI Index.[17] The sample size across countries ranged from 1,001 to 1,021 respondents.

Surveys were conducted in the native language(s) of each country, with the option to complete in English, if preferred. To ensure question equivalence across countries, surveys were professionally translated and back-translated from English to each respective language, using separate translators. See Appendix 1 for further method details.

Who completed the survey?

Country samples were nationally representative of the adult population on gender, age and regional distribution matched against official national statistics within each country. Across the total sample, the gender balance was 50% women, 49% men and 1% non-binary and other gender identities. The mean age was 44 years and ranged from 18 to 91 years.

Ninety percent of respondents were either currently employed (67%) or had prior work experience (23%). These respondents represented the full diversity of industries and occupational groups listed by the OECD.[18] Almost half the sample (49%) had a university education. Further details of the sample representativeness, including the demographic profile for each country sample, are shown in Appendix 2.

17
countries

17,193
respondents

- Australia
- Brazil
- Canada
- China
- Estonia
- Finland
- France
- Germany
- India
- Israel
- Japan
- Netherlands
- Singapore
- South Africa
- South Korea
- United Kingdom
- United States

Gender

50% Women

49% Men

1% Non-binary & other genders

Age Group

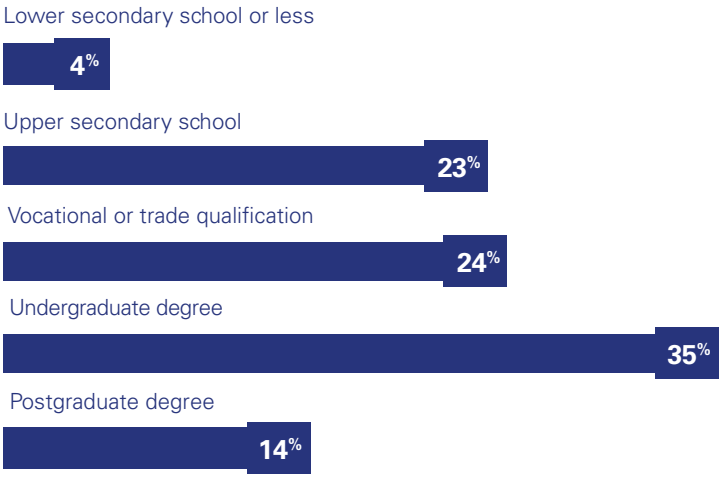
15% Generation Z (18 – 25)

33% Millennial (26 – 41)

28% Generation X (42 – 57)

24% Baby Boomer + Silent Generation (58 – 91)

Education



Occupation

31% Professional & Skilled (including army)

24% Manager

19% Administrative

13% Service & Sales

13% Manual

How we asked about AI

After asking about respondents' understanding of AI, the following definition of AI was provided.

> Artificial Intelligence (AI) refers to computer systems that can perform tasks or make predictions, recommendations or decisions that usually require human intelligence. AI systems can perform these tasks and make these decisions based on objectives set by humans but without explicit human instructions (OECD, 2019).

Given perceptions of AI systems can be influenced by the purpose and use case,[19] survey questions asking about trust, attitudes and governance of AI systems referred to one of five AI use cases (randomly allocated): Healthcare AI (used to inform decisions about how to diagnose and treat patients), Security AI (used to inform decisions about public safety and security), Human Resource AI (used to inform decisions about hiring and promotion), Recommender AI (used to tailor services to consumers), or AI in general (i.e. AI systems in general).

These use cases were chosen as they represent domains where AI is being rapidly deployed and is likely to be used by, or impact, many people.

Before answering questions, respondents were provided with a description of the AI use case, including what it is used for, what it does and how it works. These descriptions are shown below, and were developed based on current in use systems and input from domain experts working in healthcare, security, human resources, and recommender systems, respectively.

Healthcare AI

An AI system used to improve the diagnosis of disease (e.g. cancer), inform the best treatment options, and predict health outcomes based on patient data. It works by comparing a patient's health data (e.g. symptoms, test and scan results, medical history, family history, age, weight and gender etc.) to large datasets based on many patients. Doctors use Healthcare AI to inform decisions about patient diagnosis and treatment.

Security AI

An AI system used to help identify suspicious or criminal behaviour, security threats, and people of interest to police. It works by processed and analysing a range of information such as face and fingerprint scans collected in places like country borders, and photos and live footage of people and vehicles in public places collected through security cameras. Police and security agencies use Security AI to inform decisions about public safety and security.

Human Resources AI

An AI system used to help select the most suitable applicants for a job, identify workers who are most likely to perform well in a job, and predict who is most likely to quit. It works by collecting and comparing worker characteristics, employee data, and performance over time, and analysing which qualities are related to better job performance and job retention. Managers use Human Resources AI to inform decisions about hiring and promotion.

Recommender AI

An AI system used to personalise services such as news, social media content and product recommendations by providing content and products that are most relevant to the user. It works by predicting a person's choices and preferences based on their characteristics (e.g. age, gender, location), past behaviour, interests or preferences, and the behaviour of similar users. Companies use Recommender AI to tailor services to consumers.

How we analysed the data

We conducted statistical analyses to examine differences between countries, AI use cases, and demographic groups. Where significant and meaningful differences are evident between countries, we report country-level data. Further details of the statistical procedures are discussed in Appendix 1. We also report meaningful differences between groups and AI use cases.

TOPIC ONE

To what extent do people trust AI systems?

To answer this question, we asked respondents how much they trust and accept a range of AI systems, and the extent to which they perceive them to be trustworthy. We also asked people how they feel about AI.

We define trust in AI as a *willingness to accept vulnerability* to an AI system (e.g. by relying on system recommendations or output, or sharing data) based upon *positive expectations* of how the system will operate (e.g. accuracy, helpfulness, data privacy and security). [20]

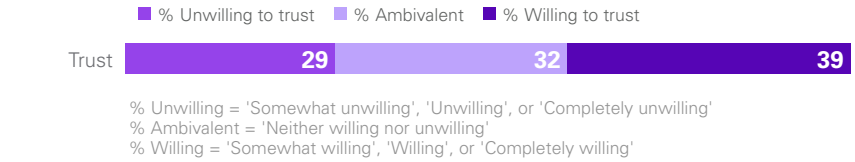
Most people are ambivalent or unwilling to trust AI systems

Three out of five people (61%) across countries are wary about trusting in AI systems, reporting either ambivalence or an unwillingness to trust (see Figure 1). In contrast, 39% report that they are willing to trust AI systems.

Mirroring these findings, most people (67%) report low to moderate acceptance of AI. Only a third of people across countries report high acceptance. There is a strong association between trust in AI and acceptance of AI (correlation r=0.71, p<0.001).

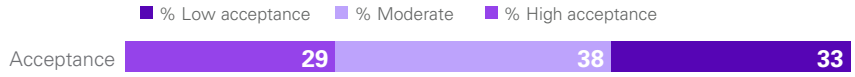
Figure 1. Willingness to trust and accept AI systems

'How willing are you to trust AI [specific application]?' [8 items]



% Unwilling = 'Somewhat unwilling', 'Unwilling', or 'Completely unwilling'
% Ambivalent = 'Neither willing nor unwilling'
% Willing = 'Somewhat willing', 'Willing', or 'Completely willing'

'To what extent do you accept the use of AI [specific application]?' [3 items]



% Low acceptance = 'Not at all' or 'Slightly'
% Moderate acceptance = 'Moderately'
% High acceptance = 'Highly' or 'Completely'

There are stark differences across countries: AI is most trusted and accepted in the emerging 'BICS' economies of Brazil, India, China, and South Africa

Our survey revealed stark differences in trust and acceptance of AI systems across countries. Figure 2 shows trust in AI is highest in India, China, Brazil, and South Africa, respectively.

These countries are each part of the BRICS alliance of major emerging economies. We use the acronym 'BICS' in this report to denote the four countries of Brazil, India, China, and South Africa included in our survey that showed a distinctively different pattern of findings to the other countries.

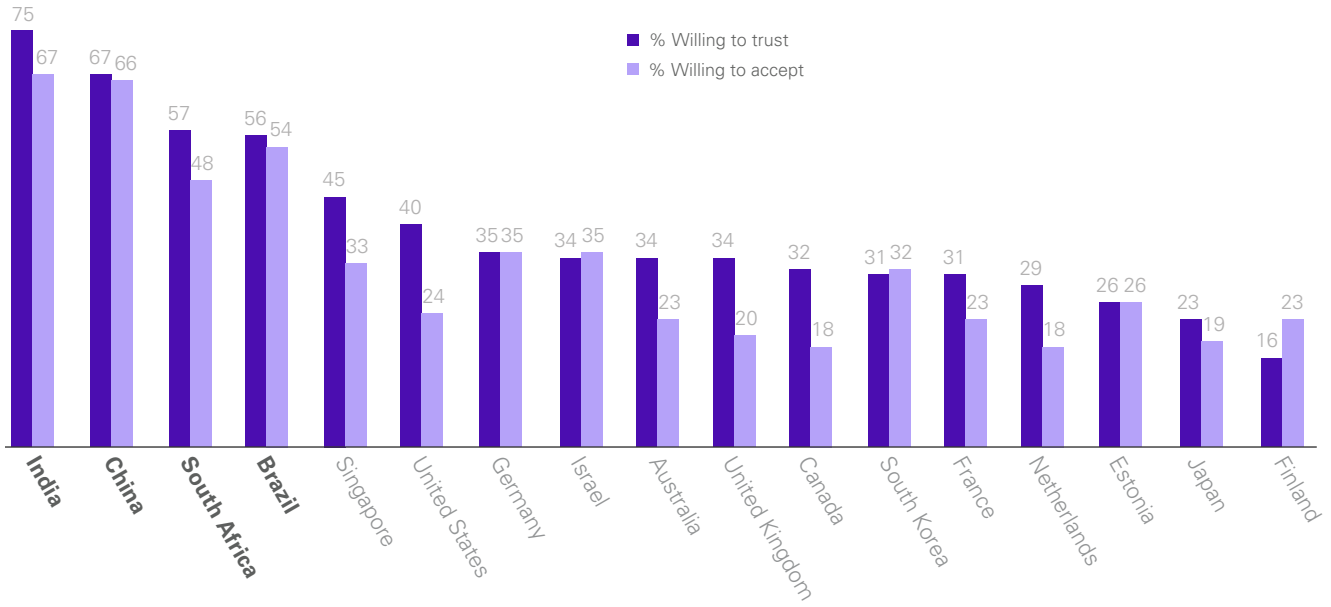
In the BICS countries, most people (56–75%) trust in AI systems, with people in India reporting the greatest willingness to trust, followed by China. In contrast, a minority of people in other countries report trusting AI, with the Finnish reporting the lowest trust (only 16%).

We see a similar pattern for acceptance of AI as we do for trust. The BICS countries are notably higher in their acceptance of AI, with 48–67% of people in these countries reporting high acceptance.

Again, India and China lead the way, with 66–67% reporting high acceptance of AI, compared to only 18% in the Netherlands and Canada, respectively. There are low levels of AI acceptance across all Western countries, with Germany reporting the most acceptance (35% high acceptance).

The higher trust and acceptance of AI in the BICS countries is likely due to the accelerated uptake of AI in these countries, and the increasingly important economic role of emerging technologies.[21] As discussed in the forthcoming sections of this report, people in the BICS countries are the most positive about AI, perceive the most benefits from it, and report the highest levels of AI adoption and use at work.

Figure 2: Willingness to trust and accept AI systems across countries



% Willing to trust = 'Somewhat willing', 'Mostly willing', or 'Completely willing'
% Willing to accept = 'Highly accept' or 'Completely accept'

Trust and acceptance depend on the application: AI use in human resources is the least trusted and accepted

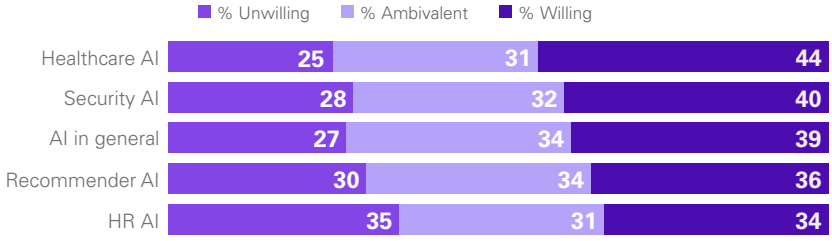
People's trust and acceptance of AI depends on the specific application or use case.

There is a tendency for people to trust the use of AI in human resources (HR) the least (34% willing, M=3.9; see Figure 3), and the use of AI in healthcare diagnosis and treatment the most (44% willing, M=4.3).

This difference likely reflects the important direct benefit that increased precision of medical diagnosis and treatments affords people, combined with the high levels of trust in doctors in most countries.

Figure 3. Trust in AI systems across applications

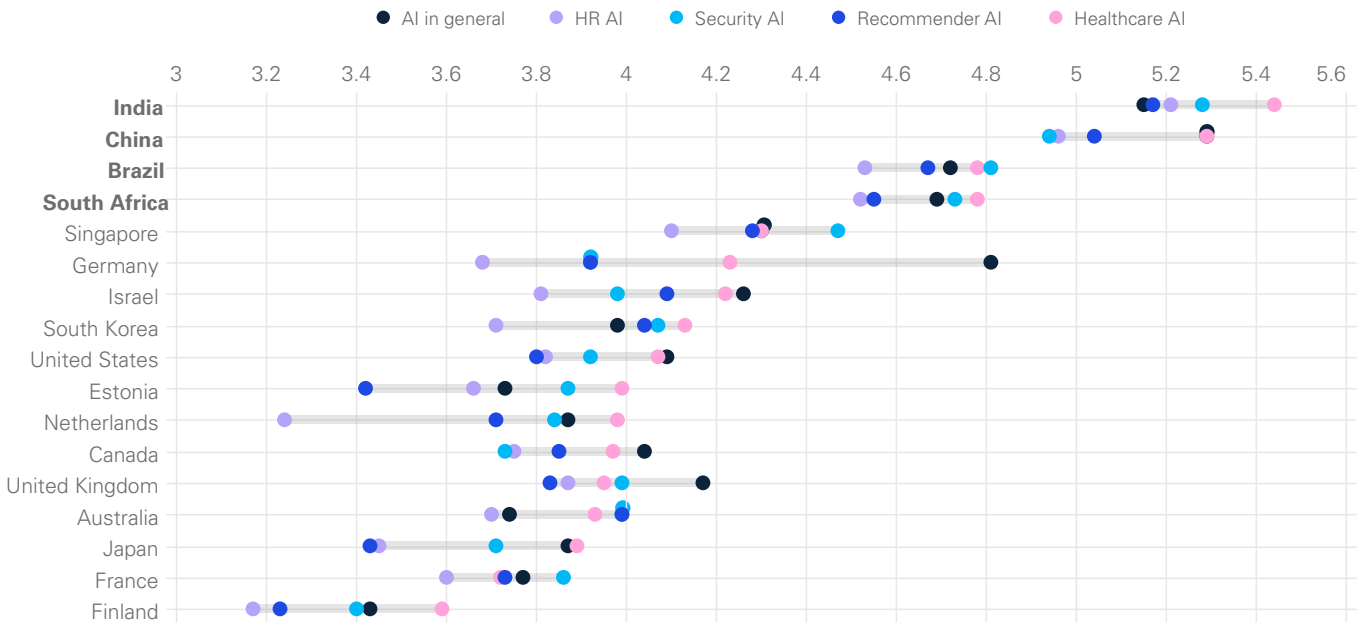
'How willing are you to trust AI [specific application]?' [8 items]



% Unwilling = 'Somewhat', 'Mostly' or 'Completely unwilling' to trust
% Ambivalent = 'Neither willing nor unwilling' to trust
% Willing = 'Somewhat', 'Mostly' or 'Completely willing' to trust

This difference between applications is meaningful in some countries, but not all. Specifically, people in eight countries – Germany, the Netherlands, Finland, Estonia, Israel, South Korea, Japan, and China – report lower trust and acceptance of Human Resources AI, than either Healthcare AI or AI in general (see Figure 4).

Figure 4. Trust in AI systems across countries



* Mean trust in AI application on 7 point scale
[Countries sorted in order of 'Healthercare AI']

People are more willing to rely on than share information with AI systems, particularly with security and recommender systems

We drilled down to examine two key ways people demonstrate trust in AI systems: reliance and information sharing.

Reliance

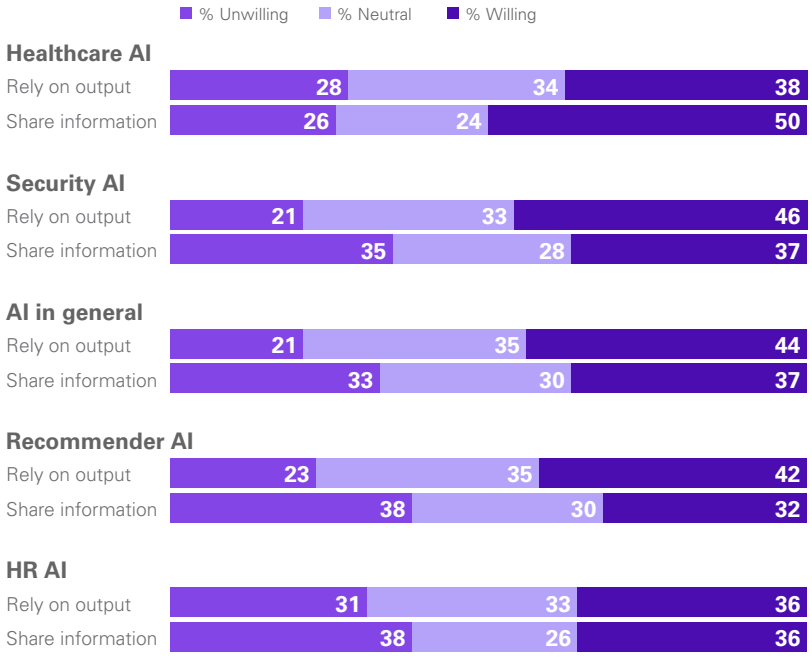
Assesses people's willingness to rely on an AI system's output, such as a recommendation or decision (i.e. to trust that it is accurate). If people are not willing to rely on AI system output, the system will not be used.

Information sharing

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

Figure 5. Willingness to rely on and share information with AI systems

'How willing are you to: rely on information provided by [specific AI application] / share information with [specific AI application]?' [8 items]



% Unwilling = 'Somewhat unwilling', 'Unwilling', or 'Completely unwilling'
% Neutral = 'Neither willing nor unwilling'
% Willing = 'Somewhat willing', 'Willing', or 'Completely willing'

People in general are more willing to rely on, rather than share information with AI systems (see Figure 5). People are particularly more willing to rely on Security and Recommender AI applications than share information with them.

However, this pattern is reversed for Healthcare AI, where respondents are usually more willing to share information than rely on the outcomes of the application. This most likely reflects that sharing information with healthcare providers and systems is normal and routine to facilitate effective care.

AI systems are perceived as more trustworthy in BICS countries

To understand how people view the trustworthiness of AI systems, we asked people about three key components of trustworthiness: ability, humanity and integrity.

Ability

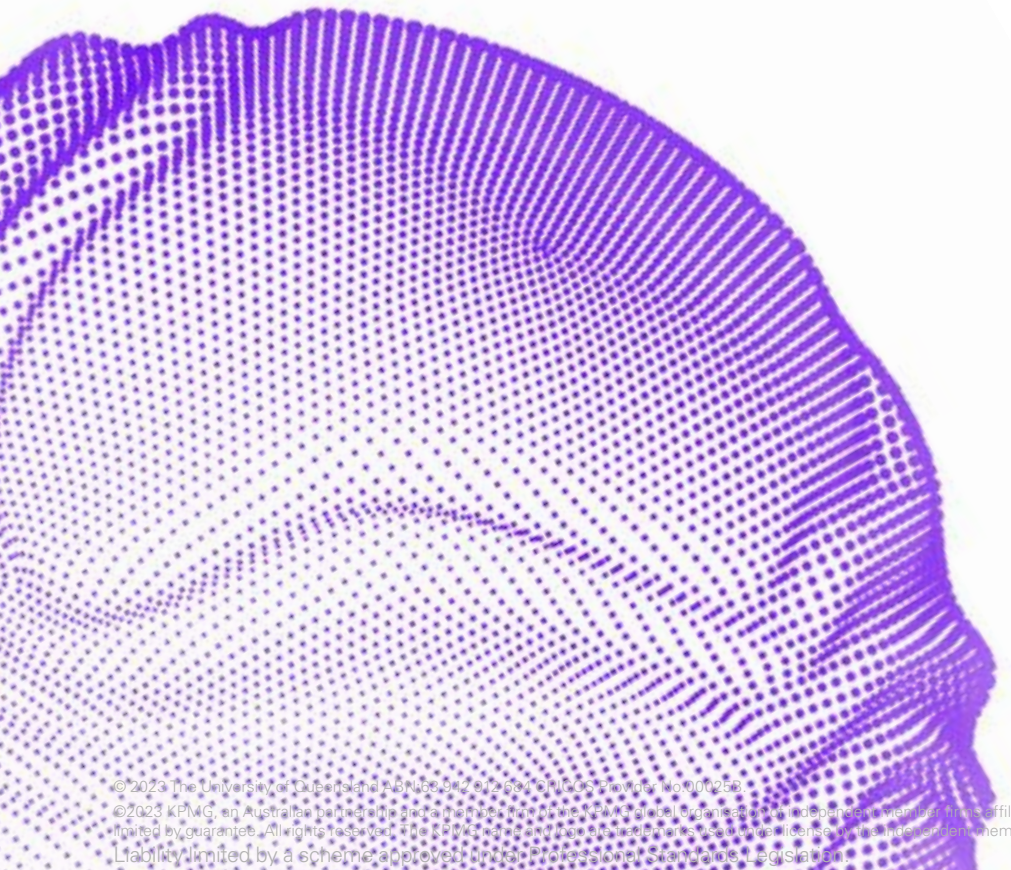
AI systems are fit-for-purpose and perform reliably to produce accurate output as intended.

Humanity

AI systems are designed to deliver beneficial outcomes for people and society, and have a positive impact.

Integrity

AI systems are safe and secure to use and adhere to commonly accepted ethical principles (e.g. fairness, do no harm), human rights (e.g. privacy) and applicable laws.

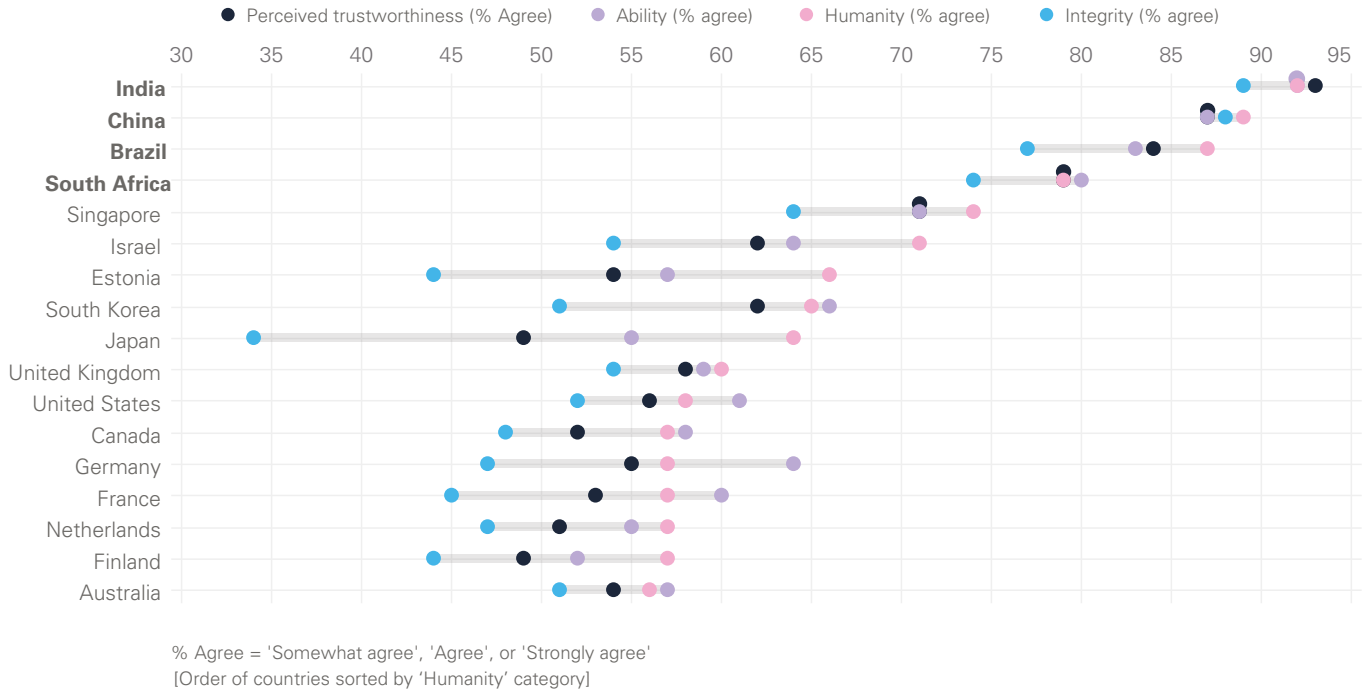


As shown in Figure 6, we see a similar pattern across countries in beliefs about the trustworthiness of AI systems, as for AI trust and acceptance. People in the BICS countries hold much more positive beliefs about the trustworthiness of AI systems, compared to all other countries, with 79–93% viewing these systems as trustworthy. Indians again have the most positive views, with 93% agreeing that AI systems are trustworthy, followed by the Chinese (87%).

