MIT Technology Review
Insights / Research

Asia's AI agenda

The ethics of AI

A research program in association with:

Founding sponsor

ADP

Gold sponsor

SG:D
SINGAPORE
EMPOWERING POSSIBILITIES

Silver sponsors

GENESYS

splunk>

ASIA School of Business
in collaboration with MIT Sloan Management

Preface

Asia's AI agenda is an MIT Technology Review Insights research program sponsored by ADP, Infocomm Media Development Authority, Genesys, Splunk, and the Asia School of Business. It is designed to comprehensively examine the development of artificial intelligence (AI) in Asia Pacific from four distinct angles: Asia's AI ecosystem, the leading use cases and business applications across the region, the evolving talent landscape, and the emerging discussions around AI and ethics. To produce this series, MIT Technology Review Insights conducted a survey of almost 900 executives across 13 markets, and a series of interviews with leading authorities from academia and industry. The research is editorially independent and the views expressed are those of MIT Technology Review Insights.

The survey

- In September and October 2018, MIT Technology Review Insights surveyed 871 senior business leaders, of whom 29 % are CIOs, CTOs, or heads of AI or analytics. More than half (53%) are C-suite and director-level executives. Almost 60 % are from large companies with over $1 billion in revenue.
- Survey respondents are based in 13 Asia-Pacific economies: Australia, China, Hong Kong, India, Indonesia, Japan, Malaysia, New Zealand, the Philippines, Singapore, South Korea, Thailand, and Vietnam, with a minimum of 50 responses from each.
- Respondents are drawn from a wide range of industries, including more than 50 from each of the following sectors: consumer goods and retail; financial services; information technology and communications; manufacturing, pharmaceuticals and healthcare; professional services; property, construction, and engineering; and transport and logistics.

Expert interviews

We would like to thank the following experts for contributing their time and insights towards this research program:

Goh Yihan, Dean, School of Law, **Singapore Management University**, Singapore

Loredana Padurean, Associate Dean, **Asia School of Business**, Malaysia

Tomaso Poggio, Eugene McDermott Professor, Department of Brain and Cognitive Sciences, **MIT**, United States

Kenji Suzuki, Professor, Center for Cybernics Research and Faculty of Engineering, Information and Systems, **University of Tsukuba**, Japan

Viriya Upatising, Chief Information Officer, **True Corporation**, Thailand

Toby Walsh, Scientia Professor of Artificial Intelligence, **University of New South Wales**, Australia

Ken Wong, Head of AI Lab, **OCBC Bank**, Singapore

Xue Lan, Professor, School of Public Policy and Management, and Dean, Schwarzman College, **Tsinghua University**, China

Zee Kin Yeong, Assistant Chief Executive, Data Innovation and Protection Group, **Infocomm Media Development Authority**, Singapore

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

Contents

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

1. Executive summary

Globally, future outlooks for artificial intelligence (AI) swing between two extremes—excited anticipation about the positive impact AI will have on economies and societies, and deepening fear of its potential to disrupt livelihoods and do harm. In Asia, governments and civil society groups are concerned about defining regulatory frameworks to guard against the latter, and all ecosystem experts are grappling with how to steer AI toward the former, more socially advantageous directions. On balance, however, the Asian business leaders surveyed for this report have great optimism about AI's positive effect on their businesses, societies, and individual well-being.

This report, combining an Asia-wide executive survey with expert interviews from industry, government, and academia, takes the pulse of public and private actors in the AI ethics debate in the region. The key findings of the report are:

- **AI will be a major growth driver for Asia in the coming decade.** The company priorities for AI are to enhance customer satisfaction, speed up decision-making, and reduce inefficiencies. The loss of some roles to automation, and the restructuring of others to take advantage of technology-created capacity, are likely. Yet reducing headcount is not a top priority in and of itself. Just one-third of survey respondents listed the need to reduce labor costs as a top-three driver for AI.

- **Biases within AI tools are potentially dangerous for Asia—but biases about AI's use in Asia could be even more so.** Asia's AI ecosystem participants are aware of and concerned about the potential for embedded biases (race, gender, or socio-economic status) within AI tools, and the harm this can cause through facilitating over-policing of minority communities, or economic exclusion. Weaponization and malicious use of AI are also ethical concerns in Asia as applications are increasingly commoditized and industrialized. While Asian decision-makers are concerned about a potentially negative impact, particularly where jobs are concerned, optimism is the more dominant sentiment, which will propel the use of AI in Asia.

- **Asian governments are building institutional capacity and frameworks to increase AI governance—but have yet to develop regulations.** Overwhelmingly, more survey respondents believe Asia will lead the world in the development of ethics and governance than any other region: 45%, as compared to only a quarter who see North America as the ethics front-runner. Across the region, from Singapore to Japan and China, governments are assembling AI institutions to guide governance, often consulting with the private sector and civil society.

- **Asian respondents are engaged in AI ethics discussions and see a constructive role for governments.** Just under half (42%) of participants say there is "vigorous debate" on ethical issues surrounding AI in their company, and the majority (55%) think AI should be government-regulated. Both Asia's governments and its businesses are committed to maximizing AI's benefits, so the fact that talent availability is the region's top AI deployment challenge (according to 58% of respondents) suggests that those with AI skills will wield significant bargaining power in the debate which lies ahead.