In contrast, people in western countries, as well as Japan, have the least favourable beliefs about the trustworthiness of AI, ranging from 49–58% viewing AI systems as trustworthy.

Figure 6. Perceptions of the trustworthiness of AI systems

'I believe [specific AI application] would: produce output that is accurate (ability) / have a positive impact on most people (humanity) / be safe and secure to use (integrity)' [14 items]



% Agree = 'Somewhat agree', 'Agree', or 'Strongly agree'
[Order of countries sorted by 'Humanity' category]

On average, 63% of people across all countries and applications perceive AI systems are trustworthy. Perceptions of trustworthiness are typically higher than trusting intentions because trust involves risk and vulnerability (e.g. by relying on AI output or sharing information with an AI system), whereas perceiving a system as trustworthy does not.

There is a strong association between perceived trustworthiness and trust in AI systems (r=0.77, p<0.001).

In line with the findings for trust and acceptance, we found differences in trustworthiness across applications. In most countries, Human Resources AI is seen as less trustworthy than other AI applications.

More people believe AI systems are capable and beneficial than believe they are safe and designed to uphold ethical principles and rights

As shown in Figure 6, people have most faith in the ability of AI systems to produce accurate and reliable output and provide helpful, beneficial services for people. In contrast, people are more sceptical about the extent to which AI systems are safe and secure to use and adhere to commonly accepted ethical principles (e.g. fairness, do no harm) and privacy rights (integrity).

This was a consistent pattern across countries, with the exception of China and India. For example, in Japan, most people (55–64%) view AI systems as technically competent (M=4.6/7) and beneficial (M=4.9/7), however, only 34% agree that AI systems uphold ethical principles and rights (integrity M=4.1/7).

This difference between AI integrity as compared to ability and humanity is evident for AI in general and two specific applications – AI use for security and AI recommender systems. There are no meaningful differences between perceptions of AI ability, humanity and integrity for Healthcare AI or Human Resources AI.

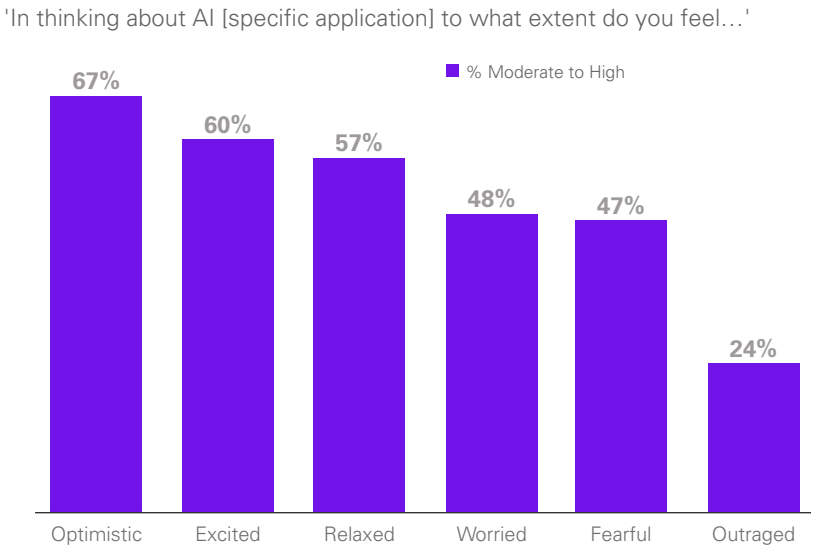
Most people are optimistic and excited about AI: however, many also feel worried and fearful

We asked people the extent to which they feel a range of emotions about the AI applications. A majority of people report positive emotions such as feeling optimistic, excited, or relaxed about these AI systems. However, just under half the people also report feeling worried or fearful about the AI applications, and just under a quarter feel outraged (see Figure 7).

People who have positive emotions towards an AI system are more likely to also trust in AI, as demonstrated by the strong, positive correlation (r=0.68, $p<0.001$). In contrast, when people feel negative emotions towards AI, this is associated with lower trust (r=-0.28, $p<0.001$).

Further analysis[22] revealed people commonly experience ambivalent feelings towards AI: 41% experience both high positive *and* negative emotions, for example feeling excited but also worried about AI. In contrast, 35% experience high positive emotions coupled with low negative emotions, and 16% have low positive emotions coupled with high negative emotions. Only 8% report feeling low positive and negative emotions towards AI.

Figure 7. Emotions associated with AI



* 5 point scale
% Moderate to High = 'Moderately', 'Very', or 'Extremely'

People in the BICS countries feel most optimistic, excited and relaxed about AI

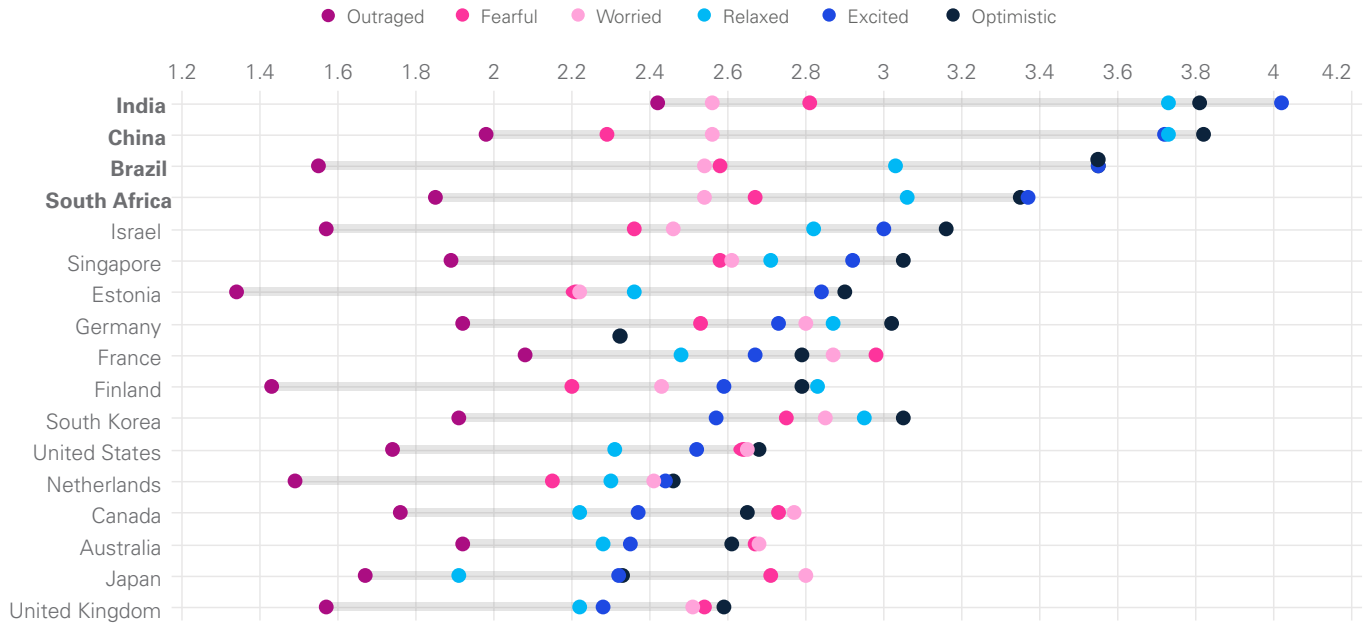
Figure 8 shows how people feel about AI in each country, ordered by the countries with people who felt most positive about AI. People in the BICS countries are the most optimistic, excited, and relaxed about AI, and people in Japan the least. Positive emotions were significantly stronger than negative emotions in the BICS countries as well as in Estonia, Finland, and Israel.

Fear or worry about AI was the dominant emotion experienced by people in Australia, Canada, France, and Japan, with people in France amongst the most fearful and worried. People generally did not feel much outrage towards AI.

While people in India are most likely to feel positive emotions about AI, they also have one of the highest levels of fear and are more likely to report outrage than other countries. This reinforces that in many countries, fear and worry about AI often coincides with optimism or excitement.

Figure 8. Emotions towards AI across countries

'In thinking about AI [specific application] to what extent do you feel… '



* 5 point scale
1 = Not at all, 2 = Slightly, 3 = Moderately, 4 = Very, 5 = Extremely
[Order of countries sorted by 'Excited' category]

Younger generations, the university educated, and managers are more trusting and accepting of AI systems, and more likely to feel positive emotions

As shown in Figure 9 and through statistical analyses, younger people, notably Generation Z and Millennials, are more trusting and accepting of AI, than older generations and view AI systems as more trustworthy than older generations.

These generational effects held across most countries and are particularly pronounced in Australia and the USA. For example, in Australia, 25% of older generations trust AI compared to 42% of Gen X and Millennials, and 13% of older generations accept AI compared to 34% of Gen Z and Millennials. In contrast, in South Korea and China, we see a reversal of this pattern, with older generations more trusting of AI than younger generations. For example, in South Korea, 23% of Gen Z and Millennials are willing to trust AI, compared to 44% of Baby Boomers and older generations.

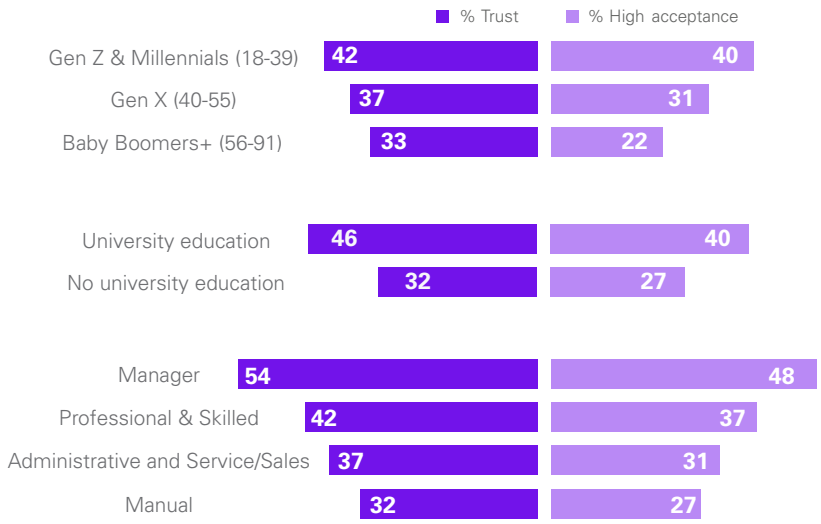
People with a university education are also more trusting and accepting of AI than those without a university degree and hold more positive views of the trustworthiness of AI. This difference was also particularly evident in Australia, with 42% of university educated Australians willing to trust AI, compared to 27% of Australians without a university education.

Managers are also more trusting and accepting of AI, and perceive it as more trustworthy than people in other occupations.

In addition, younger generations, those with a university education, and managers are more likely to feel positive emotions about AI. There are no generational, educational, or occupational differences in the experience of negative emotions about AI.

It is noteworthy that there are no meaningful differences across men, women and other genders in trust, acceptance, or emotions towards AI, however in a few countries (USA, Singapore, Israel, and South Korea, respectively), men were more trusting or accepting of AI, and reported more positive emotions, than other genders.

Figure 9: Trust and acceptance of AI systems by generation and education



TOPIC TWO

How do people perceive the benefits and risks of AI?

To answer this question, we asked people about a range of potential benefits and risks associated with AI, the likelihood of risks occurring, as well as whether the benefits outweigh the risks.

People expect AI will deliver a range of benefits but perceive more process benefits than benefits to people

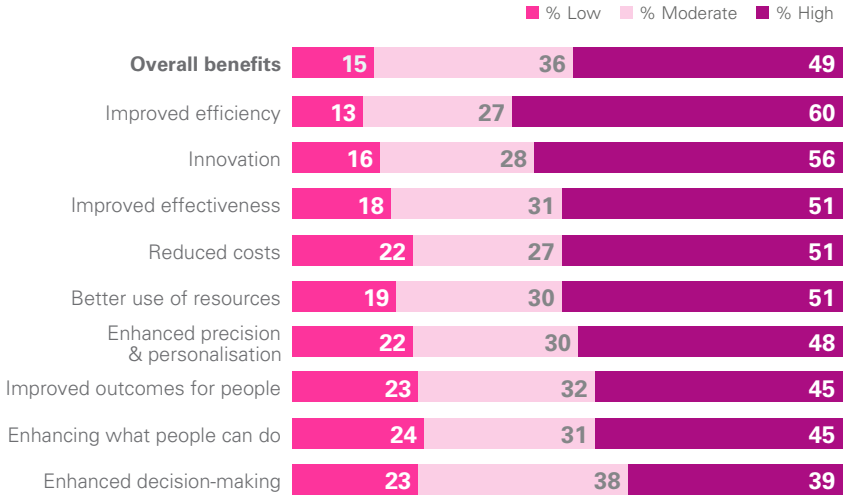
Most people (85%) believe the use of AI will result in a range of benefits, as shown in Figure 10. People who perceive more benefits from AI are also much more likely to trust in AI systems (r=0.62, p<0.001).

People have particularly high expectations that AI will improve efficiency, innovation, effectiveness, resource utilisation and reduce costs. People perceive the 'process' benefits of AI, such as improved efficiency and innovation, as greater than the 'people' benefits of AI, such as improving outcomes for people, and enhancing decision-making and what people can do.

In many countries, the benefits of using AI in Human Resources, particularly the benefits for people and effectiveness, were lower than the benefits of using other applications of AI (e.g. Security, Healthcare, Recommender systems and AI in general).

Figure 10: The perceived benefits of AI use

'To what extent do you expect these potential benefits from the use of AI [specific application]?'



Low = 'Not at all' or 'To a small extent'
Moderate = 'To a moderate extent'
High = 'To a large extent' or 'To a very large extent'

People in the BICS countries perceive the greatest benefits of AI

There are significant differences between countries in perceptions of AI benefits. As shown in Figure 11, people in the BICS countries have the most positive view of the benefits of AI (Ms=3.8–4.0/5). In contrast, people in Australia, Canada, the UK, USA, the Netherlands, Finland, and Japan, were less convinced by the benefits of AI (Ms=3.0–3.1/4).

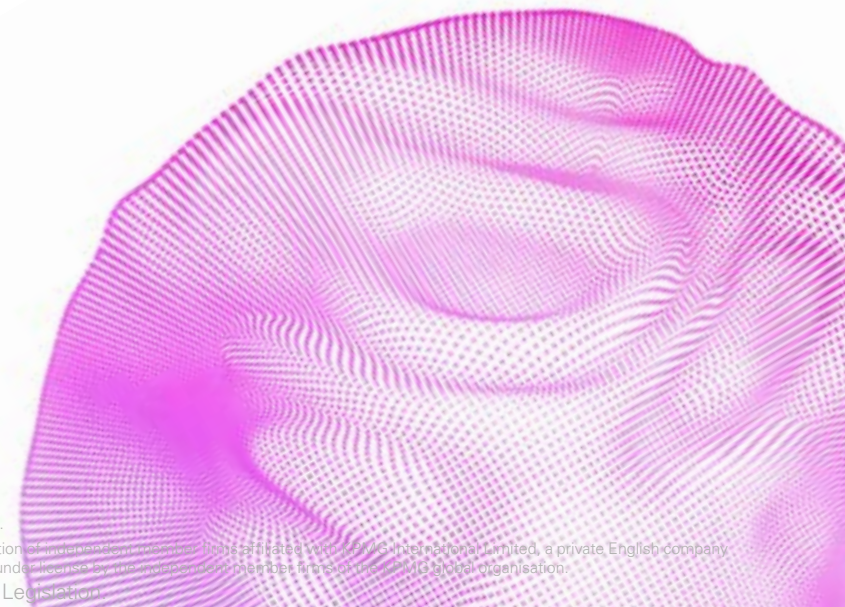
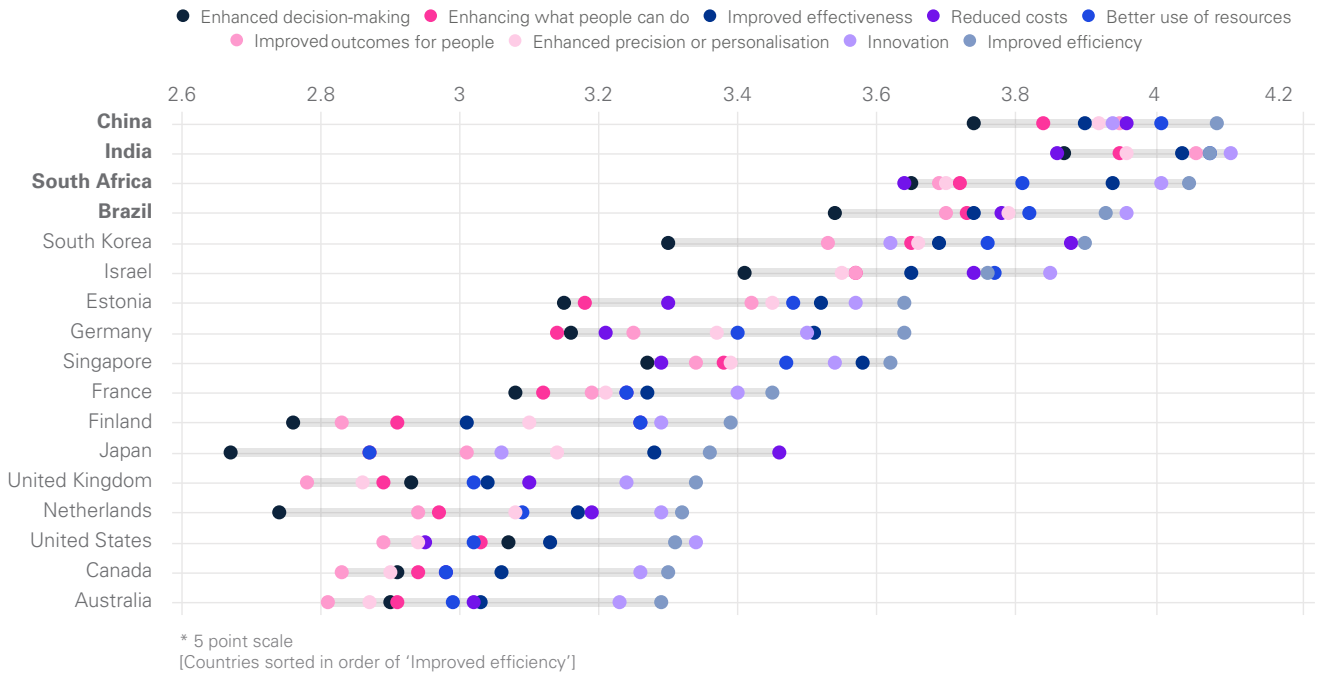


Figure 11: The perceived benefits of AI across countries

'To what extent do you expect these potential benefits from the use of AI [specific application]?'



* 5 point scale
[Countries sorted in order of 'Improved efficiency']

People are concerned about a range of potential risks from AI use, particularly cybersecurity risks

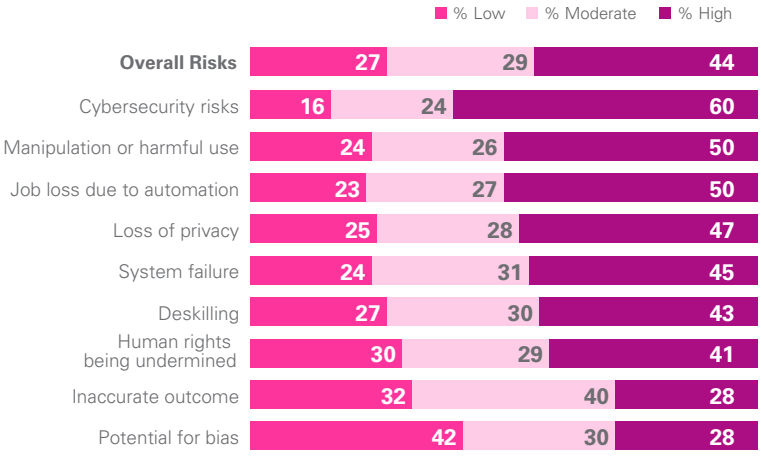
While people expect significant benefits from AI, the large majority (73%) also perceive significant potential risks from AI. People who perceive more risks of AI use, are somewhat less trusting of AI systems (r=-0.25, p<0.001).

Cybersecurity risk (e.g. from hacking or malware) is the dominant concern raised by 84% of people. Other risks of moderate to very large concern raised by more than two-thirds of people (68–77%) include manipulative or harmful use of AI, job loss and deskilling, loss of privacy, system failure, undermining of human rights and inaccurate outcomes.

In comparison, people are less concerned about the risk of bias from AI use. However, bias is still a concern for the majority of people (58%). This may reflect that the general public perceive AI systems as less biased than humans, or alternatively, are less aware of the potential risk of bias from AI systems.

Figure 12: The perceived risks of AI use

'How concerned are you about these potential risks of AI use [specific application]?'



% Low= 'Not at all' or 'To a small extent' '
% Moderate = 'To a moderate extent'
% High = 'To a large extent' or 'To a very large extent'

To complement these quantitative findings, we asked people: 'What concerns you most about the use of AI [specific application]? Thematic analysis of this open-ended data[23] reinforced that people across all countries are concerned about each of the risks shown in Figure 12, including concerns about:

- privacy breaches, cybersecurity attacks and hacking
- manipulation and harmful use, including misuse by service providers and governments (e.g. to monitor or control)
- job loss, technological unemployment and deskilling
- inaccurate outcomes and recommendations, and poor or biased decisions
- system failure or malfunction causing harm
- the loss of human control and agency, and loss of human judgement in decision-making, resulting in unintended consequences (including AI 'taking over').

The qualitative data further highlighted concerns around:

- a lack of transparency of when and how AI is being used, and how AI generates decisions and outcomes
- a lack of regulations, policies and governance to make AI use safe and ethical.

It is also important to note that some people report having no concerns.

People view the risks of AI in a comparable way across countries

In contrast to the distinct differences across countries in how people view the benefits of AI, there are many commonalities in how people from different countries perceive the risks of AI use.

In almost all countries, people are most concerned about cybersecurity risks. The exceptions are people in India and South Africa, who are most concerned with job loss due to automation, followed by cybersecurity risks. This concern about job loss may reflect the recent increase in AI-related activity in these two countries.[24]

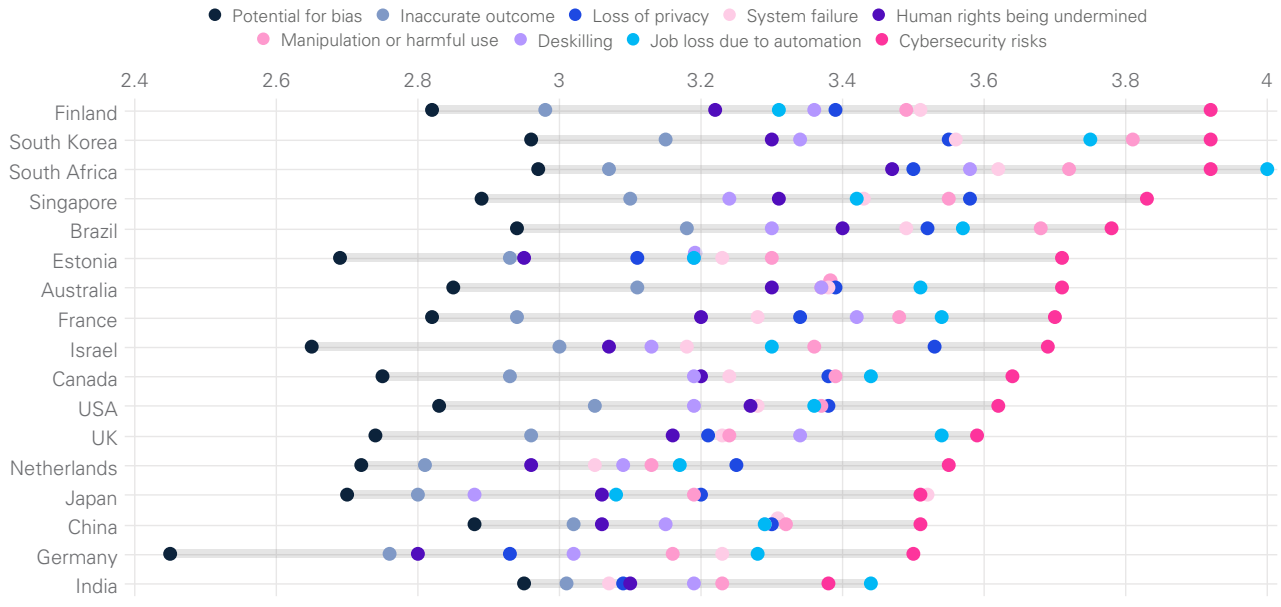
While AI acceleration clearly has the potential to provide economic benefits to these countries, it may also result in job losses. In Japan, the top concerns are AI system failure (e.g. where the AI system malfunctions or goes offline) and cybersecurity, which may reflect the heavy dependence on smart technology in Japan.

We also see that across all countries, people are least concerned about the potential risk of bias from AI use, followed by inaccurate outcomes from AI use.

People in South Africa, South Korea and Brazil perceive the risks of AI higher than people in most other countries. In contrast, people in Germany perceive the potential risks of AI lower than people in most other countries.

Figure 13: The perceived risks of AI across countries

'How concerned are you about these potential risks of AI use?'



* 5 point scale
[Countries sorted by 'Cybersecurity risks']

People are divided about the likelihood that AI risks will impact people

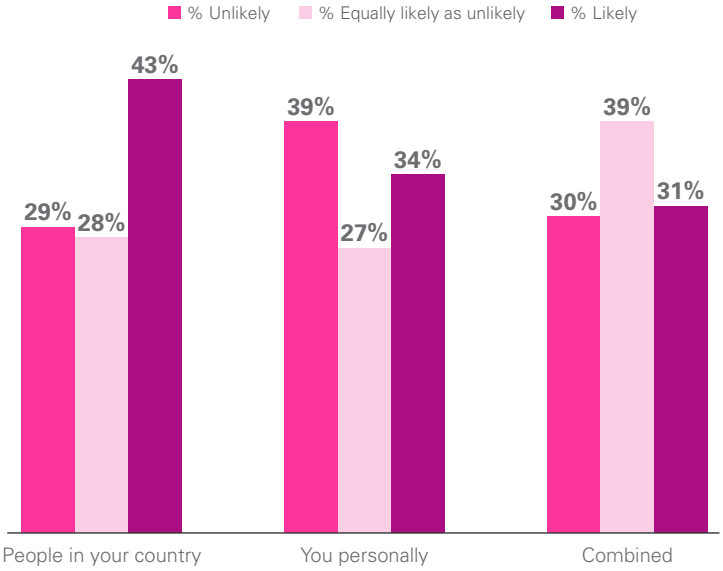
We asked respondents how likely it is that one or more of these risks would impact people in their country, as well as them personally. As shown in Figure 14 (combined score), people were split in their views, with 31% believing these risks are likely to impact people, 30% believing they are unlikely to impact, and 39% believing that it is as likely as unlikely that one or more risks will impact people.

People in western countries and Israel perceive the risks as more likely to impact other people in their country, than them personally.

People in South Korea, India and South Africa are the most likely to believe the risks associated with AI will impact people. In contrast, people in the EU countries of Finland, Estonia, France, and Germany are the least likely to believe these risks will impact people.

Figure 14: The likelihood of risks impacting people

'Within the next 10 years, how likely is it that one or more of these risks will impact.....?'



% Unlikely = 'Somewhat unlikely', 'Unlikely' or 'Very unlikely'
% Equally likely as unlikely = 'Equally likely as unlikely'
% Likely = 'Somewhat likely', 'Likely', or 'Very likely'

One in two people believe the benefits of AI outweigh the risks

We asked people whether the benefits of AI outweigh the risks, both in relation to people in their country, and to themselves personally. In both cases half agree that the benefits of AI outweigh the risks, and under a quarter (21–24%) disagreed. The remainder (26–29%) are neutral.

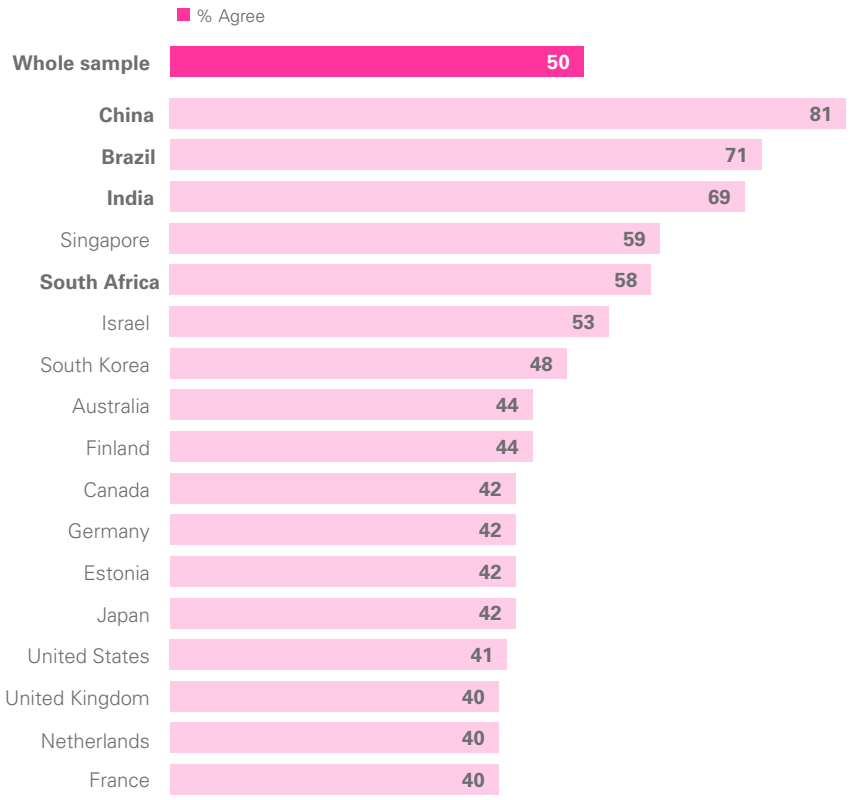
In several countries, people were less likely to believe the benefits of AI use in the Human Resources application outweigh the risks.

People are more likely to believe the benefits of AI outweigh the risks in the BICS countries and Singapore: people in western countries are more circumspect

There are large country differences in how people perceive the AI benefit-risk trade-off. As shown in Figure 15, most people in the BICS countries, Singapore and Israel (53–81%) agree the benefits of AI outweigh the risks. In contrast, people in the western countries, Japan and South Korea are less convinced and more ambivalent, with only 40–48% agreeing the benefits outweigh the risks.

Figure 15: Perceptions across countries that AI benefits outweigh risks

'Thinking about people in your country generally, to what extent do you agree the benefits of AI [specific application] outweigh the risks?' [2 items]



% Agree = 'Strongly agree', 'Agree', and 'Somewhat agree'

The university educated, younger generations, and managers perceive more benefits of AI, but there are no demographic differences in perceptions of risk

Younger generations, namely Gen Z and Millennials, view the benefits of AI more positively than people of the Baby Boomer generation or older (55% vs 37% high). People with a university education also view the benefits of AI more positively than people without a degree (56% vs 41% high) and are more likely to view the benefits of AI as outweighing the risks (51% vs 38% agree) and believe that one or more of the risks associated with AI will impact them personally (38% vs 29%).

Managers are also more likely to perceive benefits associated with AI than those with other occupations (62% vs 43–51% high), and more likely to believe the benefits of AI outweigh the risks (58% vs 36–47% agree).

There are no differences between men, women, and other genders in the perceived benefits of AI, and no differences in the perceived risks across generation, education or occupational groupings.

TOPIC THREE

Who is trusted to develop, use and govern AI?

Given the risks and benefits associated with AI, we asked people who they trusted to develop and govern AI. Specifically, we asked how much confidence people have in a variety of entities to develop and use AI, as well as regulate and govern it. We first explore the insights for the total sample, and then examine country differences.

People are most confident in universities and defence organisations to develop, use and govern AI in the best interests of the public

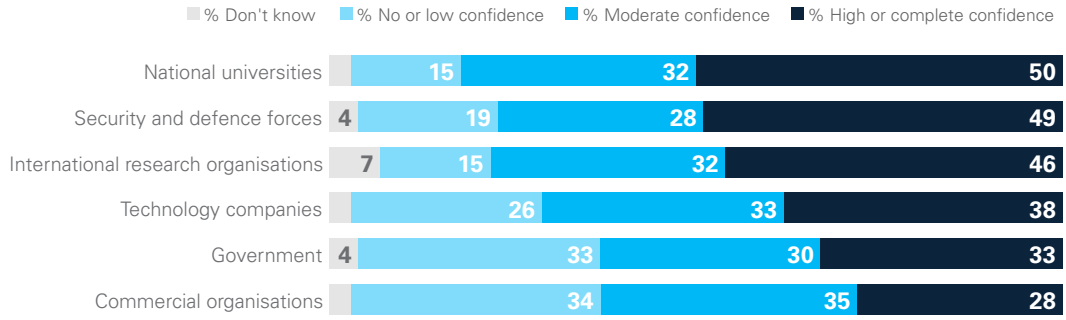
As shown in Figure 16, people have the most confidence in their national universities and research institutions, national defence forces, and international research organisations to develop and use AI in the best interests of the public, with between 77–82% reporting moderate to complete confidence (Ms=3.4–3.5/5).

Seventy-one percent report feeling confident in technology companies to develop and use AI (M=3.2/5).

People have the least confidence in government and commercial organisations (63% each, Ms=2.9–3.0), with a third of people reporting no or low confidence in government and commercial organisations to develop and use AI. A solution may be for these organisations to collaborate in AI development with more trusted entities, such as universities and research institutions.

Figure 16: Confidence in entities to develop and use AI

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



*Research institutions, defence forces, and government were country specific

There is a similar pattern regarding confidence in entities to regulate and govern AI in the best interest of the public (see Figure 17).[25] People are more confident in national universities, international research organisations, as well as security and defence organisations (76–79% confidence, Ms 3.4/5 each) to regulate and govern AI, than other entities.

People reported the least confidence in governments, technology, and commercial organisations (60–66%, Ms=2.9–3.0). About a third of people report no or low confidence in these entities to develop and regulate AI (see Figure 17).

When people are confident in entities to develop and govern AI, they are more likely to trust in AI systems (correlations ranging from 0.42 [defence forces] to 0.54 [technology companies], $p<0.001$).

Figure 17: Confidence in entities to regulate and govern AI

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

*Research institutions, defence forces, existing agencies that govern specific sectors, and government were country specific

Countries vary in their confidence in entities to develop, use and govern AI

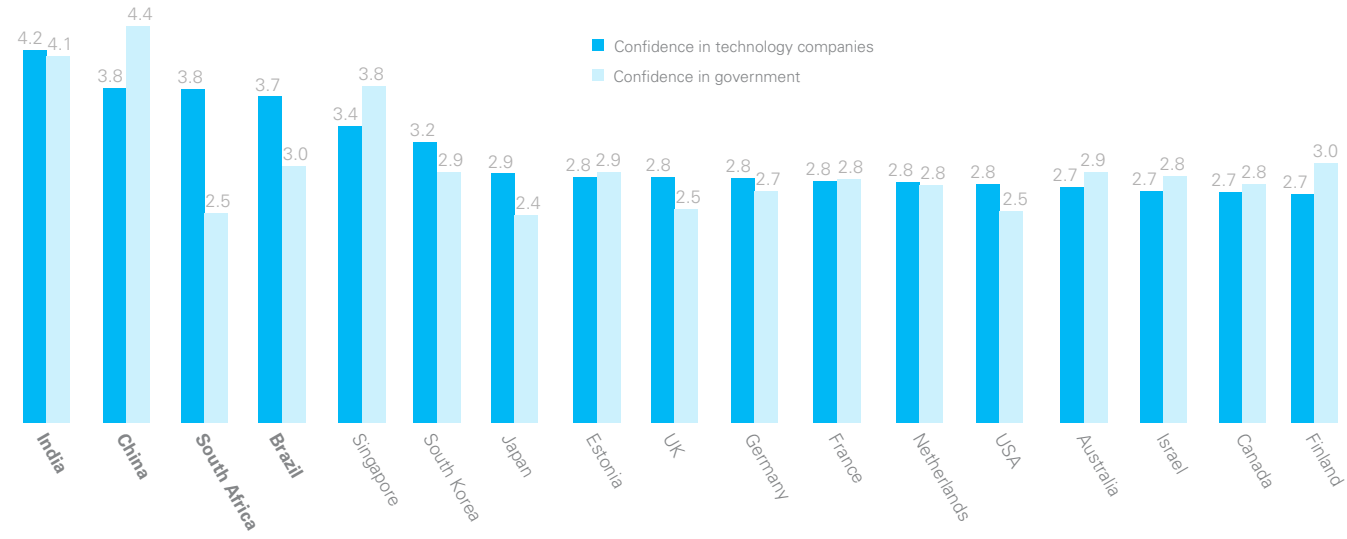
There is significant variation across countries on people's confidence in entities to develop, use and govern AI, particularly confidence in government and technology firms (see Figure 18).

A lack of confidence in government to develop, use and govern AI in the best interests of the public was reported by about half the people in South Africa (52%), the USA (49%), Japan (47%), and the UK (45%).

In contrast, many people in China (86%), India (70%), and Singapore (60%) have high or complete confidence in their governments to develop, use and govern AI.

While confidence in technology companies to develop, use and govern AI is generally low in western countries and Israel, particularly Finland, Canada, and Australia (30–40% no or low confidence, M=2.7–2.8), it is comparatively high in the BICS countries (51–73% high or complete confidence, M=3.7–4.2).

Figure 18: Confidence in technology and government entities to develop, use and govern AI



* Mean of 5 point scale amalgamating confidence to develop and use AI with confidence to regulate and govern AI

General trust partly explains confidence in entities to use and govern AI

These country differences in confidence of entities to use and govern AI can be partly explained by generalised trust towards these entities. General trust in an entity sets a foundation that influences domain specific trust in the entity to perform particular actions.

There are very high correlations between general trust in government (to do the right thing and act competently) and confidence in government to develop/use (r=0.80, p<0.001), and regulate/govern AI (r=0.82, p<0.001). The countries where people are most confident in their government to develop, use and govern AI, namely China, India, and Singapore, are also the countries with higher general trust in their governments (60–82% high trust, Ms 4.9–5.6/7). Similarly, the countries where people have the least confidence in government to develop, use and regulate AI, also have low general trust in government, namely South Africa, UK, USA, and Japan (53–65% low trust, Ms=2.7–3.1).

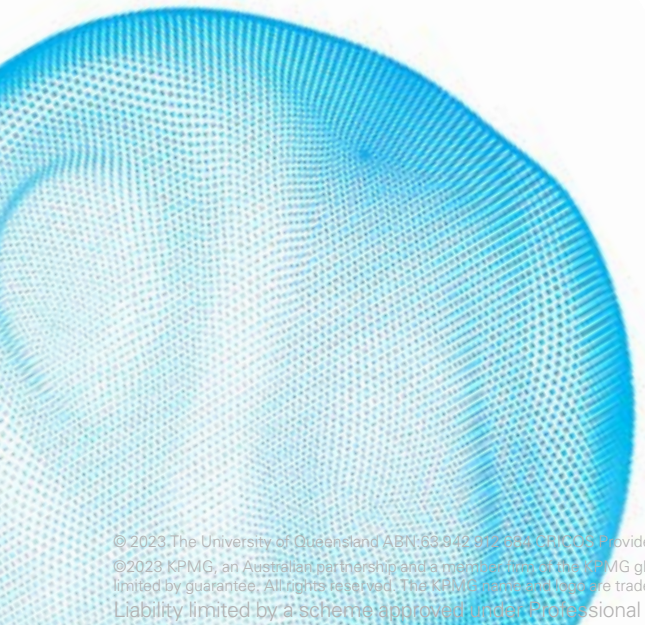
We also find high correlations (ranging between 0.55 and 0.70, p<0.001) between general trust in universities and research institutions, security forces, and business with confidence in each of these entities to develop, use and govern AI.

Younger generations, the university educated, and managers are more confident in entities to develop, use and govern AI

Generation X and Millennials are more confident in all entities, except government, to develop and regulate AI in the best interest of the public than Baby Boomers and older generations (42% vs 28% high confidence).

People with a university education are more confident in some entities to develop and regulate AI, particularly government (38% vs 23% high confidence), defence forces (50% vs 40% high confidence), and international research and scientific organisations (49% vs 38% high confidence).

Managers are more confident in government (44% high confidence vs 20–33%), commercial organisations (35% vs 19–23%), and technology companies (44% vs 25–31%) to develop and regulate AI than all other occupation groups.



TOPIC FOUR

What do people expect of the management, governance and regulation of AI?

We asked people about their expectations around AI management, governance and regulation, including the extent to which they think regulation is necessary, who should regulate, and whether current regulations and institutional safeguards are sufficient. We also asked what development and governance principles and practices are important for people to trust AI systems.

To contextualise these expectations, we first asked people their views about the impact of AI on society.

Sixty-one percent believe the societal impact of AI is uncertain

Most people (61%) believe the long-term impact of AI on society is uncertain and unpredictable (see Figure 19). The more uncertainty people perceive, the less likely they are to trust AI systems (r=-0.25, p<0.001).

While the majority of people in almost all countries agree the societal impact of AI is uncertain, people in the western countries of the USA, Australia, the UK and Canada perceive the greatest uncertainty (70–72%). In contrast, those in South Korea, Japan, Israel and Brazil perceive the least uncertainty (43–55%).

Figure 19: Perception that the impact of AI on society is uncertain

'To what extent do you agree with the following statements…?

(1) The impact of AI [specific application] is unpredictable
(2) The long-term impact of AI [specific application] on society is uncertain
(3) There is a lot of uncertainty around AI [specific application]'

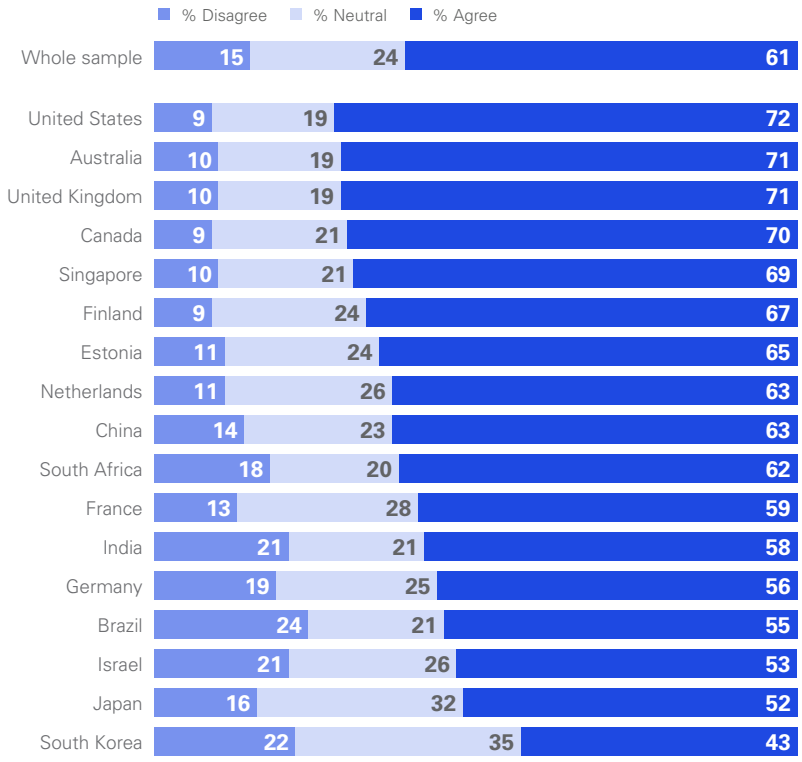
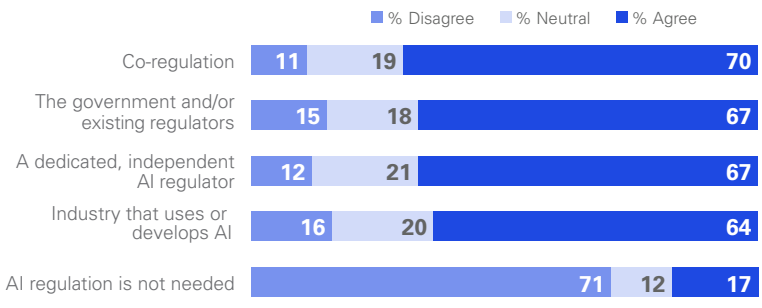


Figure 20: Expectations of who should regulate AI

'I think AI [specific application] should be regulated by....'



% Disagree = 'Somewhat disagree', 'Disagree', or 'Strongly disagree'
% Neutral = 'Neutral'
% Agree = 'Somewhat agree', 'Agree', or 'Strongly agree'

Most people believe AI regulation is required and expect some form of external, independent oversight

Given the perceived uncertain impact of AI on society, it is not surprising that most people across countries (71%) believe AI regulation is required. Less than one in five people (17%) believe AI regulation is not needed, with the remaining 12% unsure. This finding corroborates prior surveys indicating strong desire for the regulation of AI.[26]

People are broadly supportive of multiple forms of regulation. As shown in Figure 20, the majority of people (64–70%) expect a range of entities to be involved in regulating AI, including government and/or existing regulators, industry that uses or develops AI, co-regulation by industry, government, and existing regulators, and a dedicated, independent AI regulator.

Countries vary in expectations of who should regulate

In many countries, people express a preference for some form of independent regulation over regulation by industry (see Figure 21). For example:

- Australians prefer AI to be regulated by government and existing regulators, or by an independent AI body, rather than by industry (M=5.3 vs 4.6).
- An independent regulator is endorsed as a better option than industry regulation in the UK (M=5.4 vs 4.6), Germany (M=4.9 vs 4.5), and Finland (M=4.9 vs 4.4).
- Co-regulation is a preferred option compared to regulation by industry in the UK (M=5.1 vs 4.6), Canada (M=4.9 vs 4.5), Finland (M=4.9 vs 4.4), Israel (M=5.1 vs 4.5) and China (M=5.7 vs 5.3).

In contrast, in South Africa, all forms of regulation are seen as preferable compared to regulation by government (Ms=5.0–5.4 vs 4.5).