- **AI-driven unemployment narratives are counterbalanced by the potential to enhance and augment human work.** Many survey respondents share one of the world's biggest ethical fears—that the unchecked use of AI will result in massive loss of jobs and livelihoods; 42% believe that the rise of AI in Asia will destroy more jobs than it will create. However, other survey responses reveal that Asian business leaders are not overly concerned about job loss in their own organizations, and they believe that AI will benefit their employees. The majority (59%) of respondents believe their employees' job roles have been enhanced since the introduction of AI.

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

2. Introduction: Who's afraid of AI?

AI technology is far from new, but its capability has leaped forward over the past half decade, allowing it to move from the margins to the mainstream. AI is increasingly becoming industrialized and commoditized in areas ranging from logistics and health diagnostics to personal assistants and vehicles. Across Asia, the ecosystem is responding to watershed moments in global AI development with competitive zeal and redoubled efforts.

The AI shake-up of Go, the adversarial board game known for its momentous complexity, is one such recent example. When AlphaGo, a program developed by Alphabet's AI company DeepMind Technologies, defeated Korean international Go champion Lee Sedol in March 2016, it was considered a singular achievement. When a successor program, AlphaGo Master, defeated China's world champion Ke Jie a year later in May 2017, it set off what has been described as China's "Sputnik moment" (referring to the panic set off in the U.S. when the Soviet Union launched the first satellite into orbit in 1957, spurring the space race).

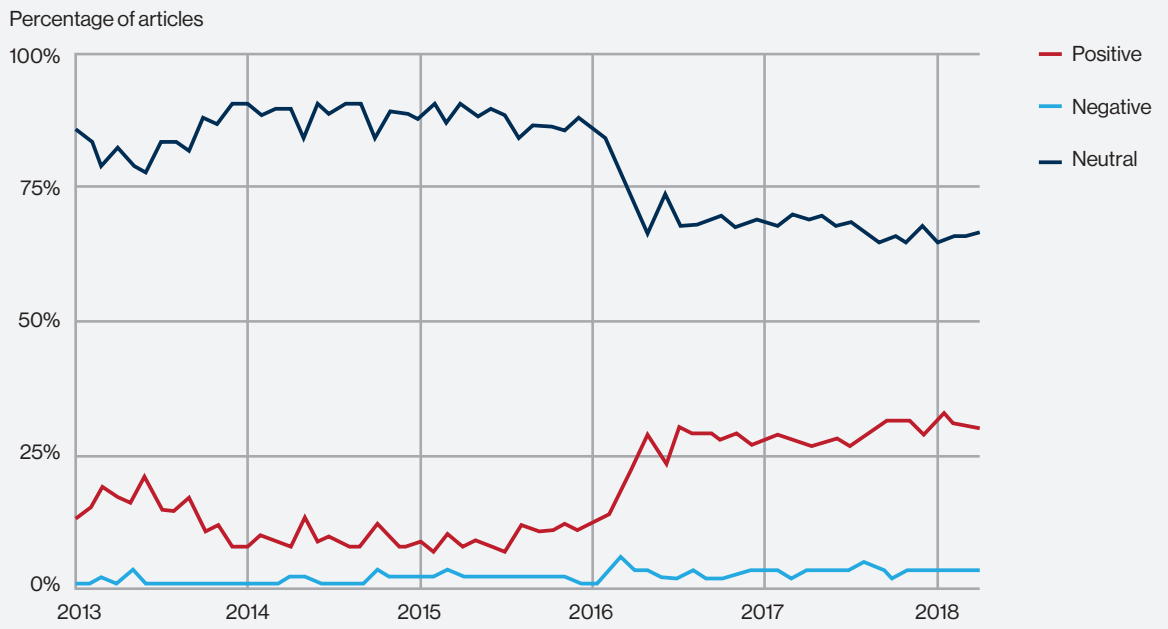
The next space race

China's accelerated AI development plan, known as the "AI 2030 agenda," was released the following month, and articulated an overarching national vision for the numerous technology initiatives already underway. Among these was a homegrown response to AlphaGo: by January 2018, a Go program called Fine Art, developed by Chinese internet giant Tencent, had also bested Ke Jie.

While media coverage of AlphaGo and Fine Art, like most stories on AI over the past few years, reported on the victories of computers over humans, articles written about AI since 2016 have become markedly more positive in their sentiment about the technology.[1]

At the same time, a recent World Economic Forum survey of public and private sector leaders

Figure 1: Sentiment analysis of articles referencing AI, 2013-18



Source: TrendKite, Artificial Intelligence Index 2018[1]

[1] *cdn.aiindex.org/2018/AI%20Index%202018%20Annual%20Report.pdf*

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

found that AI and robotics were seen as the emerging technologies with the greatest potential negative impact (see Figure 2).[2] This shows that while attitudes towards AI are generally improving, there are still grave concerns about the knock-on effects.

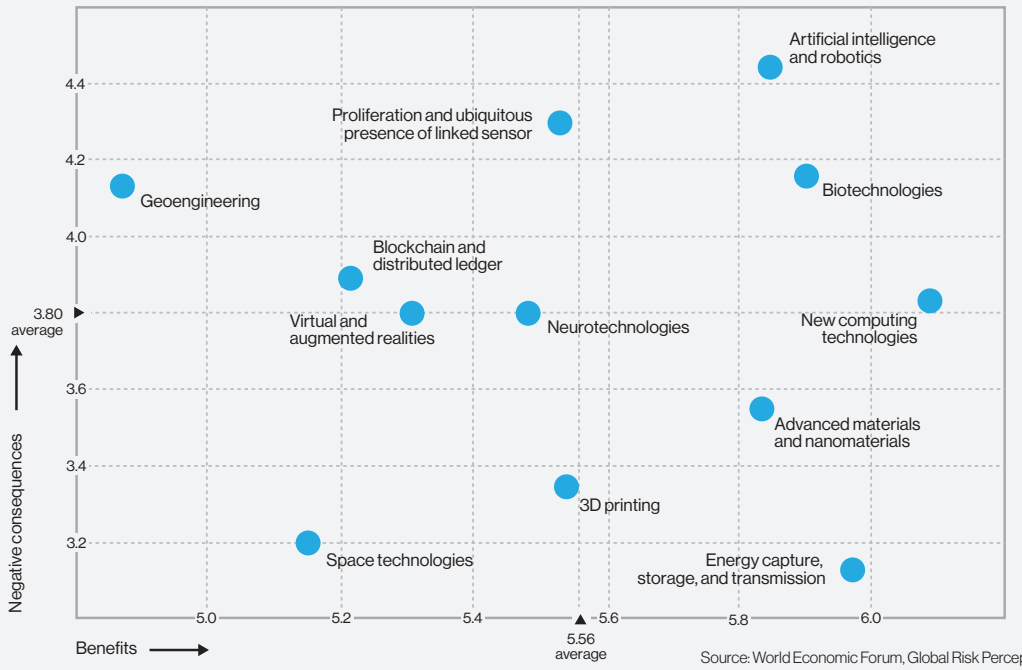
"The first car was invented in the late 1800s, but it took a century for our use of automotive technology to fully mature," says Loredana Padurean, associate dean and faculty director for action learning at the **Asia School of Business**, a collaboration between the Central Bank of Malaysia and MIT's Sloan School of Management. "AI is advancing at 10 times that speed—AI is going to advance faster than any other innovation we have ever seen," she says, noting the insufficient time for debating the potential impact on our societies and economies.

The risks of AI and autonomous systems are not hypothetical. Biases, accidents, and controversial uses (such as AI-enabled weapons and other military applications) have all been features of the AI landscape over the past two years. Asian government and business leaders are not blind to these risks. However, as this report will argue, AI ecosystem participants are more sanguine about the "upsides" of AI, and are confident that Asia has the will, the resources, and the institutional capacity to develop robust ethical frameworks. Asia's potent cocktail of pro-AI enthusiasm and pragmatism may create further ethical conundrums, but this report argues the region has experience in using technology to increase social harmony and economic sustainability.

"It took a century for our use of automotive technology to mature. AI is advancing at 10 times that speed—AI is going to advance faster than any innovation we have ever seen."

Loredana Padurean
Associate Dean
Asia School of Business

Figure 2: Perceived benefits and negative consequences of 12 emerging technologies



Source: World Economic Forum, Global Risk Perception Survey, 2017

[2] *reports.weforum.org/global-risks-2017/part-3-emerging-technologies/3-1-understanding-the-risk-landscape/*

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

3. AI battlefields

Bias is a prominent worry in the development of AI because algorithms discriminate in order to make decisions, and those bases are literally "hard-coded" by technical experts. The fact that men account for a much larger share of the AI workforce than women creates the potential for gender biases in algorithm development.

The World Economic Forum "Global Gender Gap Report, 2018" warned that, with women accounting for 22% of AI roles globally, machine learning technologies are being encoded with perspectives that are intrinsically non-representative of the societies they serve.[3] The report observed that in the three top-ranked economies by AI talent, the United States and India had significant gender skews towards male programmers (70 and 72% respectively), and Germany, at 82%, had one on the world's widest gaps. At 61%, Singapore, the report's eighth-ranked AI economy, was the only Asian country in the top 20 to have a substantially smaller gender gap.

In the July 2018 "China AI Development Report", the China Institute for Science and Technology Policy at Tsinghua University analyzed the results of an "AI interest" survey of users of Toutioa, a domestic news-aggregation platform, finding that throughout 2017, 66% of all views of AI-related articles were by men.[4] Despite efforts to bring women into scientific and technical careers, it is clear that more still needs to be done.

Figure 3: Share of professionals with AI skills, by gender and country

Country rank, AI skills	Country	Female (%)	Male (%)
1	United States	23	77
2	India	22	78
3	Germany	16	84
4	Switzerland	19	81
5	Canada	24	76
6	France	21	79
7	Spain	19	81
8	Singapore	28	72
9	Sweden	20	80
10	United Kingdom	20	80
11	Netherlands	21	79
12	Poland	16	84
13	Australia	24	76
14	Brazil	14	86
15	Italy	28	72
16	Turkey	24	76
17	Belgium	19	81
18	South Africa	28	72
19	Mexico	15	85
20	Argentina	17	83

Source: World Economic Forum, Global Gender Gap Report, 2018

Who, me?