As shown in Figure 21 (black dots), people in India, China and Singapore are more likely to see AI regulation as unnecessary, compared to people in other countries. Specifically, a quarter or more of Singaporeans (25%), Chinese (37%) and Indians (39%) view AI regulation as not needed. However, except for India, most people in all other countries believe AI regulation is required, ranging from 56% in China to 83% in Israel.

Figure 21: Expectations of who should regulate AI across countries

* 7 point scale
[Countries sorted in order of 'AI regulation is not needed']

Only two in five people believe current safeguards are sufficient to make AI use safe

The majority (61%) of people disagree or are unsure that current safeguards around AI (i.e. rules, regulations, and laws) are sufficient to make the use of AI safe and protect them from problems (see Figure 22). This pattern was strongest in the western countries together with Israel, Japan, and South Korea. Only 39% of people believe that there are sufficient structural assurances around AI use. This finding corroborates previous surveys[27] reporting people do not think current rules are effective in regulating AI and is problematic given the strong relationship between current safeguards and trust in AI (r=0.66, p<0.001).

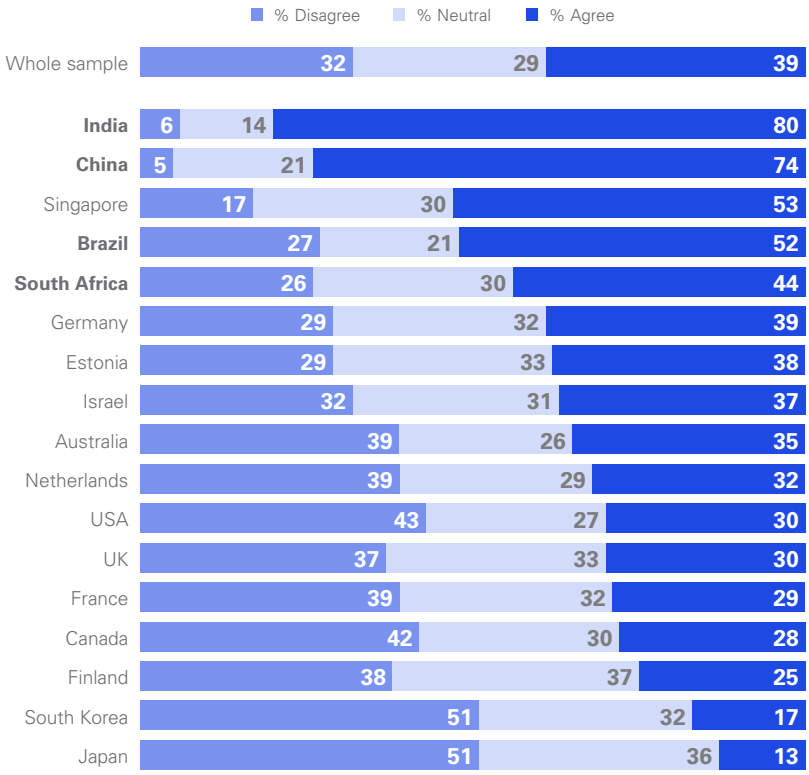
However, there are stark country differences. Most people in India (80%) and China (74%) believe appropriate safeguards are already in place, more than people in any other country. About half of people in Singapore (53%) and Brazil (52%) also believe current safeguards are sufficient. Conversely, less than one in five people in Japan (13%) and South Korea (17%) agree, with people in these countries rating current safeguards lower than all other countries.

The younger generations of Gen Z and Millennials (M=4.3) and Gen X (M=4.0) are more likely to believe there are sufficient safeguards in place to govern AI, compared to people in the Baby Boomers generation or older (M=3.8). Managers are also more likely to perceive sufficient safeguards than other occupations (M=4.6 vs 4.0–4.1)

Figure 22: Perception of current regulations, laws, and rules to make AI use safe

'To what extent do you agree with the following...

(1) There are enough current safeguards to make me feel comfortable with the use of AI [specific application]
(2) I feel assured that there are sufficient governance processes in place to protect me from problems that may arise from the use of AI [specific application]
(3) The current law helps me feel that the use of AI [specific application] is safe
(4) I feel confident that there is adequate regulation of AI [specific application]'



% Disagree = 'Somewhat disagree', 'Disagree', or 'Strongly disagree'
% Neutral = 'Neutral'
% Agree = 'Somewhat agree', 'Agree', or 'Strongly agree'

Assurance mechanisms enhance trust in AI systems

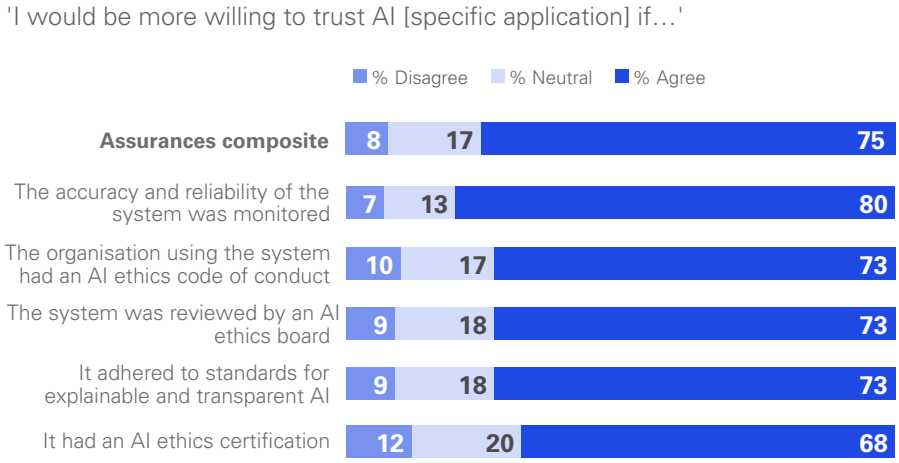
In addition to external rules, laws, and safeguards, we also asked people whether a range of assurance mechanisms available to organisations would influence their trust.

Three out of four people (75%) report they would be more willing to trust an AI system when assurance mechanisms are in place that support ethical and responsible use. These mechanisms include monitoring system accuracy and reliability, using an AI code of conduct, oversight by an independent AI ethical review board, adhering to standards for explainable and transparent AI, and an AI ethics certification (see Figure 23). These mechanisms increase perceptions of safeguards and reduce uncertainty.

Of the specific assurance mechanisms, four out of five people (80%) agree that system accuracy and reliability monitoring would enhance their trust, with fewer, but still two thirds (68%), agreeing that adherence to an AI ethics certification would enhance trust.

The assurance mechanisms influence trust most strongly in the BICS countries and Singapore, and the least in Japan.

Figure 23: AI assurance mechanisms



There is strong global endorsement of the Trustworthy AI management and governance principles: each principle is important for trust globally

A proliferation of reports and guidance documents on the development and deployment of trustworthy, ethical AI have been produced, with considerable consensus emerging on these principles.[28] One goal of this survey was to determine the extent to which these principles are important for people to trust in AI across the globe.

To answer this question, we asked about the importance of 16 practices reflecting the principles for trustworthy AI shown in Table 1. These principles primarily reflect the Principles for Trustworthy AI adopted by the European Union.[29]

Table 1: Principles and Practices for Trustworthy AI

Technical performance, accuracy and robustness

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

Data privacy, security and governance

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

Human agency and oversight

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

Transparency and explainability

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

Fairness, non-discrimination and diversity

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

Accountability and contestability

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

AI literacy

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

Risk and impact mitigation

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

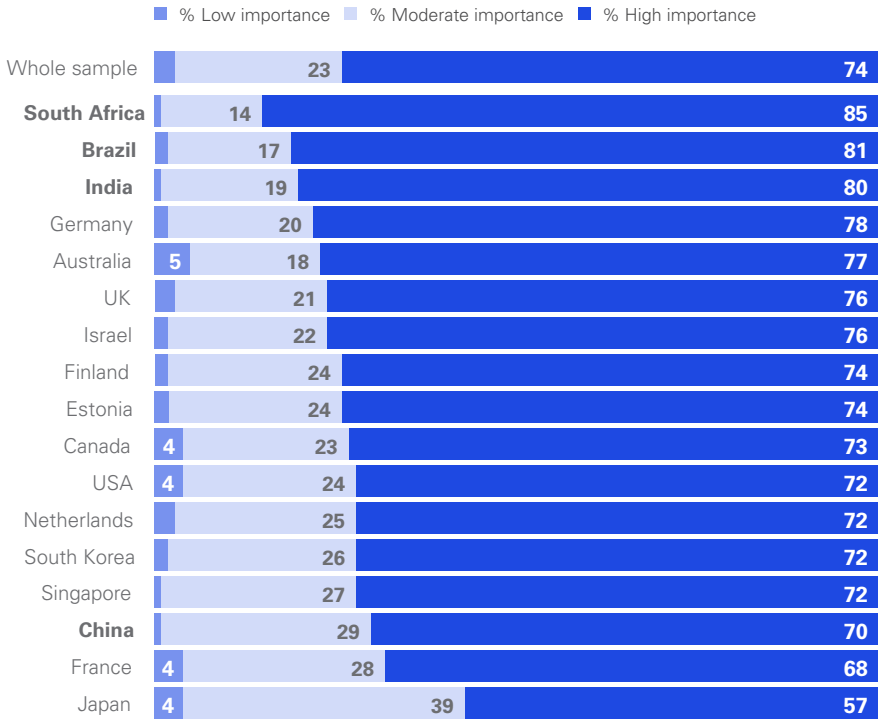
These principles were endorsed globally, with almost all people (96-99%) across all countries surveyed viewing these eight principles, and the practices that underlie them, as moderately to extremely important for trust in AI systems (see Figure 24). This finding held in relation to all AI use cases examined (i.e. AI systems in general and in healthcare, human resources, security, and recommender systems) suggesting their universal relevance.

People in South Africa viewed the principles as more important for trust than people in all other countries, other than Brazil.

Collectively these findings indicate clear public endorsement of these principles and practices and a blueprint for developing, using and governing AI in a way that supports trust across the globe.

Figure 24: Importance of the Principles for Trustworthy AI across countries

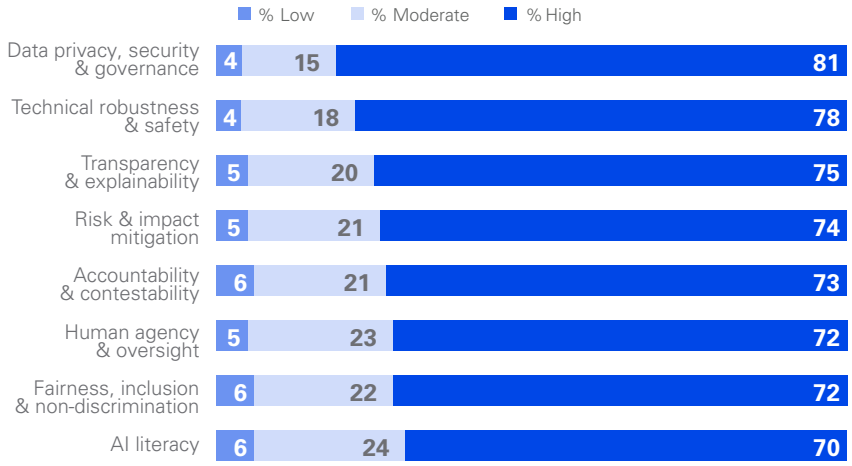
'How important are the following for you to trust AI [specific application]?



% Low = 'Not at all important' or 'Slightly important'
% Moderate = 'Moderately important'
% High = 'Highly important' or 'Extremely important'

Figure 25: Importance of the Principles for Trustworthy AI

'How important are the following for you to trust AI [specific application]?'



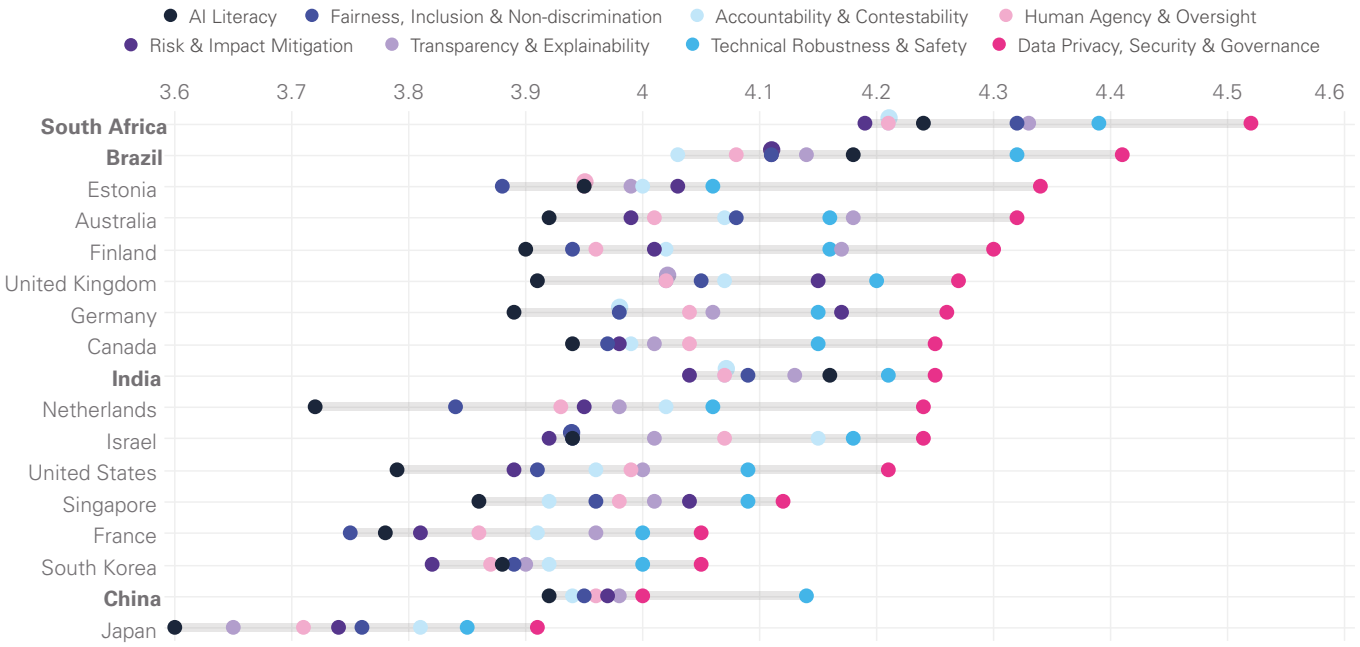
% Low = 'Not at all important' or 'Slightly important'
% Moderate = 'Moderately important'
% High = 'Highly important' or 'Extremely important'

Data privacy, security and governance is most important in all countries

While all eight principles are deemed important (see Figure 25), data privacy, security and governance practices are considered the most important for trust in AI systems in all countries, except China where it was ranked second (see Figure 26). In contrast, AI literacy practices are rated the least important in many countries.

Figure 26: Importance of the Principles for Trustworthy AI by country

'How important are the following for you to trust AI [specific application]?'



*5 point scale
1 = Not at all important, 5 = Extremely important
[Countries sorted by 'Data privacy, security & governance' category]

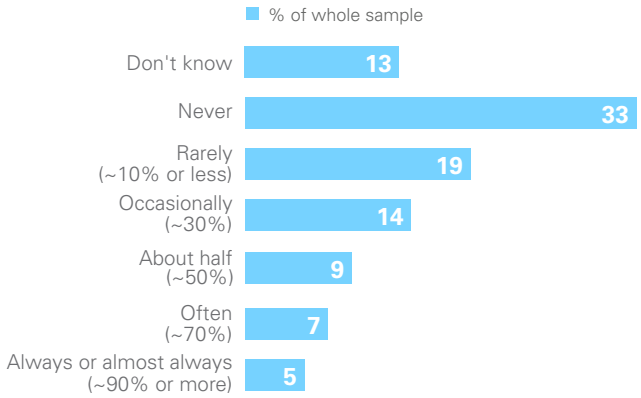
TOPIC FIVE

How do people feel about AI at work?

To answer this question, we asked people about their use of AI at work and by their employing organisation, their trust in AI for various work and decision making functions, and the impact of AI on work and jobs. Only people who were currently or previously employed completed these questions.[30]

Figure 27: Perceived use of AI in work

'How often do you use AI in your work?'



One in two people report using AI in their work

As shown in Figure 27, over half the surveyed people (54%) report using AI in their own work. A third report using AI rarely (19%) or occasionally (14%), and one in five (21%) report using it about half the time or more frequently.

In comparison, a third of the people surveyed believe they never use AI, and a further 13% 'don't know', suggesting they do not have sufficient understanding to gauge whether they use AI in their work.

AI use for work is most commonly reported in the BICS countries and least common in western countries

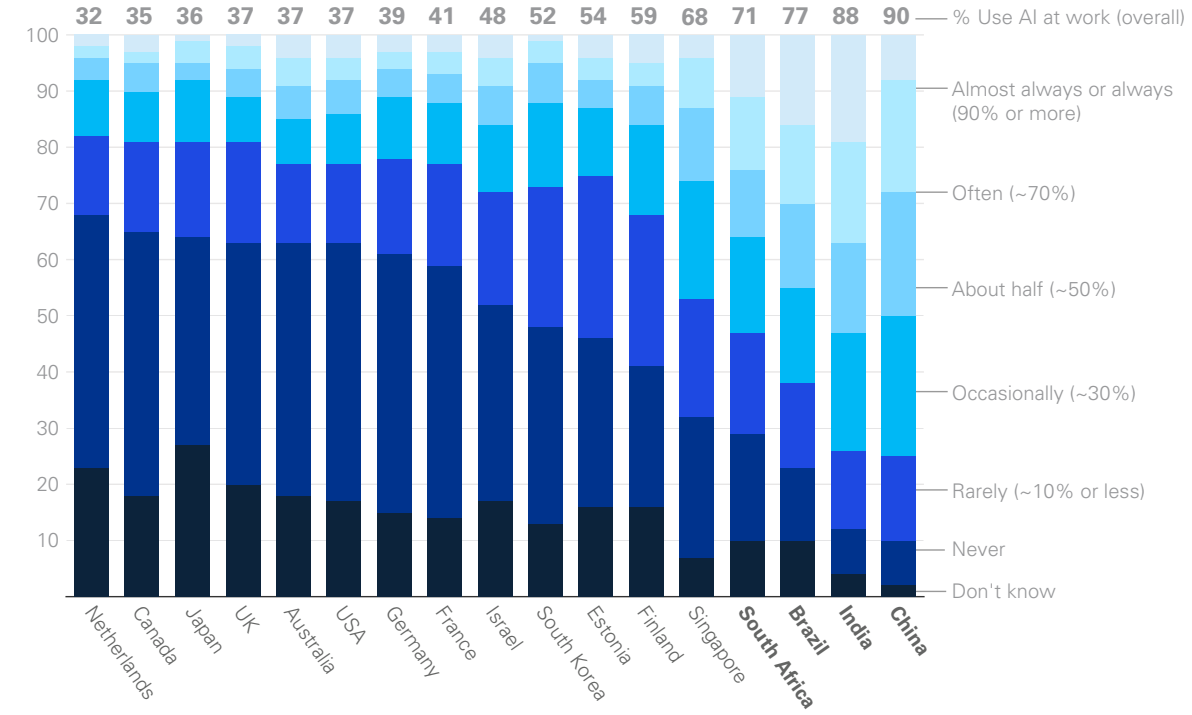
As shown in Figure 28, there are clear differences across countries in people's reported use of AI in their work. While the majority in the BICS countries (71–90%) report using AI in their work, a minority of people in most western countries (32–41%) report using AI. Exceptions are Finland and Estonia, where 54–59% of people report using AI at work.

Finland and Estonia are known for their efforts to educate people about AI through public literacy programs.

This pattern suggests that the low reporting of AI use at work may partially reflect a general lack of understanding and awareness of when and where AI is embedded in everyday technologies (e.g. internet searches, email filters etc) in other western countries (see next section). Indeed, there is a correlation between one's knowledge of when and where AI is being used and perceived use of AI at work (r= 0.32, p<0.001).

Figure 28: Reported use of AI at work across countries

'How often do you use AI in your work?'

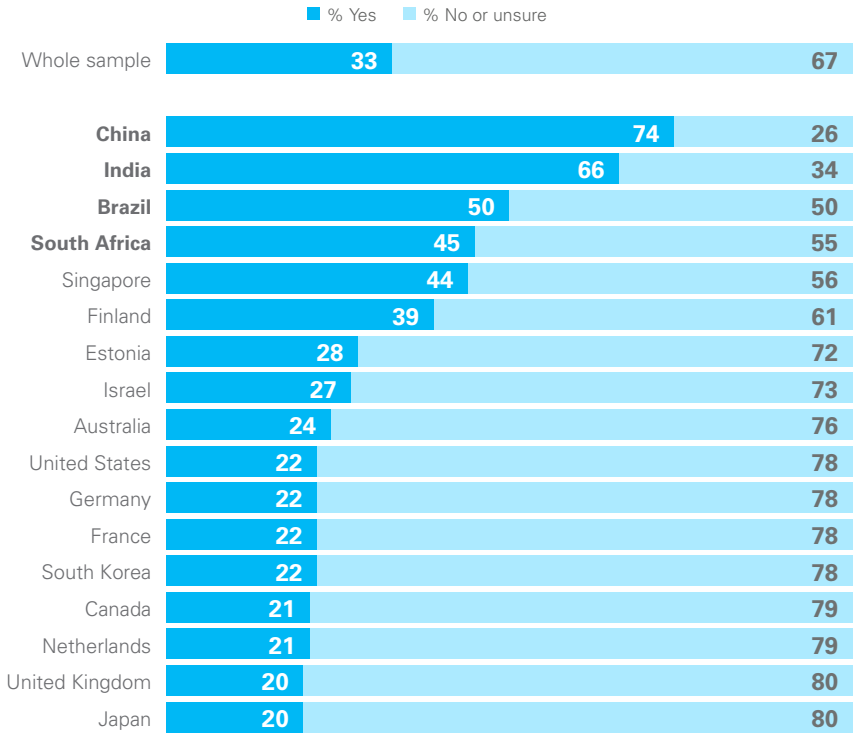


We also asked respondents whether AI is used by their employing organisation (see Figure 29). Overall, only a third of people surveyed believe AI is used by their employing organisation.

People in the BICS countries, Singapore and Finland report greater AI use in their organisations than people in other countries. This likely reflects a greater understanding of AI and when and how it is used by organisations in these countries.

Figure 29: Use of AI in my employing organisation

'Is AI used in this organisation?'

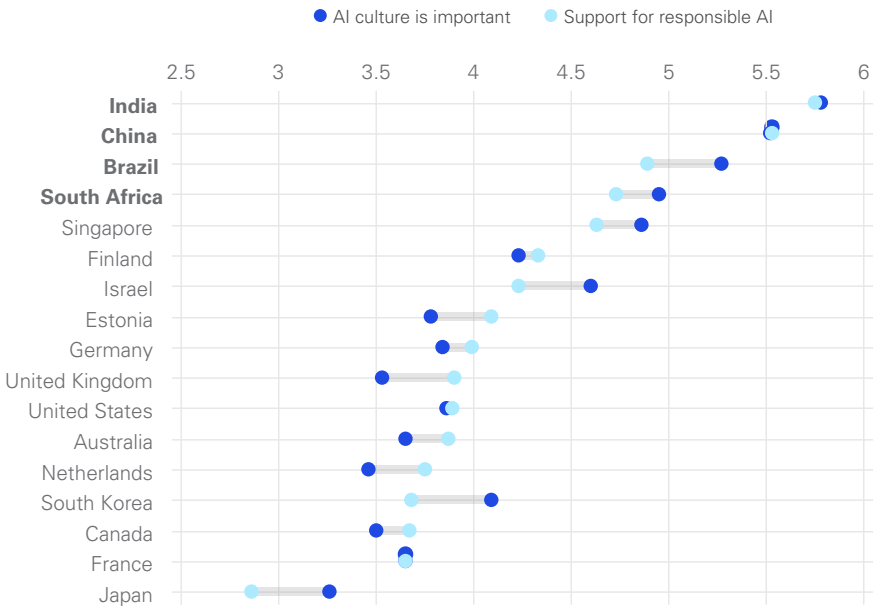


People in the BICS countries and Singapore rate their employers as more supportive of AI and its responsible use, than people in other countries

As shown in Figure 30, the majority of people in the BICS countries and Singapore (83–85%) report their employing organisation considers AI adoption strategically important and recognises efforts to integrate AI (i.e. AI culture). The majority (60–88%) also report their employing organisation supports the responsible use of AI.