Race is another bias that can AI can learn from its developers. In the United States and Europe, facial recognition has proven less effective in identifying non-white faces, and researchers have shown that some such tools are more likely to misidentify minorities as criminals. Such systems are already in use by many police forces globally, giving rise to fears that racial biases in AI-enabled facial recognition can lead to "false positives" and wrongful arrests. California's high-tech heartlands of San Francisco and Oakland may soon pass legislation banning the use of facial recognition

[3] *www3.weforum.org/docs/WEF_GGGR_2018.pdf*
[4] *www.sppm.tsinghua.edu.cn/eWebEditor/UploadFile/China_AI_development_report_2018.pdf*

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

and other digital surveillance tools by their municipal governments.[5] The need to guard against bias intensifies as companies move into sensitive sectors, such as justice and the law. "You have more legal tech startups coming into the picture using AI, but some of these startups are not even fronted by legally trained people," warns Goh Yihan, dean of the **Singapore Management University** School of Law.

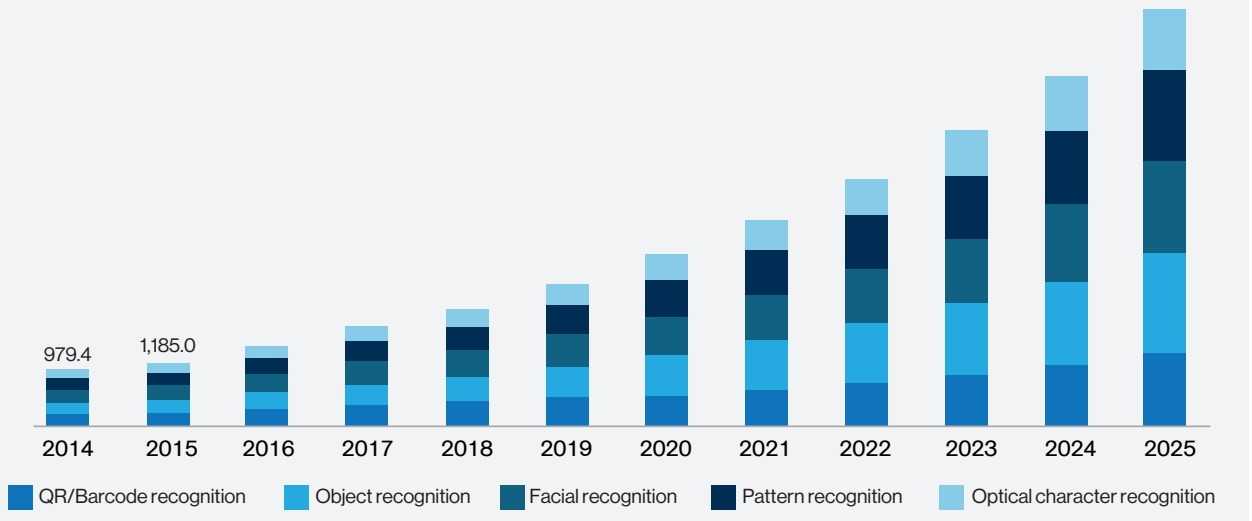
Nevertheless, Asia is likely to become a global frontrunner in the use of facial recognition technology in the coming decade. Market research firm Statistics Market Research Consulting predicts that global facial recognition technology sales will more than triple over the coming decade, from $4 billion in 2017 to nearly $15 billion in 2026.[6] Analysts estimate that half of global demand comes from Asia, and China specifically is recognized as both the world's single largest market and the leading producer of facial recognition technologies.

Since 2015 China been building the world's largest face recognition database, which by 2020 is intended to contain the biometric details of all its 1.3 billion citizens, and a growing body of evidence suggests that biases are being deployed with intent. A recent New York Times report claimed Chinese authorities are using facial recognition to identify and track ethnic Uighurs in cities outside of their traditional home province of Xinjiang.[7] A Dutch cybersecurity researcher, while identifying security breaches in Chinese facial image databases, recently discovered a cache of personal data and images on 1.8 million Chinese women being used to determine their so-called 'BreedReady' status.[8]

Other efforts to control Chinese citizens on socio-political, rather than criminal, grounds include its social credit system that uses AI, big data, and facial recognition to profile and place restrictions on those who fail to pay debts or flout other rules. China's world-leading image-recognition market raises profound questions about not only the intended use, but also the manner in which the technology is being created.

While few Asian governments have China's resources or careful orchestration between the state and industry, most do share a heightened concern for public security and social harmony. AI can, of course, foster public safety and improve resource allocation, but China's example reveals how quickly its application can stray into reinforcing political, ethnic, or ideological biases in the name of ethical goals.

Figure 4: China image recognition market size, by technique, 2014-25 (US$ m)



Source: Grand View Research, 2018

[5] *www.slate.com/technology/2019/05/san-francisco-first-city-ban-facial-recognition.html*
[6] *www.reuters.com/brandfeatures/venture-capital/article?id=48514*
[7] *www.nytimes.com/2019/05/22/world/asia/china-surveillance-xinjiang.html*
[8] *www.theguardian.com/world/2019/mar/11/china-database-lists-breedready-status-of-18-million-women?CMP=twt_gu*

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

Malicious use and weaponization

Dystopian scenarios involving AI typically revolve around what would occur should powerful tools fall into the wrong hands. A machine learning system recently developed by academic researchers at the University College of London and the Alan Turning Institute re-identified anonymized Twitter users through meta-data analysis with 96.7% accuracy.[9] If hackers can develop or appropriate such models themselves, the implications for personal data privacy are severe.

Even more severe are the implications if bad actors take over AI-controlled weapons. Lethal autonomous weapons (LAWs) are systems which can, of their own volition, identify a target and complete a lethal attack without human approval or authorization. There is no evidence that a LAW yet exists, but there are technological precursors in several Asian countries, including South Korea and China. While there is currently no international consensus on how LAWs fit into frameworks governing war crimes and accountability, Human Rights Watch believes LAWs contravene the Geneva Convention Martens Clause, which requires emerging technologies not covered by existing treaties to be judged by the "principles of humanity" and the "dictates of public conscience".[10]

Dangerous technologies, from guns to nuclear reactors, can be controlled in a way that AI cannot. "All sorts of algorithms could be repositioned by a person with a bad intent" says Toby Walsh, scientia professor of computer science and engineering at the **University of South Wales**, Sydney. He points to possibilities for radar-linked AI that identifies pedestrians in autonomous-vehicle technology being programmed for drones to hunt targets. Mitigating the risk of malicious AI use is not, he says, possible through out-and-out bans: "You're not going to stop AI technology, just like we didn't stop chemistry by banning chemical weapons, and we didn't stop biological research by banning biological weapons," he says. That said, there is evidence that the AI ecosystem has other means to limit engagement in weaponized AI and lethal weapons. Walsh, along with a coalition of over 50 AI academics, used a boycott to get the Korea Advanced Institute of Science and Technology to agree not to develop AI weapons as part of a collaboration with a defense company.

"You're not going to stop AI technology, just like we didn't stop chemistry by banning chemical weapons, and we didn't stop biological research by banning biological weapons."

Toby Walsh
Scientia Professor of Computer Science and Engineering
The University of South Wales

Malicious AI is also playing out in social media, through disinformation and bot-driven "drown out" campaigns. These reportedly played out in Saudi Arabia in the aftermath of the disappearance of journalist Jamal Khashoggi, and Myanmar's government has been accused of leveraging disinformation to kindle religious and ethnic conflict and antipathy against the Rohingya. Non-state groups also stand accused. In Indonesia, police claim to have uncovered a clandestine fake news operation geared to destabilize its government.

Inequality and unemployment

As bots carry out more complex knowledge tasks in both white- and blue-collar work, there is concern that AI will split labor pools into a "winning" class of workers—whose roles and responsibilities are augmented by AI—and a losing class whose jobs are replaced by automated tools or eliminated entirely. Overall, AI will affect one in every five Asian jobs, and automation will eliminate one in eight, according to the MIT Technology Review Insights report "AI and human capital".[11] The report also found that emerging Asian economies, which are built on labor-intensive industries and services

[9] *www.wired.co.uk/article/twitter-metadata-user-privacy*
[10] *www. treaties.un.org/doc/Treaties/1983/12/19831202%2001-19%20AM/XXVI-2-revised.pdf*
[11] *www.insights.techreview.com/asias-ai-agenda-ai-and-human-capital/*

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

from electronics and textiles to business process outsourcing, are much more at risk of job losses due to AI and automation than their wealthier counterparts—even though greater percentages of "automatable" jobs will disappear from the workforces of rich Asian countries. Poorer Asian economies are proportionately more dependent on the lower-skilled job classes endangered by AI and, unlike their wealthier peers, have fewer high-skill jobs that will be enhanced by AI, and less institutional capacity to retrain and reskill workers.

AI's potential threat to jobs is not lost on our survey respondents. Some 42% think AI will either destroy or disintermediate more jobs than it creates in Asia and only 19% disagreed. These concerns are filtering down to front-line staff in many Asian organizations. "'Is my job going away because of AI?' is a common question we are asked by staff in nearly every town hall meeting," says Virya Upatising, chief information officer at **True Corporation**, the Thailand-based telecommunications firm.

There is also growing anxiety in China, despite the country's overall culture of techno-optimism, according to Xue Lan, dean of Schwarzman College at **Tsinghua University**, who led the team responsible for the China AI Development Report. "There have been many heated debates in China about the potential impact of AI on employment. In our report we found that, in general, the public is very supportive of AI development, but they are simultaneously concerned about AI's impact on their jobs". Xue notes that the potential impact of AI is a subject to discussion in the general public and also by professional societies, leading to the creation of research groups and organizations.

The majority of Asia-based respondents to MIT Technology Review's survey agree with positive statements about the impact of AI. Four-fifths believe AI will increase productivity overall, 59% believe their employees have had their roles enhanced by AI, and nearly half believe that overall employee satisfaction is higher since the introduction of AI. It is possible that these positive views are borne out of necessity or an "anchoring bias"; some 70% of respondents say that AI is essential to their future competitiveness.