In contrast, a minority of people (15–49%) in western countries and the Asian countries of Japan and South Korea perceive their organisation has these attributes. People in Israel and Finland tend to view their organisations higher on AI culture and support for responsible AI use, compared to these other countries.

Figure 30: AI culture and support for responsible AI in my employing organisation



* 7 point scale
[Countries sorted by 'Support for responsible AI' category]

Most people view AI use in managerial decision-making as acceptable, and prefer AI involvement to decision-making by humans only

To understand people's views of the use of AI in managerial decision-making, we asked people to choose the most acceptable weighting between human and AI involvement in decision-making.[31]

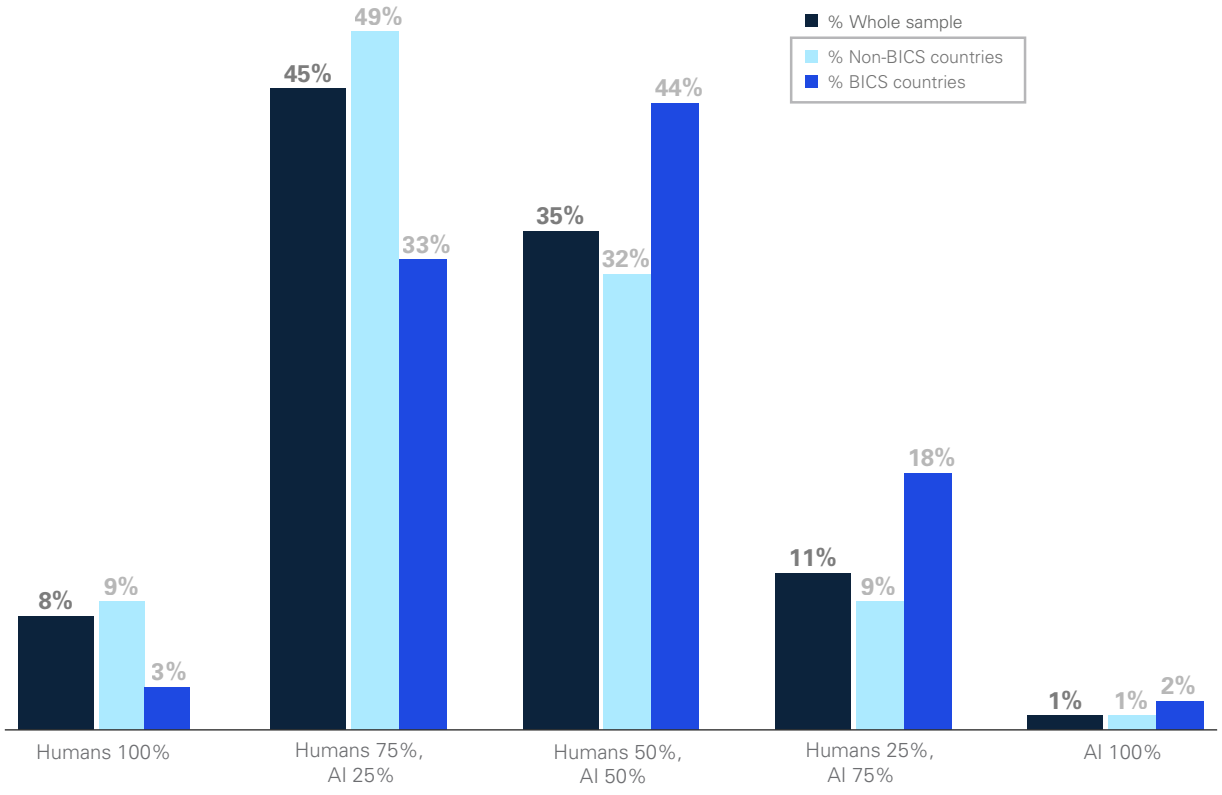
As shown in Figure 31, most people believe AI should have a role in managerial decision-making, with only a few people (8%) advocating for humans to have sole decision-making power. Almost half (45%) view a collaboration of 75% human/25% AI split in decision-making to be the most acceptable configuration. The next most popular proposal is a 50/50 split, advocated by a third of people (35%).

However, few people (12%) believe AI should be dominant in managerial decision-making. This demonstrates a lack of support for fully automated managerial decision-making or a dominance of AI over humans in decision-making.

People in BICS countries are more likely to view a 50/50 split as the most acceptable configuration, whereas in all other countries a 75% human/25% AI split was the most popular option.

Figure 31: Human–AI collaboration in managerial decision-making

'Which of the following proposals do you find most acceptable for human manager-AI collaboration in managerial decision-making activities?'



* BICS countries include Brazil, India, China, and South Africa

Most people are comfortable with AI use to augment and automate tasks but not for HR and people management

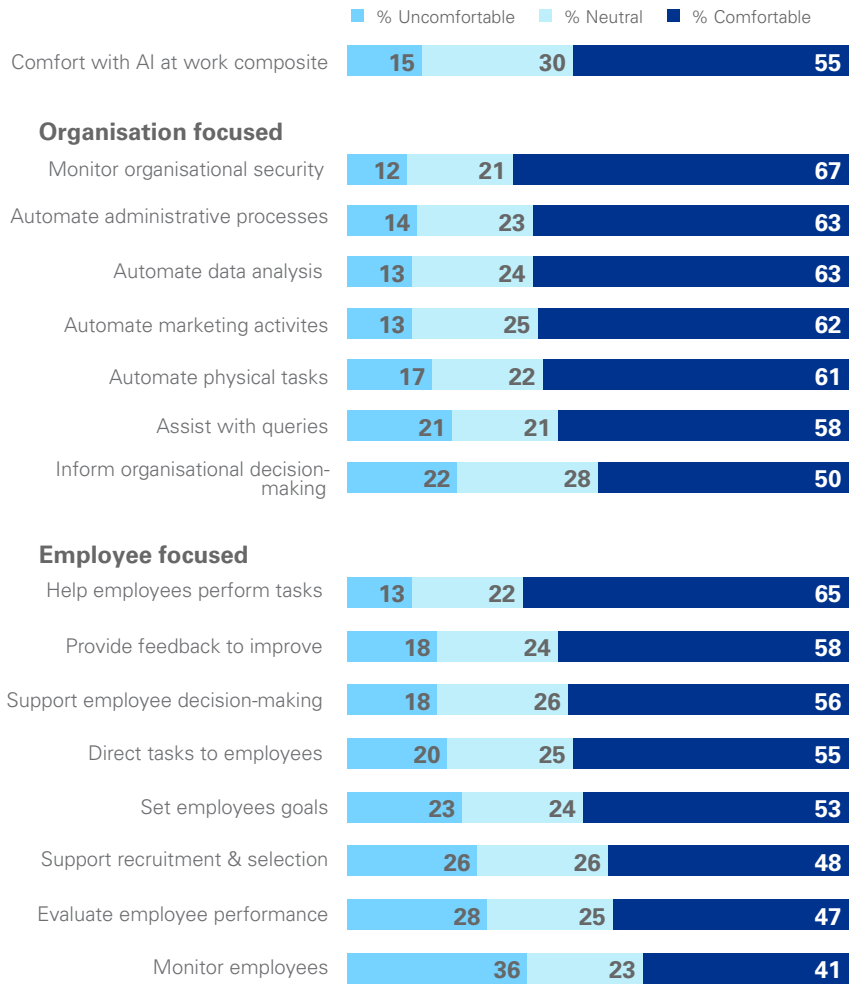
We asked respondents how comfortable they are with the use of AI for a range of activities at work. While over half (55%) feel comfortable with the use of AI at work, comfort varies depending on the purpose or application of the AI in the workplace.

Over half of people surveyed feel comfortable with the use of AI for organisation-focused purposes to augment and automate tasks. This includes tasks such as monitoring security, automating administrative, analytic, marketing, and physical tasks, and assisting with queries. People are less comfortable with the use of AI for informing organisational decision-making.

We also asked about the use of AI for employee-focused activities. Here we found a distinction between AI use for augmenting employees at work, and AI use for human resource management purposes. Most people are comfortable with AI use for the purpose of augmenting employee performance and decision-making, for example by providing feedback on how to improve performance and supporting employees to make decisions. People are notably less comfortable with AI use for human resource management, such as to monitor and evaluate employees, and support recruitment and selection processes.

Figure 32: Comfort with the use of AI at work

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



% Uncomfortable = 'Somewhat', 'Mostly', or 'Completely uncomfortable'
% Neutral = 'Neutral'
% Comfortable = 'Somewhat', 'Mostly', or 'Completely comfortable'

There are notable differences across countries. People in the BICS countries and Singapore are the most comfortable with AI use at work, with 67-87% comfortable across the various AI work applications (see Figure 33).

This is consistent with the pattern of higher use of AI at work by people in these countries. In contrast, people in France, Germany, Japan, Australia, and Canada are the least comfortable (33–46% across AI work applications).

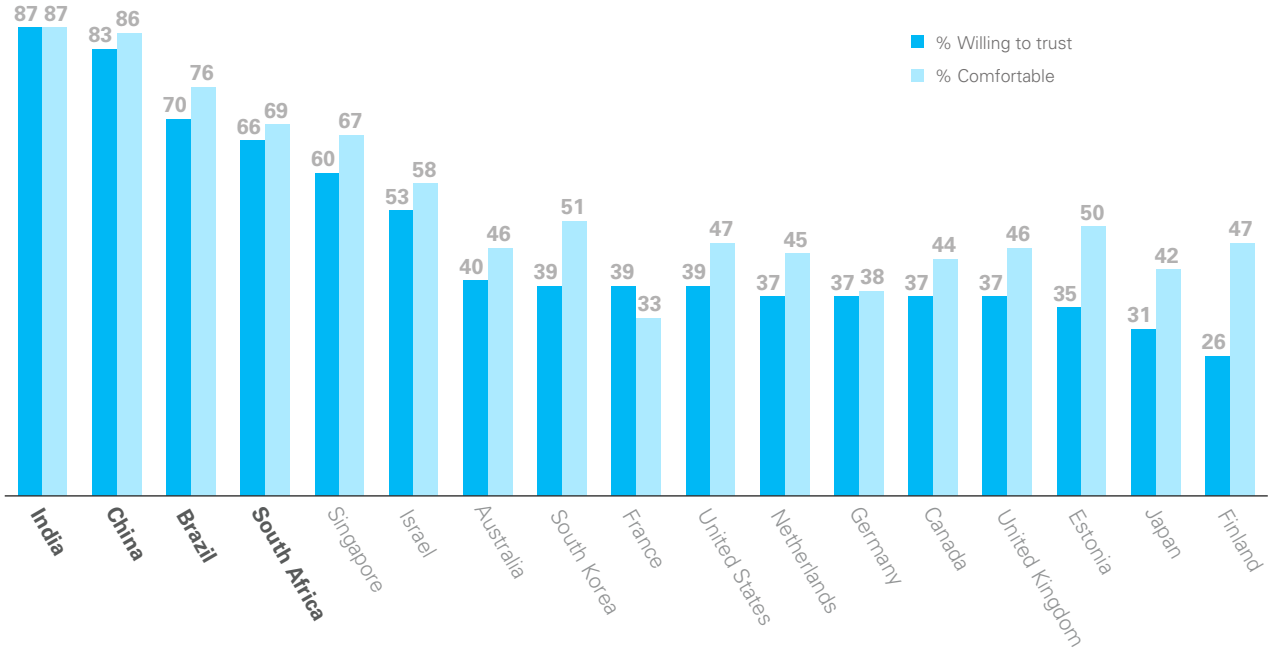
Almost half of people trust AI at work, with trust highest in the BICS countries

We asked people how willing they are to trust AI systems at work by using the systems for work purposes, relying on the information and decisions they provide, and allowing data and relevant information about themselves to be used.

We found almost half (48%) of people surveyed are willing to trust AI at work, with stark differences between countries. Trust is highest in the BICS countries, with two-thirds or more of people (66–87%) in these countries trusting AI at work, significantly more than all other countries. Singapore and Israel had the next highest levels of trust at work.

In contrast, people are the least willing to trust AI systems at work in the western countries, Japan, and South Korea, with only 26–40% willing to trust. People's trust in AI at work is strongly associated with their trust in AI systems more broadly (r=0.71, p<0.001).

Figure 33: Trust and comfort with the use of AI at work across countries



% Willing to trust = 'Somewhat', 'Mostly', or 'Completely willing'
% Comfortable = 'Somewhat', 'Mostly', or 'Completely comfortable'

Half of people believe AI will improve their competence and autonomy at work

We asked people how AI use would affect their sense of competency, autonomy, and relatedness at work. Prior research shows the fulfilment of these three basic psychological needs at work enhances people's wellbeing, motivation, and performance.[32] In contrast, when these needs are unfulfilled, it can result in maladaptive behaviours and loss of motivation.

Almost one in two people believe the use of AI in their work would make them feel more competent and effective in their job (49%) and enhance their autonomy (48%) by giving them more choice in how they do their work.

Almost two in five (38%) believe using AI would help them feel a sense of relatedness and connection with other people and groups at work, with a third unsure and a little under a third disagreeing.

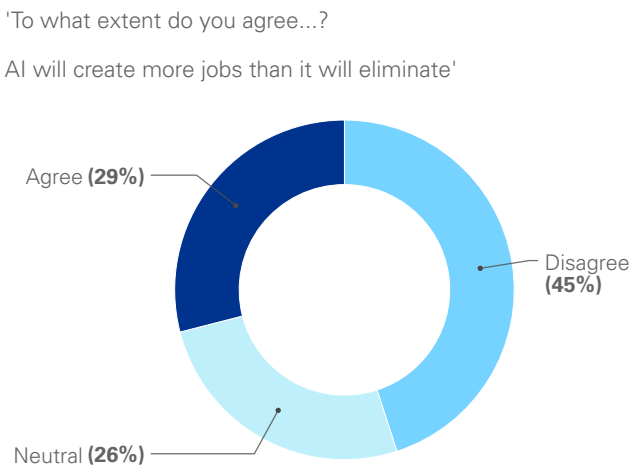
People in the BICS countries had more positive views about how AI would impact their competency, autonomy, and relatedness at work, compared to people in other countries. For example, over 80% of people in India believe AI will increase their effectiveness, choice, and connection to others at work. In contrast, in many western countries, a third or fewer people believed AI would result in these positive effects.

Fewer than one in three people believe AI will create more jobs than it will eliminate

Most people (71%) disagree or are unsure that AI will create more jobs than it will eliminate (see Figure 34). This finding supports prior surveys reporting concerns about job loss from AI and automation.[33]

People in China, India, Brazil, and Singapore are the most positive about job creation from AI and more positive than all other countries, yet there is significant variation even among these more optimistic countries. For example, about two-thirds (63–67%) of people in China and India, and 37–48% of people in Singapore and Brazil, agree AI will create more jobs than it will eliminate, compared to less than 30% in other countries.

Figure 34: Perceived impact of AI on jobs generally



% Disagree = 'Somewhat disagree', 'Disagree', or 'Strongly disagree'
% Neutral = 'Neutral'
% Agree = 'Somewhat agree', 'Agree', or 'Strongly agree'

As shown in Figure 35, people are split in their beliefs about whether AI will impact jobs in their area of work, with 42% believing AI will replace jobs in their area, 39% disagreeing and 19% unsure.

People in the BICS countries, as well as Singapore, South Korea, and Israel, are more likely to report that AI will replace jobs and key aspects of work in their area (48-74% agree).

Younger people and the university educated are more likely to report using AI at work and be more trusting and comfortable with AI at work

Younger people are more trusting and comfortable with AI use at work than older respondents. As shown in Figure 36, 65% of Generation Z and Millennials are trusting of AI at work, compared to 39% of older respondents (Baby Boomers and older), and 60% of younger generations are comfortable with the use of AI at work, compared to 46% of older respondents.

Similarly, the university educated are more trusting (56% vs 40%) and comfortable (62% vs 48%) with AI use at work, than those without a university degree.

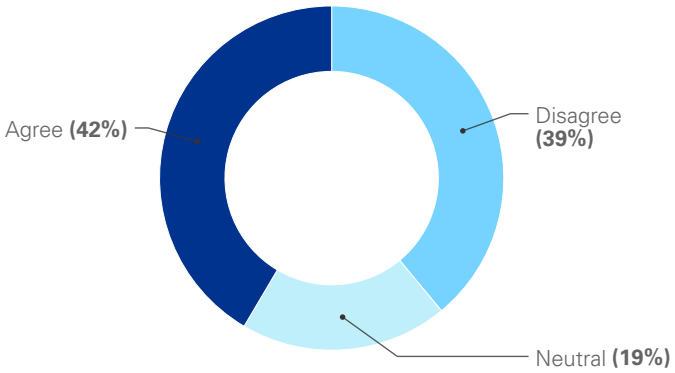
This pattern most likely reflects the fact that younger generations and the university educated are also more likely to report using AI in their own work and more likely to report that the organisation they work for uses AI, has an AI culture, and supports the responsible use of AI (see Figure 36).

There are no gender differences in the use of AI at work or attitudes towards AI at work.

Figure 35: Perceived impact of AI on my work

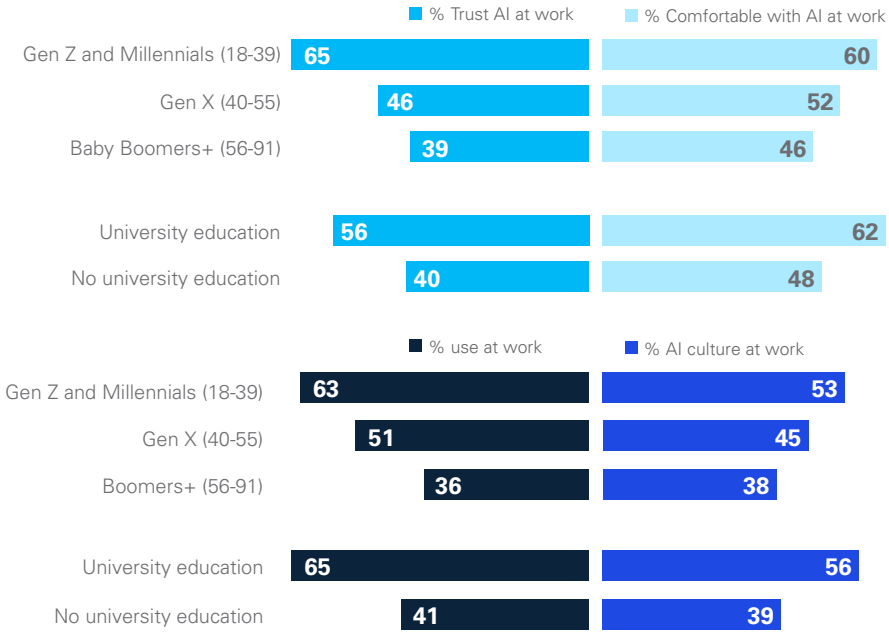
'To what extent do you agree...?

(1) AI will replace jobs in my area of work
(2) Key aspects of my work could be performed by AI'



*Amalgamated global mean responses
% Disagree = 'Somewhat disagree', 'Disagree', or 'Strongly disagree'
% Neutral = 'Neutral'
% Agree = 'Somewhat agree', 'Agree', or 'Strongly agree'

Figure 36: Generational and education differences in perceptions of AI at work

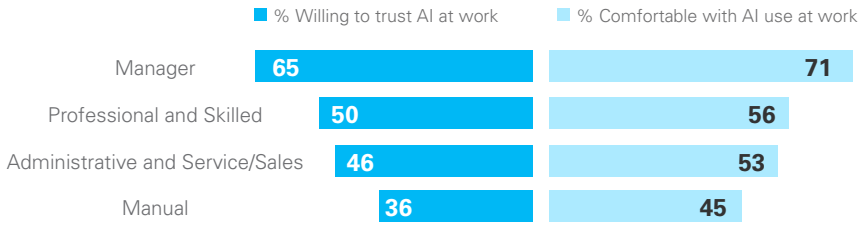


Managers are more trusting and comfortable with the use of AI at work than non-managers: manual workers are the least trusting and comfortable

As shown in Figure 37, managers are more willing to trust AI use at work than all other occupations (65% vs 36–50% willing), and are also more comfortable with its use (71% vs 43–56% comfortable). Of the codable occupations,[34] manual workers are the least trusting of (36%) and comfortable with (45%) AI use at work.

As with younger respondents and those with a university education, the trust and comfort managers have in the use of AI at work may reflect that they are more likely to work for organisations that use AI (51% vs 20–38%), believe their organisations have an AI culture (65% agree vs 35–51%) and support responsible AI use (66% agree vs 40–51%), and are more likely to use AI in their own work (72% vs 33–61%). Conversely, manual workers are less likely to work for organisations that use AI (20%), believe their organisations have an AI culture (35%), or use AI in their own work (40%) than other occupations.

Fig 37: Occupational differences in perceptions of AI at work



Younger generations and the university educated are more likely to believe AI will create jobs, even though it can replace aspects of their work

The younger generations of Gen Z and Millennials are more likely than Baby Boomers and older generations to believe key aspects of their work can be performed by AI (49% vs 34%) but also that AI will create more jobs than it will take away (34% vs 22% agree).

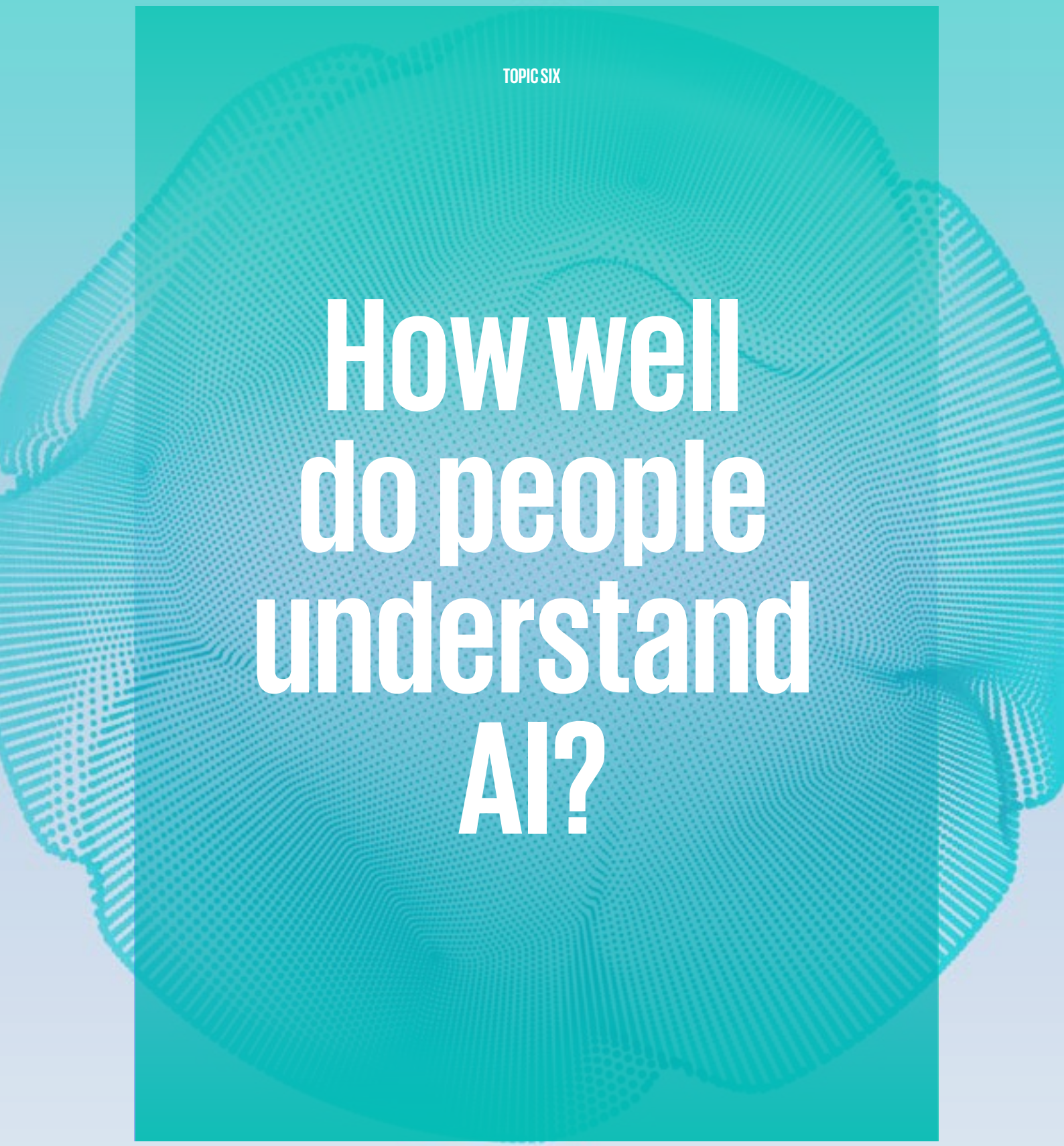
Similarly, people with a university education are also more likely than those without a degree, to believe key aspects of their work can be performed by AI (49% vs 37%) and that AI will create jobs (36% vs 22% agree).

Managers are more likely to believe AI will create jobs than people in all other occupations

Managers are more optimistic that AI will create more jobs than it will eliminate compared to all other occupations (43% agree vs 22–31%). However, managers are also more likely to agree that AI could perform key aspects of their work than all other occupations (56% agree vs 31–41%).

Managers (48% agree) and administrative, sales and service workers (43%) are more likely to believe AI will replace jobs in their area of work than all others (34–38%).

Manual workers are least likely to believe AI will create jobs (22%), perform key aspects of their work (31%), or replace jobs in their area of work (34%).

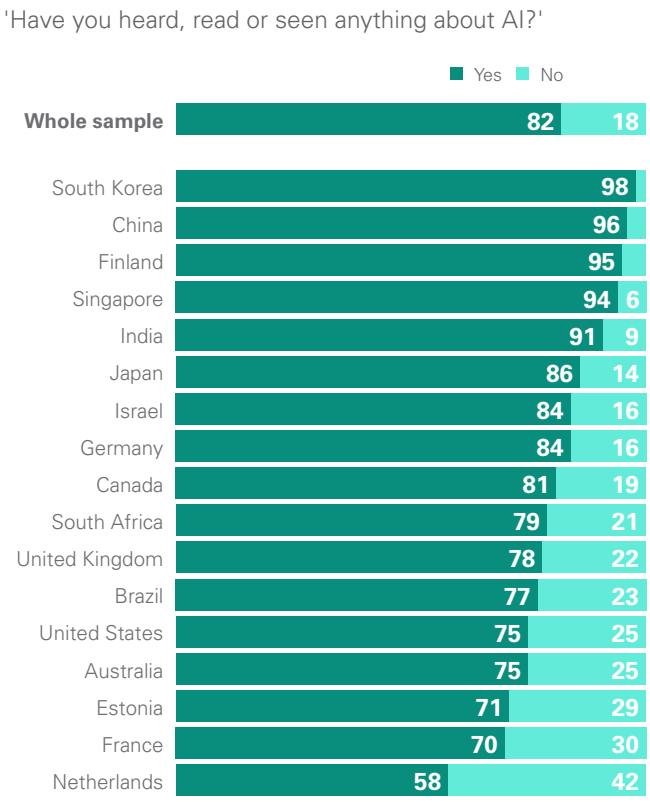


TOPIC SIX

How well do people understand AI?

To identify how well people understand AI, we asked about AI awareness, subjective and objective knowledge of AI and interest to learn more.

Figure 38: Awareness of AI across countries



Four out of five people are aware of AI: Awareness is highest in Asian countries in Finland

Eighty-two percent of people across all countries had heard, read or seen something about AI.

As shown in Figure 38, awareness of AI varies across countries. Asian countries and Finland have the highest levels of AI awareness, with people from the Netherlands reporting the lowest awareness. The high rates in Finland compared to other western nations may partially reflect investment in public AI education created and championed in Finland.[35]

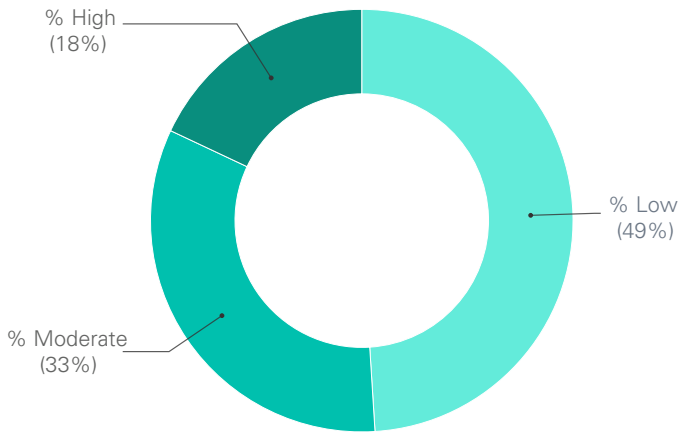
One in two people don't feel they understand AI or when it is used

Half the people (see Figure 39) have low subjective knowledge of AI, reporting that they do not understand AI or when and how it is used.[36] A smaller proportion (18%) report high subjective knowledge of AI, and a third report moderate understanding.

Subjective knowledge is associated with trust in AI (r=0.36, p<0.001), as well as technology efficacy, one's self-assessed ability to use digital technologies effectively (r=0.34, p<0.001). Technology efficacy is also associated with trust (r=0.42, p<0.001).

Figure 39: Subjective knowledge of AI

'To what extent do you...
(a) Feel you know about AI
(b) Feel informed about how AI is used
(c) Think you understand when AI is being used'



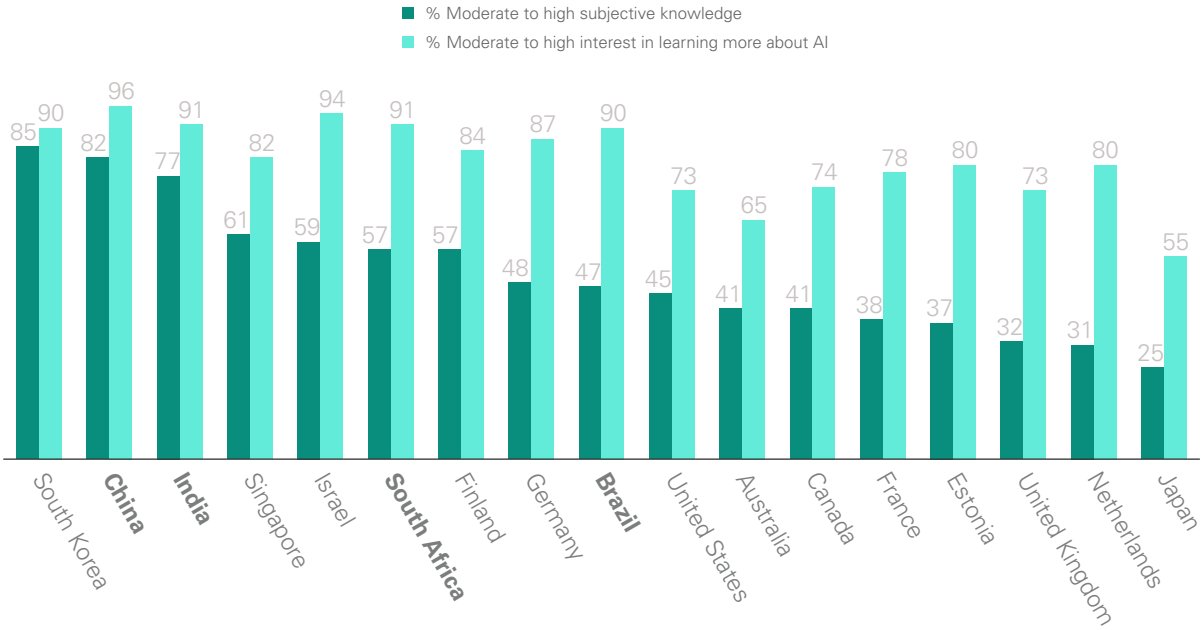
Low = 'Not at all' or 'To a small extent'
Moderate = 'To a moderate extent'
High= 'To a large extent' or 'To a very large extent'

Subjective understanding of AI is highest in Asian countries

There are stark differences in subjective knowledge across countries. As shown in Figure 40, people in Asian countries tend to report greater subjective understanding of AI, particularly people in China and India.

A notable exception is Japan: people in Japan reported the lowest level of subjective AI understanding. This may reflect cultural differences in orientation to technology, as well as age, noting that Japan has the world's oldest population. People in the Netherlands, UK, France, and Estonia also rate their knowledge as lower than people in other countries.

Figure 40: Subjective knowledge of AI and interest in learning more about AI



Moderate to high subjective knowledge = 'to a 'moderate', 'large' or 'very large' extent'
Moderate to high interest in learning more about AI = 'to a 'moderate', 'large', or 'very large' extent'

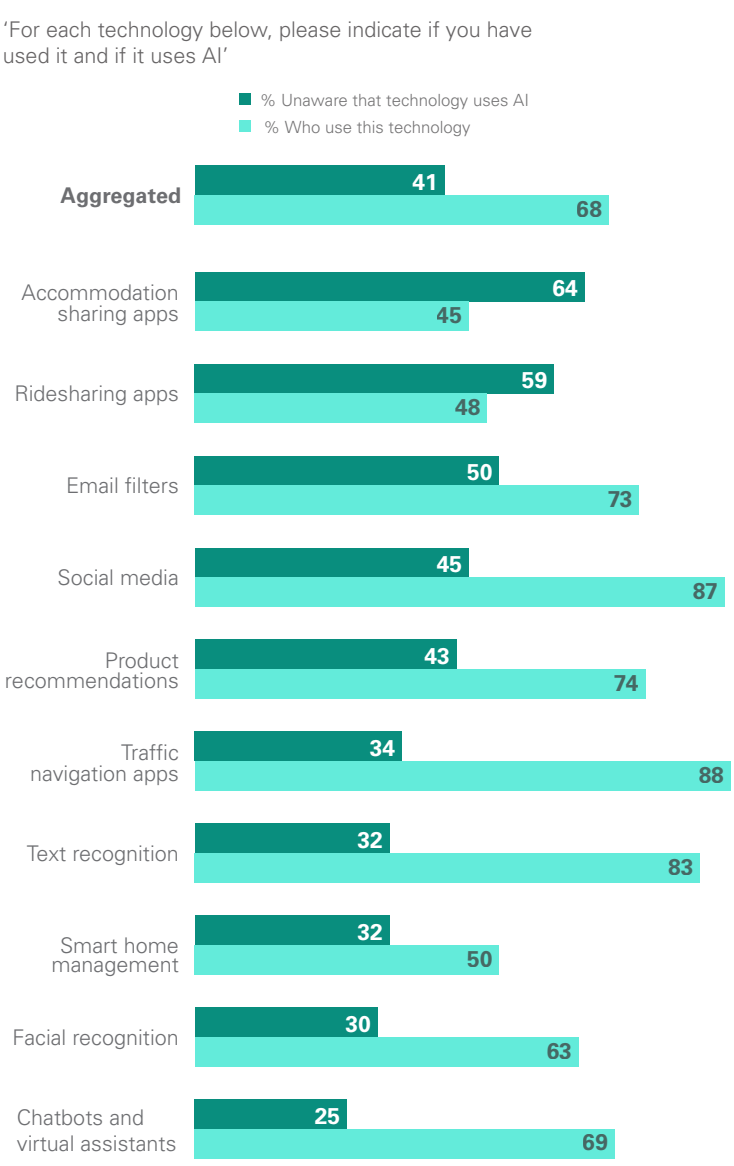
Most people want to learn more about AI with the greatest desire in BICS countries

Most people (82%) are interested in learning more about AI. Only 19% report no or low interest in learning more about AI (ranging from 4% of Chinese to 45% of Japanese).

People in the BICS countries, South Korea and Israel have the strongest desire to learn more about AI, with over 90% of people interested in these countries. Of the Western countries, people in Germany and Finland report the strongest interest.

People in Japan and Australia have notably lower interest in learning about AI compared to other countries. Only 55% of people in Japan and 65% of people in Australia expressed a desire to learn more about AI.

Figure 41: Use of AI technologies and awareness of their use



* % Unaware that technology uses AI includes both "No" and "Don't know" responses

Two out of five people are unaware that AI enables common applications they use: 45% don't know AI is used in social media

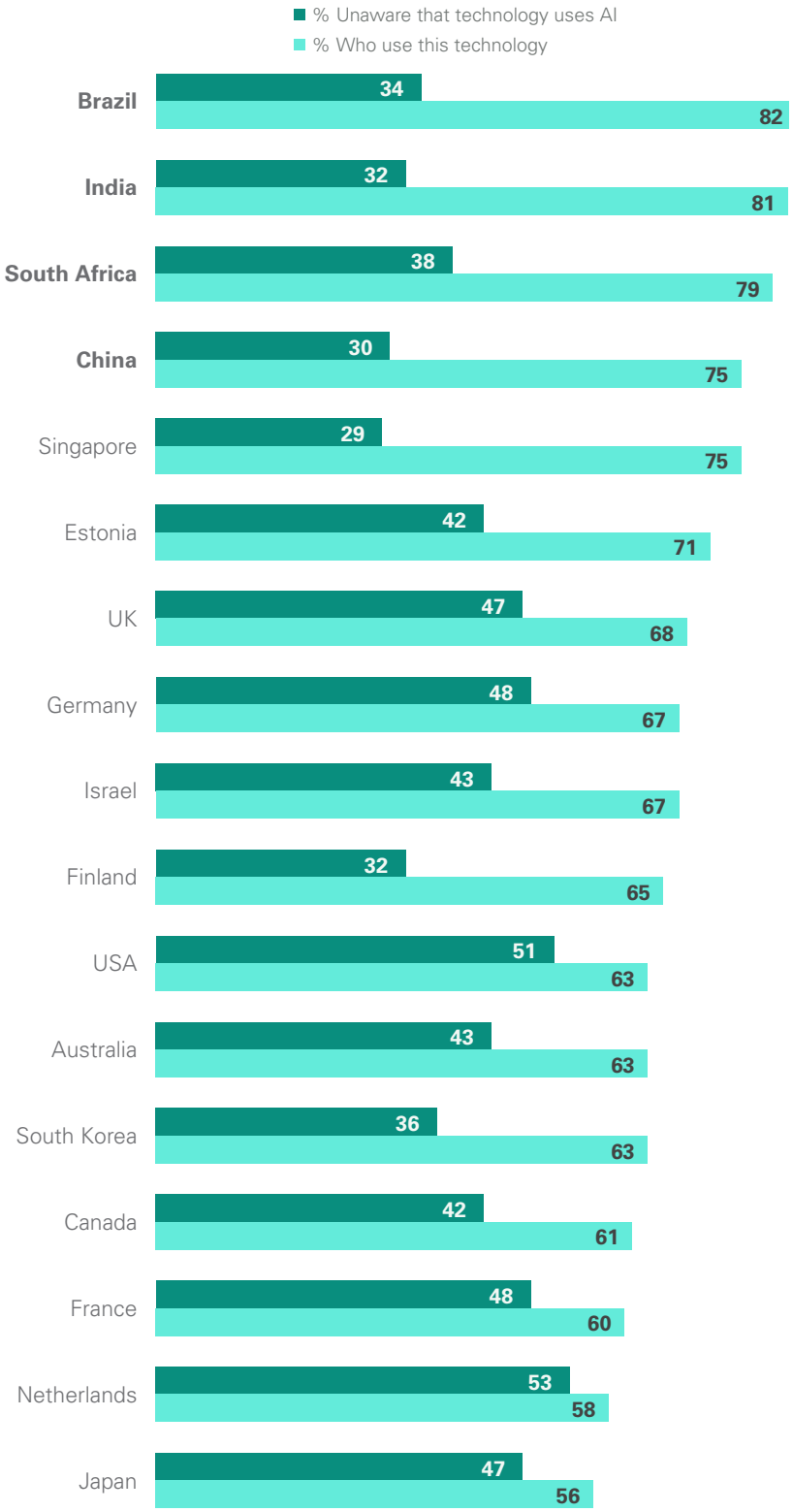
As an indicator of people's objective awareness of AI use, we asked if the common technologies shown in Figure 41 use AI. Overall, we found that although 68% of people had used these common AI-enabled technologies, two in five people (41%) were unaware that these technologies use AI. There is a small association between objective awareness of AI use and trust in AI (r=0.20,p<0.001).

The use of the technology did not necessarily translate into an increased understanding of whether AI is part of the technology. For example, while 87% of people report using social media, 45% are unaware that social media relies on AI. As shown in Figure 41, this pattern of using an application without awareness that it relies on AI, is particularly strong for social media, email filters and accommodation and ride sharing apps. These are all examples of embedded forms of AI (i.e. AI use with no physical or vocal manifestation).

In contrast, there is more awareness of AI when it is used in embodied forms (e.g. with voice or voice activation) such as virtual assistants (75%) and smart home management (68%), as well as when AI is used for facial recognition (70%).

Figure 42: Use of AI technologies and awareness of their use across countries

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



* '% Unaware that technology uses AI' and '% Who uses technology' reflect an average of 10 different technologies for each country.

The finding that 68% of people across countries report using AI-enabled applications demonstrates the high global penetration of AI into people's daily lives. However, there are notable differences across countries.

As shown in Figure 42, people in the BICS countries and Singapore are more likely to use these common AI-enabled applications than people in other countries. Awareness of AI use is also higher in three of the BICS countries (Brazil, India, and China, but not South Africa), as well as Singapore and Finland, than in all other countries. On average two-thirds of people or more (66–71%) in these countries are aware of AI use in common applications.

In contrast, people in the European countries of the Netherlands, France, Germany, and the UK, together with the USA and Japan, were less knowledgeable of AI use, with only 47–53% aware of AI use in these common applications.

People with a better understanding of AI are more likely to perceive greater benefits

Subjective knowledge of when and where AI is being used is associated with perceived benefits of AI (r=0.37, p<0.001) but has only a very small association with perceived risks (r=0.04, p<0.001). A similar pattern emerges for objective awareness of AI with perceived benefits (r=0.22, p<0.001) and perceived risks (r=0.02, p=0.019).

Younger generations, the university educated, and managers are more knowledgeable of AI and more interested to learn about AI

Younger generations and those with a university education have greater knowledge of AI, both subjectively and objectively (i.e. are better able to identify when AI is used), and have greater interest in learning about AI, than older generations and those without a university education (see Figure 43).

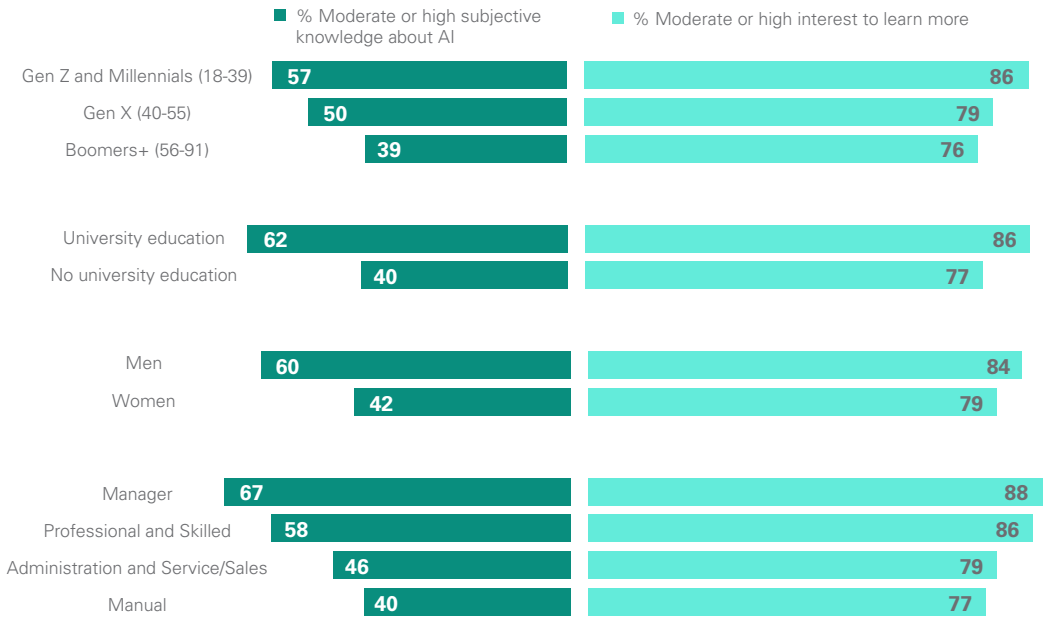
Over half (57%) of Gen Z and Millennials report they understand AI, compared to only 39% of Baby Boomers and older generations. Gen Z and Millennials are also more likely to correctly identify AI use in common applications than Baby Boomers and older: 48% of older respondents are unable to identify AI use across applications, compared to 39% of younger respondents. Younger generations are also more likely to express interest in learning more about AI (86% vs 76%).

We see a similar pattern across education levels. Fifteen percent more people with a university education have heard of AI compared to people without a university education. Sixty-two percent of the university educated feel they understand AI compared to only 40% of those without a university education. Those with a university education are also more likely to correctly identify AI use in common applications than those without a university education: 34% of those with a university education are unable to identify AI use across applications, compared to 48% of those without a university education. The university educated are also more likely to express interest in learning more about AI (86% vs 77%).

There is a significant gender gap in awareness of AI, and subjective but not objective understanding of AI. Men are more likely to have heard of AI than women (87% vs 77%), and report more subjective knowledge about AI than women (60% vs 42%). This is one of the few gender differences found. But men and women do not differ in their ability to objectively identity AI use in common applications, or in their interest to learn more about AI.

Managers report having a stronger understanding of AI than other occupation groups (67% vs 40-58%), and are more aware of AI use across common applications than other occupations except professional and skilled workers. Managers are also more interested in learning more about AI than all other occupations except professional and skilled workers (88% vs 77-79%).

Figure 43: Demographic differences in subjective knowledge of AI and interest to learn about AI



TOPIC SEVEN

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

To identify the most important drivers of trust and acceptance of AI systems examined in this report, we used a statistical technique called structural equation modelling.[37] We explain the model in Figure 44, together with notes on interpreting the model.

Trust is central to AI acceptance

The model shows that trust is a key driver of AI acceptance (B=0.87[38]), and explains 74% of the variance in acceptance. This finding empirically supports why trustworthy AI matters: if people perceive AI systems as trustworthy and are willing to trust them, then they are more likely to accept these technologies.

Trust acts as the central mechanism through which other drivers impact AI acceptance. It is important, therefore, to understand what influences trust in AI systems. To do this, we examine four distinct "pathways to trust" – institutional, motivational, uncertainty reduction, and knowledge – and examine their comparative importance in predicting trust[39].