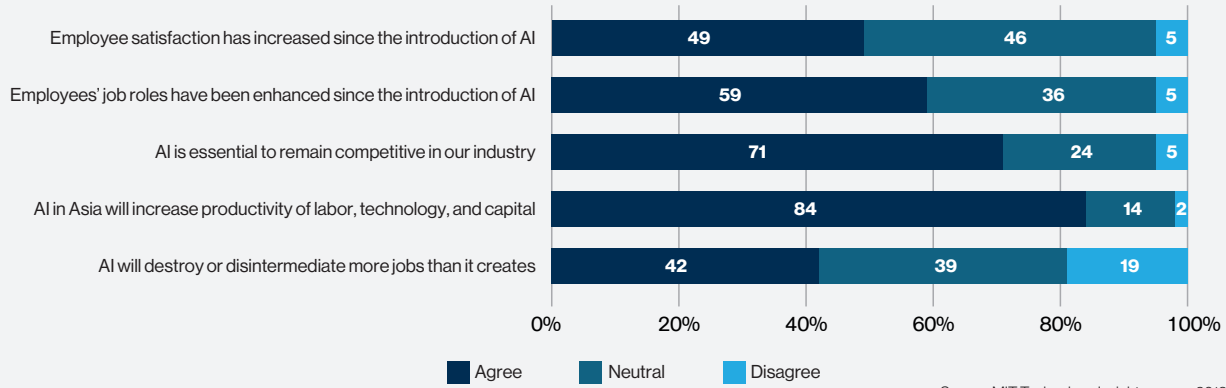
Key takeaway

Race and gender bias—Improving the race and gender diversity of Asia's AI developer pool is critical for mitigating biases - however genuinely unintended.

The all-seeing eye—China's facial recognition lead will be of benefit to some but not others. There will be countless new use cases boosting efficiency, security, and public safety. The price will be diminished civil liberties.

Jobs will go—Only 19% of survey respondents in Asia do not believe that AI will destroy or disintermediate more jobs than it creates.

Figure 5: Perceived impact of AI in Asia (% of respondents)



Source: MIT Technology Insights survey, 2018

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

4. Emerging governance for AI

Concrete responses to today's AI ethics debate from governments and companies are at a nascent stage. Yet with Asia likely being at the forefront of AI development globally over the coming decade, it is also probable that the region will run into ethical dilemmas early on.

To date, the AI-ethics debate has been largely diagnostic: identifying the risks to individuals and societies posed by current or future technologies. Concrete responses from governments and AI technology companies are at a nascent stage. Yet with Asia likely to be at the forefront of AI development and adoption globally over the coming decade, it is also likely that Asia will run into ethical issues and dilemmas early on. Nearly half of survey respondents believe Asia will lead the development of ethics and governance frameworks, compared to only 25% seeing North America as an ethics leader. The real challenge for Asia's economies will be to temper the tendency to rush the deployment of AI to overcome critical infrastructure gaps and to increase citizen welfare with a careful examination of the potential risks to society that may be created as a result.

New institutions and global partnerships

Early movers in convening expertise and coordinating policy positions are to be found in Europe. The UK government's Centre for Data Ethics and Innovation launched in 2017 and Germany followed with the Data Ethics Commission the following year. Canada and France have formed an international coalition on AI ethics, set to include other Group of Seven and EU members. But Asian governments are not far behind in these efforts. Singapore formed the Advisory Council on the Ethical use of AI and Data in 2018, with advisory members drawn from industry, government, and civil society. The Indian government has tasked its agencies, notably NITI Aayog, to take up the policy work program for AI and publish guidance. Chinese authorities convened an ethics committee in 2018, led by leading AI scientist Chen Xiaoping, and senior figures have called for more international institutional collaboration.

Asian universities are also working closely with governments and global partners on ethics and policy. The Hong Kong University of Science and Technology (HKUST) was the first Asian university to join the "Partnership on AI to Benefit People and Society", an international alliance of more than 80 companies, that has a number of goals for shaping the development of AI in aspirational and benevolent ways. HKUST's addition has significance, given that Sino-US technology collaboration has slowed in the mounting trade war, and is a reminder of Hong Kong's unique position as a transition node for transferring technology and capital between the two superpowers. There are now many university-led AI ethics initiatives, including New Zealand's University of Otago forming a Centre for Artificial Intelligence and Public Policy to work with the ministry of broadcasting, communications and digital media, and the government's digital services division on policy options.

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

Codes and principles

Goals and broad ethical principles are beginning to be shaped. Policy documents include Japan's AI Technology Strategy (March 2017), China's Next Generation Plan (July 2017), and India's #AIforAll national strategy (June 2018), with key principles including ensuring AI has broad benefits for the nation's development. The Australian government recently earmarked AU$ 29.9m (US$ 20.6m) for a four-year AI and machine learning program, including the development of an "AI Ethics Roadmap". Malaysia's government has announced the development of a national framework for AI to be completed by the end of this year. Most of these frameworks only address the ethical issues in broad terms, and largely informed by pragmatic concerns around job loss and reskilling requirements. The Malaysia Digital Economy Corporation (MDEC), the country's technology promotion and investment coordination body, which is leading the AI framework development, is using it in large part to coordinate knowledge-sharing and best-practice development between Malaysia's AI ecosystem participants (academia, government bodies, and established and startup enterprises) to build coordinated responses to the implications that AI will have on skills, livelihoods, and economic competitiveness.