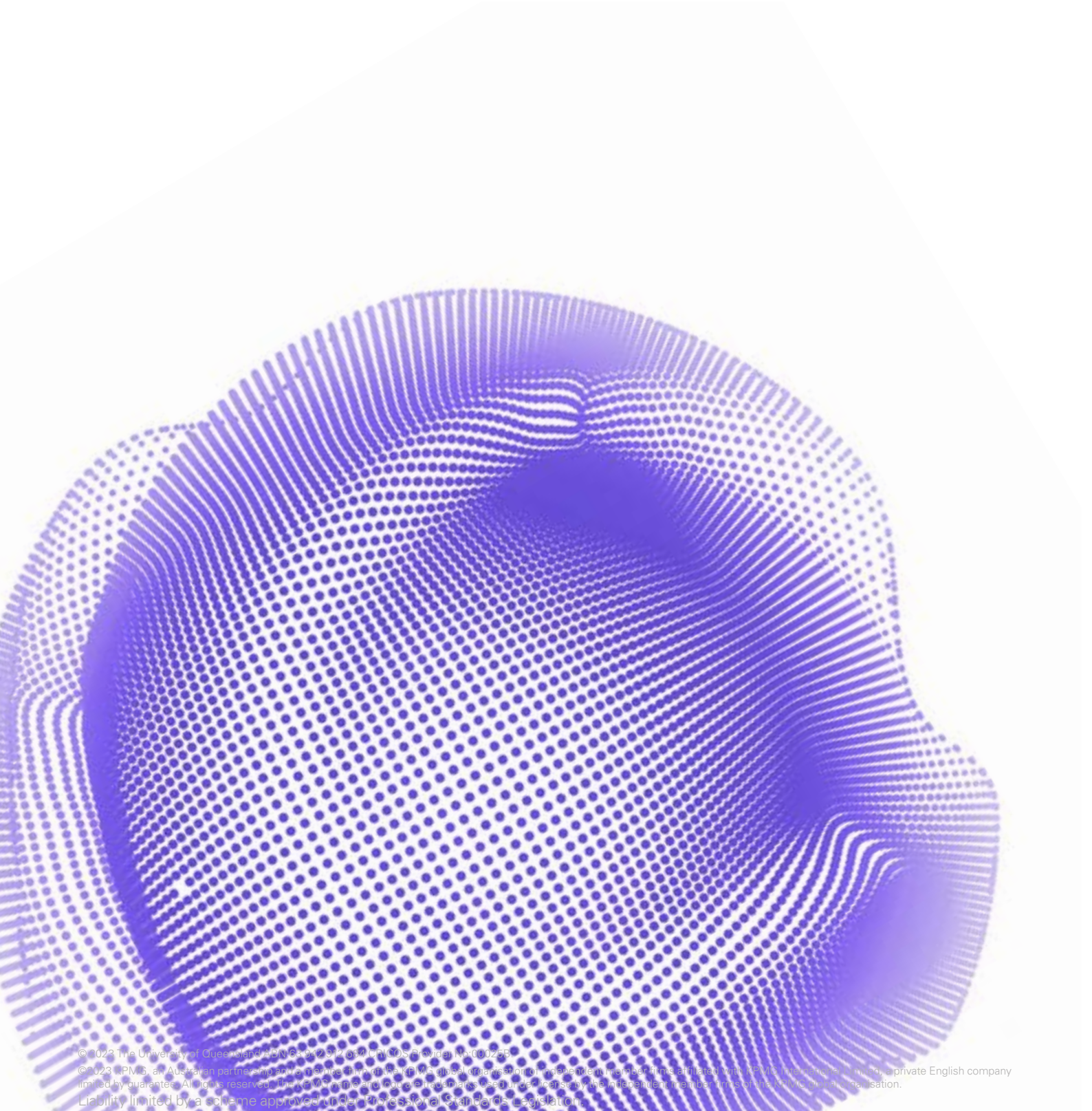
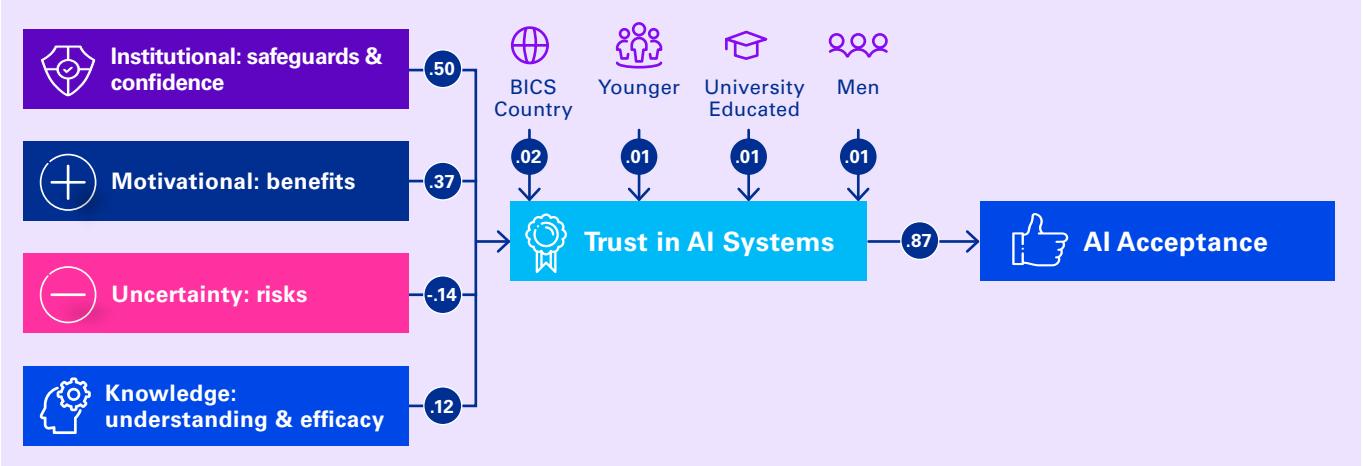


Figure 44: Model of the key drivers of trust and acceptance of AI systems



Institutional drivers include:

- Safeguards: the belief that current laws, rules and governance are sufficient to ensure AI use is safe
- Confidence in government and technology/commercial organisations to develop, use and govern AI

Motivational drivers are the perceived benefits of AI: the extent to which people expect a range of benefits to arise from the use of AI systems

Uncertainty drivers are the perceived risks of AI: the extent to which people are concerned about a range of risks related to the use of AI systems

Knowledge drivers include:

- Subjective knowledge: the extent to which people feel they understand AI, and when and where it is used
- Tech efficacy: people's assessment of their ability to use digital technologies and online services

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

The extent to which people accept and approve of AI systems

Country and demographic differences had a small impact on trust:

- People in the BICS countries are more trusting of AI
- Younger generations are more trusting of AI
- University educated people are more trusting of AI
- Men are more trusting of AI

How to read the model

When reading the model, follow the arrows from left to right. The left boxes show the four drivers of trust, with notes explaining each driver in the boxes below the model. The values on the arrows indicate the relative importance of each driver in influencing trust and acceptance: the larger the number, the stronger the effect. The positive values for institutional safeguards and confidence, benefits, and knowledge, indicate that when these drivers increase, so does trust. The negative value for uncertainty indicates that when perceived risks increase, trust decreases.

The model is based on all data (across countries and AI applications). All relationships shown are significant ($p<0.005$).

Institutional safeguards and confidence are the strongest drivers of trust

The institutional pathway is grounded in the view that people often defer to, and expect, authoritative sources and institutional processes to provide assurance of the safety and reliability of new technologies and practices.

As shown in the model, people are more trusting of AI systems when they believe current regulations and laws are sufficient to make AI use safe, and have confidence in the government, technology and commercial organisations to develop, use and govern AI. This institutional pathway is the strongest predictor of trust (B=0.50) and significantly more important than the other drivers.

Given our results show most people are unconvinced that current governance and regulations are adequate to protect people from problems associated with AI, a critical first step towards strengthening trust and the trustworthiness of AI is ensuring it is governed by an appropriate regulatory and legal framework.

The perceived benefits of AI motivate trust

The motivational pathway to trust is grounded in evidence that the more people perceive benefits from using AI, the more motivated they will be to trust AI. This motivational pathway is the second strongest predictor of trust (B=0.37) and is a stronger driver of trust than the perceived risks of AI. This helps explain why people are willing to use technologies that provide immediate benefits (e.g. convenience), despite concerns about potential risks.

This finding highlights the importance of designing and using AI systems in a way that delivers demonstrable benefits to people, in order for these systems to be trusted and accepted.

The perceived risks of AI create uncertainty and reduce trust

The uncertainty pathway is based on evidence that it is more difficult to trust in uncertain and risky contexts. The model shows that the more concerned people are about the perceived risks of AI use, the less likely they are to trust in AI systems (B=-0.14). This includes both technical risks associated with AI use (e.g. cybersecurity and privacy risks, risk of inaccurate or biased outcomes) and broader societal risks (e.g. manipulation, deskilling and job loss). This is the third strongest driver of trust.

This finding underscores the importance of ongoing action to mitigate the risks associated with AI at multiple levels, as well as communicating these risk mitigation strategies to reassure people.

Understanding AI and how to use technology influences trust

The knowledge pathway is based on evidence that knowledge and understanding enhance trust in technology.

The model shows that people are more likely to trust and accept AI when they feel they understand when and how AI is used and are sufficiently skilled to use digital technologies (B=0.12). This knowledge pathway is the fourth driver of trust and highlights the importance of supporting people's technological and digital literacy and skills.

Country and demographic factors have a smaller impact on trust

After taking into account the institutional, motivational, uncertainty and knowledge drivers of trust in the model, we found four other factors have a small but meaningful influence on trust.

People in the BICS countries (B=0.02), as well as younger generations (B=0.01), the university educated (B=0.01), and men (B=0.01) are more trusting of AI.

Collectively, the drivers in our model account for 84% of the variance in trust.

Qualitative insights on what enhances trust in AI systems

To supplement our understanding of the key drivers of trust, we asked people: *What would enhance your trust in this AI system?* People responded to this question in relation to one of the AI applications (i.e. Healthcare AI, Security AI, Recommender AI, Human Resources AI, or AI systems in general).

We conducted a thematic analysis to identify the key themes, supplemented by a word cloud to identify the most frequently mentioned terms (see Figure 45).

This qualitative data reinforces the importance of the following pathways for trust:

- *Knowledge and understanding of AI* including learning how AI works, and skills, experience and familiarity in using it.
- *Risk mitigation* by protecting the *privacy and security* of personal data and information, ensuring the *accuracy and reliability* of AI output and performance, and *retaining human control* and oversight.
- *Institutional safeguards* particularly 'regulations', 'laws', 'rules', 'appropriate governance' and 'monitoring' mechanisms that 'guarantee' responsible use and prevent problems.
- Showing *benefits* to people and society, including improving the quality, effectiveness, efficiency and ease with which work is done, reducing costs, and 'improving the lives of people' and 'betterment' of society, including taking a people-centred approach.
- *Demonstrating trustworthiness* by being 'transparent' about how AI is being used and implemented, including how data and personal information is used and how outcomes are reached, and proving the 'trustworthiness' and 'safety' of AI systems 'over time'.

The qualitative analysis also revealed that a proportion of people feel 'nothing' can be done to enhance their trust, with some perceiving the technology as too risky.

Figure 45: Word cloud of responses on what enhances trust in AI systems

TOPIC EIGHT

How have trust and attitudes towards AI changed over time?

In 2020, we conducted a survey examining trust in and attitudes towards AI in five countries: Australia, Canada, Germany, the UK, and USA. To understand how attitudes towards AI have shifted over time, we examined changes in constructs that were measured in the same way in these five countries from 2020 to 2022.[40]

Trust in AI systems has increased over time

The willingness to trust in AI systems and the perceived trustworthiness of these systems meaningfully increased from 2020 to 2022. This increase occurred for AI systems in general (trust M= 3.7 vs 4.0; trustworthiness M= 4.1 vs 4.7) and Human Resources AI (M trust M= 3.5 vs 3.8; trustworthiness M= 4.0 vs 4.4).

The largest increase occurred for the perceived trustworthiness of AI systems in general, which rose from 35% agreeing AI systems are trustworthy in 2020 to 56% in 2022. Increases occurred in all countries, with the largest increase in Germany. Trust in AI in general increased from 28% to 36%.

Trust and trustworthiness of Healthcare AI did not increase meaningfully in the whole sample but did so in Germany. For example, the proportion of Germans who perceived Healthcare AI to be trustworthy increased from 37% to 58% (M=4.2 vs 4.7).

This increase in trust most likely reflects increased use, understanding, and familiarity with AI, as documented in the next section.

Awareness of AI and use of AI in common applications has increased

Awareness and objective understanding of AI has increased over time. More people report using common applications underpinned by AI, such as social media and navigation apps (56% to 67% use), and more people are aware that these applications use AI (46% vs 56% aware). The proportion of people who had read or heard about AI also increased (62% to 78%). In particular, people are more aware of AI use in social media, text recognition, facial recognition, virtual assistants, and traffic navigation applications. While awareness and use increased in all countries, the greatest shift occurred in Germany.

Despite this increased awareness of AI, people's subjective understanding of how and when AI is used did not increase over time.

More people in the UK and USA believe AI regulation is needed but there has been no increase in the perceived adequacy of current regulation and safeguards

More people in the UK (66% vs 80%) and USA (57% vs 66%) report that AI regulation is needed. There was no change in the other countries.

There was no change in perceptions of the adequacy of current safeguards in the whole sample. However, there was one country- and application-specific increase: people in Germany perceive more current safeguards in Healthcare AI in 2022 (41%, M=4.2) compared to 2020 (33%, M=3.7). People in Germany also perceived less uncertainty around the use of AI, with no meaningful changes in other countries. These unique findings for Germany may reflect the EU's progress in regulating data and AI (e.g. the EU AI Act).

Confidence in entities to develop, use and govern AI has not changed

There were no meaningful changes in people's confidence in entities to develop, use or regulate and govern AI in any country.

Fewer people in the UK and Canada believe AI will create more jobs than it will eliminate

There was no change in the belief that AI will create more jobs than it will eliminate in the whole sample, with 20% agreeing with this sentiment in 2022, compared to 22% in 2020. However, fewer people in the UK (21%, M=3.4 vs 15%, 3.0) and Canada (23%, M=3.5 vs 16%, M=3.2) believe AI will create more jobs than it will eliminate.

Similarities in the drivers of trust across time and samples

In our 2020 report, we presented a model that included some similar drivers to the model reported in the previous section (e.g. institutional safeguards, familiarity with AI, AI uncertainty). These models differ in terms of: a) the inclusion of additional and refined measures in 2022 to account for the global nature of the survey and to better represent the theoretical pathways to trust; b) the use of a more advanced statistical modelling technique (structural equation modelling instead of path analysis); and c) the larger and more diverse sampling (17 countries across all global regions vs 5 western countries).

Despite these differences, we see several similarities in the findings between models. Institutional safeguards is the strongest predictor of trust in both models. Familiarity and knowledge of AI are also significant predictors of trust with similar impacts.

We also see similarity in the predictors of AI uncertainty (2020) and perceived risks (2022) with both having a significant negative impact on trust. Trust is a strong predictor of acceptance in both models.

In addition, we re-ran the model presented in the previous section using the sample of respondents from the five western countries surveyed in both 2020 and 2022 (USA, UK, Canada, Australia and Germany). Given the different and more positive pattern of response for the BICS countries, we wanted to determine if the model and drivers of trust are consistent for these five western countries.

Results indicate that the model holds for these five countries. While the strength of relationships vary slightly, the four key pathways and drivers of trust remain the same. Specifically, the institutional path remained the strongest, followed by the motivational, uncertainty and knowledge paths. Trust remains a strong predictor of acceptance.

Conclusion and implications

Together, the findings of this global survey provide a clear overview of the current and future challenges to trust and acceptance of AI systems, as well as opportunities for overcoming these challenges. The findings inform four pathways for how AI can be integrated into society in a way that is responsible and engenders trust.

Trust and attitudes towards AI vary across countries: western countries, together with Japan, South Korea, and Israel, generally have lower trust and less positive attitudes than people in the emerging economies.

Across our findings, we see a general pattern of lower trust, greater ambivalence, and less positive views towards AI in western countries, Japan, South Korea and Israel, compared to people in the BICS countries, and to some extent, Singapore.

Our findings suggest that the more positive attitudes in the BICS countries do not reflect blind optimism or lack of awareness of the potential risks of AI use. Rather, we see some evidence of the opposite, with people in Brazil and South Africa (together with South Korea) rating the risks of AI higher than other countries, and Indians and South Africans more likely to believe AI risks will impact people in their country.

Rather, our analysis suggests the varying levels in trust and acceptance across countries largely reflect three key factors:

- Differences in the perceived *benefits* of AI and the extent to which they outweigh potential risks: people in western countries and Japan are generally less convinced of the benefits of AI, and together with South Korea and Israel, less likely to believe the benefits of AI outweigh the risks, compared to people in the BICS countries and Singapore.
- Perceptions of *institutional safeguards:* there are differences across countries in the perceived adequacy of safeguards and regulations to make AI use safe, and confidence in the institutions responsible for this. Fewer people in western countries, Japan, South Korea and Israel view current laws and regulations for safeguarding AI as sufficient, and report less confidence in companies to develop, use and govern AI, compared to people in Brazil, India, China and Singapore.
- Familiarity and *understanding of AI:* people in western countries generally report less use of AI at work, and lower use and knowledge of AI in common applications, compared to people in the BICS countries and Singapore.

The key commonality across the BICS countries, despite differences in economic strength, institutional arrangements, and technological advancement, is the emerging nature of their economies. This may encourage a growth mindset in relation to the acceptance of technology, as it provides a means to accelerate economic progress and prosperity.

An implication of the greater trust and acceptance of AI in the BICS countries (and to some extent Singapore), combined with the emerging nature of their economies, is that these countries may be uniquely positioned to accelerate innovation and technological advantage and offer a more supportive environment to attract businesses wishing to invest in AI development and innovation, supporting a competitive advantage. Over time, this may contribute to a disruption of traditional economic hegemonies.

However, although people in the BICS countries are more likely to perceive current AI regulations and laws as adequate, it is noteworthy that on international rankings these countries rank low on governance and regulation to ensure the ethical and responsible use of AI.[41] In contrast, the European Union (EU) and Canada are viewed as leaders in AI and data governance and ethics. The EU's AI Act will set limits and conditions on the use of AI systems based on a risk classification and restrict the types of AI products and services that can be developed and sold in the EU, which are likely to influence AI development and governance practices in other countries.

Sustained competitive advantage requires trust and acceptance to be preserved. We turn now to discuss the four pathways identified by our modelling for strengthening the responsible use of AI to secure and sustain public trust: the institutional, motivational, uncertainty reduction and knowledge pathways.

Key pathways for strengthening the responsible and trusted use of AI

The Institutional Pathway: institutional safeguards and confidence in entities to use and govern AI is the strongest driver of trust

- Our findings highlight the important role institutions and institutional safeguards play in laying a foundation for trust and reassuring people that trust in AI is warranted. Yet, while the large majority of people expect AI to be regulated, many view the current regulatory and legal framework as insufficient to make AI use safe and protect people from the risks.
- The public clearly want appropriate regulation that is fit-for-purpose to manage the risks and uncertainties associated with AI. Our findings show that people broadly support multiple forms of regulation, including co-regulation with industry, but expect some form of external, independent oversight, such as regulation by government or a dedicated independent AI regulator.
- A key challenge here is that a third of people report low confidence in government, technology and commercial organisations to develop, use and regulate AI in society's best interest. This is problematic given regulation is a function of government, and governments currently rely heavily on the self-regulation of AI by business. It is also problematic given the increasing prevalence and scale with which governments and businesses adopt AI.
- Our comparative analysis over time suggests that there has been no perceived improvement in the adequacy of current institutional safeguards around AI, nor any increase in confidence in governments and commercial organisations around their AI activities, in the five western countries surveyed in 2020 and 2022. This suggests that the institutional frameworks supporting AI are lagging and failing to keep pace with community expectations. The implementation of the EU's AI Act provides an opportunity to examine how dedicated regulation affects trust and acceptance over time.
- Given the public has the most confidence in universities and research organisations to develop, use, and govern AI systems, a potential solution is for business and government to partner with these organisations around AI initiatives.
- As general trust in government and commercial organisations is strongly associated with confidence in these entities to use and govern AI, strengthening trust generally in these institutions is an important foundation for supporting trust in AI.

The Uncertainty Reduction Pathway: similarities in the perceived risks across countries and strong consensus on expectations for trustworthy AI support a global approach to risk mitigation

- A salient finding that was reinforced in both our quantitative and qualitative data is that people view the risks of AI in a comparable way across countries. From a business point of view, this suggests that businesses operating in multiple markets can anticipate a common set of risks across these markets and use similar strategies to manage and mitigate these risks (while noting there may be some cultural nuances in expectations of mitigation strategies).
- More broadly, this insight supports the merit and necessity of global collaborative approaches to AI governance and international standard setting, such as by the JTC1, IEEE, the OECD and the Global Partnership on AI, to help mitigate AI risks and support responsible use. It also underscores the importance of striving for consistency in AI regulatory and legislative frameworks across countries and markets.
- We found strong global public consensus on the principles and related practices organisations deploying AI systems are expected to uphold in order to be trusted. Organisations can directly build trust in their use of AI systems by developing capabilities and practices for upholding these principles, and investing in assurance mechanisms that demonstrate and support the responsible deployment of AI systems, such as regularly monitoring system accuracy and reliability, implementing AI codes of conduct, independent AI ethics reviews and certifications, and meeting international standards. However, the 2022 IBM Global AI Adoption Index indicates that a majority of businesses deploying AI have not taken steps to ensure their AI use is trustworthy.[42]
- A key trust-enhancing practice is the retention of human involvement and oversight in decisions that impact people. For example, we find that most people are comfortable with AI-human collaboration in managerial decision-making, and prefer AI involvement to sole human decision-making, with the caveat that humans retain equal or greater input. We also find strong universal endorsement of the principle and practice of human oversight and control of AI systems.
- This carries an important implication for technology leaders striving for the highest levels of automation: while full automation may maximise efficiency and cost reduction, it can undermine trust and acceptance. Balance is required.
- Our findings further reinforce the critical importance of cybersecurity in the digital age and protecting people's data and privacy from cybercrime. Combined with our finding that people are generally more willing to rely on the output of AI systems than share information with these systems, this underscores an inherent tension in the trustworthiness of AI systems: larger datasets typically enable greater model accuracy and robustness, but augment the risks associated with a data breach.

The Motivational Pathway: demonstrating the benefits of AI to people and society motivates trust

- Our modelling revealed that an important pathway for strengthening and preserving trust comes from demonstrating the tangible, beneficial impact of AI for people and society. This highlights the importance of human-centred AI design and having a clear beneficial purpose at the outset of AI projects, as well as co-designing AI-enabled services and products with key stakeholders and end-users.
- The pattern in our data suggests people often perceive more benefits of AI use for the organisations deploying it rather than for people or society more broadly. However, our data indicates people trust AI more when it has a clear beneficial purpose to people (e.g. Healthcare AI to enhance diagnosis and treatment) rather than a process or organisational benefit (e.g. to increase efficiency in Human Resource processes).
- The fact that one in two people do not see the benefits of AI as outweighing the risks highlights that an integrated approach is required: augmenting benefits to people while proactively mitigating risks. It further highlights that a communication exercise may be required to ensure people are aware of the benefits of AI-enabled services and products, while guarding against 'benefit selling'.
- Collectively, the findings underscore that trust in AI systems is contextual and dependent on the perceived benefits, risks, norms, and safeguards related to the specific AI use case.

These four pathways are each important for the responsible stewardship of AI into society and provide complementary ways to build and maintain public trust in AI systems.

Given the rapid and widespread deployment of AI, it will be important to regularly re-examine public trust and expectations of AI systems as they evolve over time, to ensure AI use is aligned with and meeting changing societal expectations.

The Knowledge Pathway: people trust AI systems more when they feel they understand AI and are skilled in using digital technologies

- A key finding is that the public generally has low understanding of AI and its use in everyday life. However, a large majority of the community are interested in learning more about AI and report that supporting people to understand AI is important for their trust. This last insight is further supported by our path model, which identified understanding of AI and efficacy in using online and digital technologies is an important driver of trust and acceptance of AI.
- Our findings highlight that many people use applications without any awareness that they involve AI, particularly older people and those without a university education. This raises the question of whether it is important to be transparent in disclosing the use of AI (similar to disclosing the use of cookies), to enable people to make more informed decisions about use, benefits and risks. For example, should people be made aware when they are interacting with a chatbot as opposed to a human, and that social media and recommender systems use AI in ways that may potentially be manipulative?
- Organisations also need to consider that different cohorts in the workplace and community have different understandings and views about AI, with younger people, the university educated, and managers more aware, knowledgeable about and accepting of AI. In contrast, other cohorts are likely to need more reassurance and evidence of the trustworthiness of these technologies. As such, a one-size-fits-all approach is unlikely to work.
- While university education has a beneficial impact on trust, our findings raise interesting questions about the role of AI public literacy programs and trust. As an example, Finland has a high level of public education about AI, and our results indicate that Finns are more aware of and knowledgeable about AI, more likely to use it at work and have more positive than negative feelings towards AI than their western counterparts, yet they report low trust and moderate acceptance. This suggests that being better informed does not equate simplistically to greater trust and acceptance. Education should play a role in informing people of potential risks and benefits, as well as methods for safe and responsible use, and should, over time, help to achieve a balance that enables trusted adoption.
- Taken together, these findings suggest close collaboration is required between government, universities and business to uplift public and consumer literacy and understanding of data and technology.