Other ethical framework implementation efforts are more incentives-driven: Australia's chief scientist Alan Finkel recently proposed that his

Making ethics frameworks actionable

Experts believe the plethora of government guidelines need to balance detail and specificity with flexibility to evolve as technologies do. "Rather than add to existing pool of ethical principles that have been generated by governments, we thought it would be more useful to produce a framework to translate ethical principles into things companies can do in the office," says Zee Kin Yeong, assistant chief executive of the **Infocomm Media Development Authority** (IMDA) of Singapore, and deputy commissioner at the Personal Data Protection Commission. "How do we ensure that ethical principles are properly incorporated into the company's internal governance framework as part of their risk assessments, as part of their code of ethics? And how do we make use of existing corporate governance structures to ensure that there are proper decisions being made at the right level, proper monitoring to identify any incidents, and proper ways of figuring out what are the best ways of using AI to help make decisions within the company?"

IMDA has released a model framework for public consultation to help organizations ensure that decisions made by, or with assistance from, AI are explainable, transparent, and fair to consumers, and that AI solutions are "human-centric". The model framework, in many ways, mirrors government oversight used in corporate governance generally. For instance, it suggests that organizations using AI need to allocate roles and articulate responsibilities for assessing the risks involved in AI adoption and managing the AI model adoption process, as well as reviewing outputs and ensuring effective transparency and customer communication.[15]

In its launch, the IMDA argued that such a framework would enhance trust in, and understanding of, AI and drive acceptance of how decisions are made. Building on a June 2018 discussion paper by the country's data protection commission and IMDA, it maps four ethical principles applying to AI deployment in four areas: internal governance, risk management in autonomous decision-making, operations management, and customer relationship management.

Goh Yihan, dean of the Singapore Management University School of Law, emphasizes the need for frameworks to be agile and evolving. "There is no point in having a framework that is so detailed but is also inflexible and doesn't react to the fast-moving pace of technology. A framework needs to be easily comprehensible. If you have academics write the framework with lawyers, it may run into complex language, and that's not helpful because AI and algorithms are a tool used not only by industry or certain professions, but across different industries."

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

government implement a certification process to award firms that demonstrate they have adopted AI in a responsible manner, in terms of job-loss management and sustainable manufacturing process. Finkel calls it the Turing Certificate—after the computer scientist (and recognized "founding father" of AI) Alan Turing—and believes that making it mandatory for firms seeking government contracts will send a strong signal to both the Australian and global economy.[12]

Despite the region's prodigious road-mapping, however, articulated positions on ethics are lagging in Asia, relative to other regions, and in terms of other AI development priorities. In an analysis of 18 national and regional AI strategic plans, the Canadian Institute For Advanced Research found that the strategies coming from Asia placed relatively lower priority on ethics than those developed in other regions, although in one related area—the use of AI to promote social and economic inclusion—India scored highest globally.[13] The heat map shows that Asian economies are still largely focused on pointing AI at industrial development strategies.

Along with the ethical chapters in national strategies, governments are also issuing codes and charters, although most lack legal status. In 2012, South Korea issued a "Robot Ethics Charter", which covers standards, illegal use, and data protection. These are written as guidelines rather than laws. Japan issued a "Robot Strategy" three years later (2015) covering policy, ethics, and safety standards.

Figure 6: Heatmap of national AI strategy priorities

Source: CIFAR, 2018

[12] *www.governmentnews.com.au/government-should-lead-ai-certification-finkel/*
[13] *www.cifar.ca/docs/default-source/ai-society/buildinganaiworld_eng.pdf*
[14] *www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/A-Proposed-Model-AI-Governance-Framework-January-2019.pdf*

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

China's "Next Generation AI Development Plan", published in July 2017, envisions AI "to improve social management capacity" and pledges research on civil and criminal liability, privacy and IP, information safety, accountability, design ethics, risk assessment, and emergency responses, and commits to participate in AI global governance.

In 2018, Singapore's monetary authority introduced principles to promote fairness, ethics, accountability, and transparency (FEAT) in AI and data analytics in finance. The FEAT Principles require firms to demonstrate that their use of AI Data Analytics (AIDA) does not result in systematic disadvantages for individuals; that analytic models and frameworks are regularly reviewed to reduce or remove biases, and that any decisions based on AIDA are held to the same codes of conduct and same level of ethical scrutiny as would be human-driven decisions.[15]

Companies and the AI ethics debate

Ethical AI is not just a question for government regulators; companies are also part of the debate and potentially the solution, through participation in international fora, and their internal development agendas. Some 42% of Asian business leaders in our survey indicated that there was "vigorous debate" around ethical questions in their companies. Evidence to date shows their activities have included entering international networks, advocating for rules and clarity from government, developing internal codes, and embarking on ethical and pro-social AI product development.

Asian firms have become more active participants in international ethics arenas. Baidu became the first mainland Chinese representation in the Partnership on AI, an alliance with stakeholders including Apple, Google and Amazon. Fujitsu Laboratories of Europe, part of Japan-based Fujitsu, is one of the founding partners of the AI4People initiative, a European forum on the social impacts of AI.