Appendix 1: Method and statistical notes

In this section, we explain our research methodology and statistical approach.

Survey procedure

The research was approved by and adhered to the Guidelines of the ethical review process of The University of Queensland and the National Statement on Ethical Conduct in Human Research.

The survey was divided into five sections with questions in each section focused on the respondent's: 1) use and understanding of AI; 2) attitudes towards AI systems (including trust, acceptance, risks, benefits, impacts and emotions); 3) attitudes towards AI governance and management; 4) attitudes to the use of AI at work; and 5) education and individual differences (e.g. technology efficacy, disposition to value privacy). At the end of the survey, respondents were asked open-ended questions to understand what would enhance trust in AI and what concerns them most about the use of AI.

After completing the first section on use and understanding of AI, participants read the OECD's definition of AI. Questions in sections two and three of the survey, as well as the open-ended questions in section 5, referred to one of the AI use cases described on page 10, or referred to 'AI systems in general'. Respondents were randomly allocated to one of these five use cases, providing equivalent numbers of responses across use cases.

Before answering this subset of questions, respondents were provided with a brief description of the specific AI system, including what it is used for, what it does and how it works (see full descriptions on page 10). The research team developed these descriptions based on a range of in use systems and input from domain experts working in healthcare, human resources, security, and recommender systems, respectively.

Survey measures, piloting and translations

Where possible, we used or adapted existing validated measures from academic research (e.g., Lu et al., (2019), McKnight et al., 2002; Pavlou et al., 2007; Zhang & Moffat, 2015) or from previous public attitude surveys (e.g. Eurobarometer, 2017; Ipsos, 2017; Zhang and Dafoe, 2019).

Trust in each specific AI application was measured using a reliable 8 item scale from Gillespie et al. (2021). Example items are: How willing are you to… rely on information provided by the AI system; depend on decisions made by the AI system; share relevant information about yourself to enable the AI system to perform a service or task for you; allow your data to be used by the AI system. Perceived trustworthiness was measured using a 14 item measure assessing positive expectations towards the AI system, adapted from McKnight et al (2002). Example items include: I believe [specific AI application] would: produce output that is accurate; have a positive impact on most people; be safe and secure to use.

Acceptance was measured using a reliable 3 item scale adapted from Zhang & Moffat (2015). Items included: To what extent do you accept the use of the AI application, approve of the use of the AI application, embrace the use of AI application.

The psychometric properties of all multi-item constructs were assessed to examine reliability and dimensionality. All were found to be reliable with Chronbach alphas ranging from 0.78 (uncertainty avoidance) to 0.93 (basic psychological needs). All constructs were unidimensional except for AI system trustworthiness, which is conceptualised and reported as three dimensions: 1) Ability (system perceived to perform accurately, reliably, and as intended); 2) Humanity (system perceived to create benefits, provide a helpful service, and have a positive impact); and 3) Integrity (system perceived to uphold ethical principles and data privacy rights, and operate in a fair and safe way).

We extensively piloted and refined the survey before full launch. To ensure survey equivalence across countries, we conducted translation and back-translation of the English version of the survey into the native languages dominant in each country, using separate professional translators. Respondents could also opt to complete the survey in English if preferred. The following languages were offered in their respective countries: French (France and Canada), German, Finnish, Estonian, Simplified Chinese, Japanese, Portuguese, Hebrew, Dutch, and Korean.

Reporting percentages and rounding

Most survey measures used either a 5 or 7-point Likert scale. When reporting percentages we adopted the following cut-off values for reporting low, moderate and high values, unless otherwise reported.

- 5-point scales:
  - Low = (mean) values ranging from 1.0 to 2.49
  - Moderate = (mean) values ranging from 2.50 to 3.50
  - High = (mean) values ranging from 3.51 to 5.0
- 7-point scales:
  - Low = (mean) values ranging from 1.0 to 3.49
  - Moderate = (mean) values ranging from 3.50 to 4.50
  - High = (mean) values ranging from 4.51 to 7.0

When percentages did not add up to 100% due to rounding, we distributed the remaining value based on decreasing order of the values' decimal part, as per the Largest Remainder Method.

Reporting differences between countries, applications, and people

Our reporting of between-country, between-application, between-people and within-person differences was informed by statistical analyses and adhered to well established benchmarks for interpreting between- and within-subject effect sizes (see Cohen, 1988; Lakens, 2013).

We used one-way analysis of variance (ANOVA) to examine differences between countries, AI applications and people (e.g. generational differences). We took several steps to ensure the responsible reporting of only meaningful differences in the data. First, we adopted a stringent cut-off of $p<0.01$ to interpret statistical significance. Where there are statistically significant differences between groups ($p<0.01$), we examined the omega-squared effect size to determine the magnitude of difference between the groups. Differences with an effect size less than 0.01 were deemed practically insignificant and are not reported. Meaningful patterns of between country findings that exceeded the 0.01 effect size are reported[43].

We performed paired-sample t-tests to examine within-person differences (for instance, the variability in one's willingness to rely on versus share information with an AI system). We used a measure of effect size to determine the magnitude of statistically significant effects. Specifically, we used *Hedges' g* with a cut-off of 0.30 to indicate a robust and practically meaningful difference.[44]

Appendix 2: Country samples

In this section, we describe the demographic profile of each country sample.

The demographic profile of each country sample was nationally representative on age, gender and regional location, within a 5% margin of error, based on official national statistics within each country.

Across countries, the gender balance was 51% women, 49% men and 1% other genders for all countries, with Estonia having the highest representation of women (53%) and India the lowest (47%).

The mean age across countries was 44 years and ranged from a mean of 36 years (India) to 51 years (Japan). In three countries (India, Israel and Estonia), the sample of respondents was slightly younger than the respective population average due to under-representation in the 55+ age bracket in these countries: Estonia (55+ expected: 34%, achieved: 14%), India (55+ expected: 18%, achieved: 10%) and Israel (55+ expected: 31%, achieved: 8%).

Country samples varied on education. Samples from emerging economies (Brazil, India, China, and South Africa) represented considerably more university educated people than their respective general populations (using OECD 2021 education data as a comparison[45]). A higher representation of educated people is common in survey research from the BICS countries. For instance, Edelman (2022) and Ipsos (2022) both note that online samples in Brazil, India, China and South Africa are more educated, affluent, and urban than the general population.[46]

However, our analyses show that the more positive responses reported by people in BICS countries is not due to the higher education level of these samples. We examined differences between BICS and non-BICS countries on key indicators when controlling for the effects of education and age, using analysis of covariance (ANCOVA) tests.

We found significant differences between country groupings on all key indicators, even when controlling for education and age, with higher levels of trust and more positive attitudes for people in BICS countries.

Furthermore, people *without* a university education in the BICS countries demonstrate higher mean values on key constructs than those without a university education in non-BICS countries. For example, people without a university education in BICS countries have more trust in AI (M=4.8 vs 3.8) and perceive more benefits of AI (M=3.8 vs 3.2), than those without a university education in all other countries. There are also no educational differences in trust in AI for China or South Africa

Table A1:

The demographic profile for each country sample

Country	% Gender			Age (yrs)	% Generation				% Education				
	W	M	O	Mean	Z	M	X	BB	<SS	SS	Qu	UG	PG
Australia	51	48	1	47	14	30	23	33	4	26	28	31	11
Brazil	52	47	1	40	19	37	31	13	4	27	17	38	14
Canada	51	48	1	48	12	28	25	35	5	24	31	31	9
China	48	52	0	43	16	37	25	22	1	7	15	70	7
Estonia	53	46	1	38	23	41	24	12	8	29	24	24	15
Finland	51	48	1	45	14	30	29	27	11	24	29	22	14
France	50	49	1	46	13	28	30	29	10	23	24	26	17
Germany	51	49	0	46	11	30	31	28	5	14	59	13	9
India	47	53	0	36	26	43	25	6	1	6	4	40	49
Israel	47	52	1	38	20	44	26	10	1	25	25	34	15
Japan	50	50	0	51	9	20	33	38	2	29	17	47	5
Netherlands	50	49	1	46	14	30	28	28	10	28	26	23	13
Singapore	51	49	0	44	13	32	34	21	1	12	31	44	12
South Africa	51	48	1	39	19	43	23	15	1	31	26	35	7
South Korea	52	48	0	45	12	25	41	22	1	19	3	65	12
UK	50	49	1	46	13	33	26	28	2	27	29	30	12
USA	51	48	1	45	15	28	26	31	7	30	26	24	13

Notes: Gender: W = women, M = men, O = Non-binary or other gender identity; Generation: Z = Generation Z, M = Millennials, X = Generation X, BB = Baby Boomers and older generations; Education: <SS = Lower secondary school or less, SS = Upper secondary school, Qual = Vocational or trade qualification; UG = Undergraduate degree; PG = Postgraduate degree

Appendix 3: Key indicators for each country

In this section, we report the key indicators for each country.

Country	Trust	Twthy	Accept	Benefits	Risks	Benefit -Risk	Current Safeguards	Subjective Knowledge
Australia	3.9/7	4.5/7	2.7/5	3.0/5	3.3/5	4.1/7	3.8/7	2.4/5
Brazil	4.7	5.4	3.6	3.8	3.4	5.0	4.5	2.5
Canada	3.9	4.5	2.7	3.0	3.2	4.1	3.6	2.4
China	5.1	5.6	3.8	3.9	3.2	5.4	5.4	3.5
Estonia	3.7	4.5	3.0	3.4	3.1	4.1	4.0	2.2
Finland	3.4	4.4	2.8	3.1	3.3	4.2	3.7	2.7
France	3.7	4.6	2.8	3.2	3.3	4.1	3.7	2.2
Germany	4.0	4.6	3.1	3.4	3.0	4.2	4.1	2.5
India	5.2	5.8	3.8	4.0	3.2	5.1	5.5	3.4
Israel	4.1	4.8	3.1	3.7	3.2	4.4	4.0	2.7
Japan	3.7	4.5	2.7	3.1	3.1	4.3	3.3	2.1
Netherlands	3.7	4.5	2.6	3.1	3.1	4.0	3.8	2.1
Singapore	4.3	5.0	3.1	3.4	3.4	4.7	4.6	2.8
South Africa	4.7	5.3	3.4	3.8	3.5	4.7	4.3	2.6
South Korea	4.0	4.8	3.1	3.7	3.5	4.4	3.4	3.1
UK	4.0	4.6	2.7	3.0	3.2	4.1	3.8	2.2
USA	4.0	4.7	2.7	3.1	3.3	4.1	3.7	2.5

Notes: Trust = Trust in AI systems, Twthy = Perceived trustworthiness of AI systems, Accept = Acceptance of AI systems, Benefits = Perceived benefits of AI systems, Risks = Perceived risks of AI systems, Benefit-Risk = Perception that benefits of AI systems outweigh the risks, Current Safeguards = Perceived adequacy of current laws and regulations governing AI, Subjective Knowledge = Self-reported knowledge of AI.

Endnotes

1. McKinsey (2022). *The state of AI in 2022 – and a half decade in review*.
2. BRICS is the acronym used to describe the five major emerging economies of Brazil, Russia, India, China, and South Africa (see https://en.wikipedia.org/wiki/BRICS). We did not include Russia in our sampling due to the current invasion in Ukraine, and therefore use the acronym BICS in this report.
3. European Commission (2019). *Ethical guidelines for trustworthy AI*
4. Eurobarometer. (2017). *Attitudes towards the impact of digitisation and automation on daily life* (Report no. 460). Selwyn, N., Gallo Cordoba, B., Andrejevic, M, & Campbell, L. (2020) *AI for Social Good? Australian public attitudes toward AI and society*. Zhang, B., & Dafoe, A. (2019). *Artificial intelligence: American attitudes and trends*.
5. Eurobarometer (2019). *Europeans and Artificial Intelligence*. The European Consumer Organization (2020). *Artificial Intelligence: what consumers say. Findings and policy recommendations of a multi-country survey on AI*.
6. See the Government AI Readiness Index 2022. See also the Responsible AI Sub-Index of the Government AI Readiness Index 2020.
7. For a full description of the 2020 Trust in AI survey, see Gillespie, N., Lockey, S., & Curtis, C. (2021). *Trust in Artificial Intelligence: A Five Country Study*.
8. Our finding that trust influences AI acceptance supports prior research. For example, see the following studies:

   Choung, H., David, P., & Ross, A. (2022). *Trust in AI and its role in the acceptance of AI technologies*. *International Journal of Human–Computer Interaction,* 1-13.

   Fan, W., Liu, J., Zhu, S., & Pardalos, P. M. (2020). *Investigating the impacting factors for the healthcare professionals to adopt artificial intelligence-based medical diagnosis support system*. *Annals of Operations Research, 294*(1), 567-592.

   Gillespie, N., Lockey, S., & Curtis, C. (2021). *Trust in artificial intelligence: A five country study*. The University of Queensland and KPMG Australia.

   Zhang, S., Meng, Z., Chen, B., Yang, X., & Zhao, X. (2021). *Motivation, social emotion, and the acceptance of artificial intelligence virtual assistants: Trust-based mediating effects*. *Frontiers in Psychology,* (12) 3441.
9. Gillespie, N., Curtis, C., Bianchi, R., Akbari, A., & Fentener van Vlissingen, R. (2020). *Achieving Trustworthy AI: A Model for Trustworthy Artificial Intelligence*. KPMG and The University of Queensland Report. Tabassi, E. (2023), *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST Trustworthy and Responsible AI, National Institute of Standards and Technology, Gaithersburg, MD, [online]. *OECD AI Policy Observatory Tools for Implementing Trustworthy AI*
10. Morning Consult (2022) *While Tech Fawns Over AI, Consumers Need More Convincing* Loureiro, S. M. C., Guerreiro, J., & Tussyadiah, I. (2021). *Artificial intelligence in business: State of the art and future research agenda*. *Journal of Business Research, 129,* 911-926.
11. OECD (2019) *Artificial Intelligence in Society*.
12. Multiple international and pan governmental organisations, including the OECD, The European Commission, and the G7 Innovation Ministers, note the importance of trust in AI and developing 'trustworthy' AI systems, to support continual AI adoption. This is also recognised in the AI roadmaps and strategic plans of the five countries examined in this report (see for example: the *UK AI Roadmap* and the *US National Artificial Intelligence Research and Strategic Development Plan)*
13. Prior public attitude surveys have examined general acceptance and support of AI. For example, see: The European Consumer Organization (2020). *Artificial Intelligence: what consumers say. Findings and policy recommendations of a multi-country survey on AI*. Eurobarometer (2019). *Europeans and Artificial Intelligence*. Ipsos (2022) *Global Opinions and Expectations about Artificial Intelligence*. Morning Consult (2022) *While Tech Fawns Over AI, Consumers Need More Convincing*. Pew Research Centre. (2020). *Science and Scientists Held in High Esteem Across Global Publics: Yet there is ambivalence in many publics over developments in AI, workplace automation, food science*. Zhang, B., & Dafoe, A. (2019). *Artificial intelligence: American attitudes and trends*.
14. Data was collected from research panels sourced by Qualtrics, a global leader in survey panel provision.
15. The global regions of 1) East Asia, 2) Eastern Europe, 3) Latin America and the Caribbean, 4) Middle East and North Africa, 5) North America, 6) Pacific, 7) South and Central Asia, 8) Sub-Saharan Africa, and 9) Western Europe.
16. We focused primarily on the 2021 *Government AI Readiness Index*. This index ranks and provides a total score for 160 countries on AI readiness across three pillars: Government (e.g., existence of a national AI strategy, cyber-security), Technology (e.g. number of AI unicorns, R&D spending), and Data and Infrastructure (e.g. telecommunications infrastructure, households with internet access). We supplemented this with data from the *2021 Stanford AI Index*, which examines country-level private investment in AI and acceleration in AI activity over time to enable identification of countries that are rapidly emerging in AI in regions that historically lacked AI capacity and investment (e.g. South Africa, Brazil, and India). The countries selected maintained their ranking at or near the top for their region on the 2022 Government AI Readiness Index.
17. The Responsible AI Index ranks countries according to the readiness of their governments to use AI in a responsible way according to four pillars that correspond with the OECD Principles on AI: Inclusivity, Accountability, Transparency and Privacy. It is a sub-index of the *Government AI Readiness Index* produced by Oxford Insights.
18. Occupational groupings sourced from OECD International Standard Classifications of Occupations
19. See for example Gillespie, N., Lockey, S., & Curtis, C. (2021). *Trust in Artificial Intelligence: A Five Country Study*.
20. This definition aligns with dominant interdisciplinary definitions of trust (e.g. Mayer et al., 1995; Rousseau et al., 1998), including trust in technological systems (see McKnight et al., 2002, 2011).
21. The *2021 Stanford AI Index* reports accelerated use of AI in the major emerging economies, as well as the increasing economic importance of AI in these countries. Our pattern of findings aligns with two prior surveys reporting more positive sentiments towards AI in emerging economies (see *Ipsos, 2022; Pew Research Center, 2020*).
22. To determine if people experience ambivalent emotions towards AI, we categorized positive (e.g. optimistic, excited, and relaxed) and negative (e.g. fearful, worried, and outraged) emotions into high (ratings of 'moderately' and above, i.e. 3-5 on the five point scale) and low (ratings of 'slightly' and below, i.e. 1-2 on the five point scale). Participants experiencing one or more of the positive emotions 'moderately' or above, were classified high for positive emotions, and similarly for negative emotions.

Endnotes

23. 27% of respondents opted to answer this open-ended question.
24. Stanford University (2021). *Artificial Intelligence Index Report 2021*.
25. We asked respondents to rate their confidence in twelve entities. Given similarities in the data across some of these entities, for simplicity we amalgamated responses to the following six entities into three entities: 'The federal government' and 'My state/province government' to form 'Federal/State/Provincial Government'; 'Independent regulatory bodies funded by the government' with 'Existing agencies that regulate and govern specific sectors' to form 'Existing regulatory agencies'; 'Intergovernmental research organisations (e.g. CERN)' with 'Non-government scientific organisations (e.g. AAAI)' to form 'Intergovernmental and non-governmental research organisations'.
26. Eurobarometer. (2017). *Attitudes towards the impact of digitisation and automation on daily life* (Report no. 460). Selwyn, N., Gallo Cordoba, B., Andrejevic, M, & Campbell, L. (2020) *AI for Social Good? Australian public attitudes toward AI and society*. Zhang, B., & Dafoe, A. (2019). *Artificial intelligence: American attitudes and trends*.
27. Eurobarometer (2019). *Europeans and Artificial Intelligence*. The European Consumer Organization (2020). *Artificial Intelligence: what consumers say. Findings and policy recommendations of a multi-country survey on AI*.
28. Jobin, A., Ienca, M., & Vayena, E. (2019). *The global landscape of AI ethics guidelines*. *Nature Machine Intelligence, 1*(9), 389-399. See the OECD AI Policy Observatory.
29. European Commission (2019). *Ethical guidelines for trustworthy AI*
30. Ninety percent of the sample were either currently in work (67%) or had worked in the past (23%). This translated into a sample size of 15,409 (of the total 17,193) respondents completing questions about the use AI at work.
31. We adapted a measure from Haesevoets, de Cremer, Dierckx & van Hiel (2021) *Human-machine collaboration in managerial decision making*. *Computers in Human Behavior, 119*
32. See Van den Broeck, A., Ferris, D. L., Chang, C.H., & Rosen, C. C. (2016). *A review of self-determination theory's basic psychological needs at work*. *Journal of Management, 42*(5), 1195–1229. See also https://selfdeterminationtheory.org/topics/application-basic-psychological-needs/
33. Edelman AI. (2019). *2019 Edelman AI Survey*. Eurobarometer. (2017). *Attitudes towards the impact of digitisation and automation on daily life* (Report no. 460). Retrieved from https://ec.europa.eu/; Zhang, B., & Dafoe, A. (2019). *Artificial intelligence: American attitudes and trends*.
34. We grouped into an "other" category people who are not currently in work, but had worked previously (e.g., retirees, homemakers, the unemployed), those who could not be reliably coded into another category (e.g., "employee", "self-employed", "various", etc.), and those who selected "Other" but did not provide any details of their occupation. Given the miscellaneous nature of this category, we do not report on it.
35. For example, The Elements of AI course is a free online course created by the University of Helsinki and MinnaLearn. It has been completed by over 850,000 people.
36. Responses to the three items assessing subjective knowledge were aggregated to produce an overall score.
37. Structural equation modelling is a set of multivariate modeling techniques that has advantages over traditional multivariate techniques including (a) the assessment of measurement error to provide less biased effect estimates, (b) the estimation of latent variables via observed variables, and (c) the ability to assess the fit of the model to the data. Our model fit the data well: $\chi 2$ (N = 17134, df = 1740) = 46803.72, p < .001; CFI: .93, TLI: .93, SRMR: .06, RMSEA: .04. For an accessible guide to the structural equation modeling process, see Kline, R. B. (2015). *Principles and Practices of Structural Equation Modeling* (4th ed.). Guilford Press: New York.
38. 'B' refers to the standardised beta coefficient, which indicates the strength of the effect of each independent variable (i.e., driver) on the dependent variable (i.e., outcome). Beta coefficients can be compared to indicate the relative strength of each independent variable.
39. This model differs to models in our prior reports due to the inclusion of different and new predictors (e.g., perceived benefits, perceived risks, technology efficacy), the use of a more advanced statistical modelling technique (structural equation modelling instead of path analysis), and the larger and more diverse sampling (17 countries across all global regions vs. 5 western countries).
40. In 2020, we asked questions about three AI use cases: AI in general, Healthcare AI and Human Resources AI. As such, we only include 2022 comparative survey data based on these same three use cases. While the samples collected in 2020 and 2022 are based on the same methodology and panel provider, they are independent of each other. As such, our analyses examine general trends rather than a longitudinal analysis of the same respondents.
41. See the Government AI Readiness Index 2022. See also the Responsible AI Sub-Index of the Government AI Readiness Index 2020
42. *https://www.ibm.com/watson/resources/ai-adoption*
43. See Field, A. (2013). *Discovering statistics using IBM SPSS statistics (4th ed.)*. Sage: London.
44. As a rule of thumb, a Hedges g value of 0.2 is considered a small effect size, 0.5 a medium effect size, and 0.8 or larger, a large effect size (see Lakens, D. (2013) *Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAS*. *Frontiers in Psychology, 4,* 863). We chose a cut-off of 0.3 to ensure a practically meaningful and robust difference.
45. Comparative data sourced from https://data.oecd.org/eduatt/adult-education-level.htm#indicator-chart or from UNESCO where not available from OECD.
46. See Edelman (2022). *Edelman Trust Barometer 2022*. Ipsos (2022) *Global Opinions and Expectations about Artificial Intelligence*.

Key contacts

The University of Queensland

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

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

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

KPMG

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

Rita Fentener van Vlissingen
Associate Director, ESG and Impact Relations
KPMG Australia
T: +61 2 9346 6366
E: ritafentener@kpmg.com.au

Jessica Wyndham
Associate Director, KPMG Banarra, Human Rights & Social Impact
KPMG Australia
T: +61 2 9335 8573
E: jwyndham@kpmg.com.au

Richard Boele
Chief Purpose Officer
Partner, KPMG Banarra,
Human Rights & Social Impact
KPMG Australia
T: +61 2 9346 585
E: rboele@kpmg.com.au

uq.edu.au

KPMG.com.au

February 2023. 982845372CC.