Yet survey data shows that Asian companies are looking to government to take the lead—the majority think it should be government-regulated. Chief executives of Baidu and Tencent, Robin Li and Pony Ma, have recently stated that China must develop guidelines in areas that raise ethical questions, such as autonomous vehicles, gene editing, and data privacy.[16] Li advocates government-led research into ethical AI, telling a meeting of the Chinese People's Political Consultative Conference that "only by establishing a sound set of ethical norms … can we reap more benefits from AI."

There are numerous examples of corporate AI guidelines: Microsoft has a research program called FATE (Fairness, Accountability, Transparency and Ethics), exploring how to combine AI innovation and ethical use; Google's DeepMind has formed an Ethics and Society division; Sony has issued a seven-point AI ethics guideline.[17] Asian companies, or global firms with an Asia presence interviewed for this study, indicate they are developing their own. Genesys, one of the sponsors of this report, has a five-point guide to AI ethics covering transparency, fairness, accountability, data protection, and social benefit. This is not merely a public relations exercise. The firm predicts that by 2021, algorithm opacity, decision bias, malicious use of AI, and data regulations will result in a doubling of spending on relevant governance and compliance staff.

Key takeaway

Frameworks over enforcement—Asian governments and businesses are becoming more active in setting goals and guidelines for the ethical development of the AI industry. Yet thus far, there are no oversight or enforcement mechanisms. Rather than tackling ethics head on and providing solid legal backing and recourse, Asian policymakers are fostering trust between consumers, users of AI, and AI developers so that the industry as a whole can grow.

[15] *www.mas.gov.sg/-/media/resource/news_room/press_releases/2018/Annex%20A%20Summary%20of%20the%20FEAT%20Principles.pdf*
[16] *www.abacusnews.com/future-tech/china-wants-make-its-own-rules-ai-ethics/article/3001025*
[17] *www.microsoft.com/en-us/research/group/fate/*

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

5. Challenges for AI ethics: the view from the private sector

Building robust and comprehensive responses to the ethical challenges, real and hypothetical, raised by AI is far from straightforward.

AI is not yet replacing workers outright

While it is clear that AI is assisting companies in making routine tasks more efficient, the link between job losses and AI is still opaque. Moreover, should we not grasp at opportunities to pass off low-value work so that humans can focus on higher-order cognitive tasks? One executive interviewed for this report asked "How ethical is it to keep people doing menial tasks when they could actually be taking a step up, when computers can do this for them?"

It depends on how we think about the workforce as a whole, argues Loredana Padurean at the Asia School of Business, rather than a simple question of whether certain jobs will come or go. "If we do it well, AI can improve our life dramatically, because we remove people from working in toxic or demeaning environments. The question is, what do we do with these people? Can we create societies that share the spoils of victory, and find better places for these people in society?"

Survey participants concur that unemployment hype is over-baked; only 34% viewed reducing labor costs as an important business driver for deploying AI. Improving decision-making speed and quality, in contrast, was voted a top driver by a far higher share of 51%.

There is no "global"

It is intuitive to discuss AI ethics at a global level, but many research studies have shown that there are major cultural differences in what could be considered the "right" way to articulate ethics. "Ethics is linked with morality and culture, which differs from region to region and culture to culture,

"If we do it well, AI can improve our life dramatically, because we remove people from working in toxic or demeaning environments. The question is, what do we do with these people? Can we create societies that share the spoils of victory, and find better places for these people in society?"

Loredana Padurean
Associate Dean
Asia School of Business

and even within a single country," says Zee Kin Yeong at IMDA. Cultural factors shape how citizens view the penetration of AI into realms like the home. 'Care-bots' for providing home-based health care and social services have taken root in Japan, where robots are ascribed human qualities. "People [in Japan] think about AI not as a computing machine, but more like a human being," says Professor Kenji Suzuki, Center for Cybernics Research and Faculty of Engineering at the **University of Tsukuba**. "Different attitudes to AI lead to different approaches to developing it within society."

The infamous Moral Machine trolley experiment posits that a car's brakes fail, leaving it to either stay on course and kill three elderly people—two men and a woman—who disobeyed a "do not cross" signal, or swerve and kill its three passengers: an adult man, a woman, and a boy. The Moral Machine project analyzed 40m decisions from 233 countries

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

to reveal that respondents from collectivist cultures, like China and Japan, were less likely to spare the young over the old.[18] The element of risk-taking, that some pedestrians were breaking the rules, also played a part in the decision-making of the respondents in some countries.

Unintended consequences

A third critique is that of unintended consequences from proposed regulatory responses. Algorithmic transparency, for instance, seeks to improve accountability, transparency, and fairness by mandating the sharing of information about the functioning of algorithms responsible for automated decisions. It is already outlined in the European Union's General Data Protection Regulation and elsewhere, including France's Digital Republic Bill. Yet some tech companies and technology industry groups warn that such approaches could, by revealing how algorithms work, enabling hacking, the "gaming" of algorithmic systems, and even intellectual property theft.[19] It is also unclear how the working of complex algorithms can be translated into language that the unskilled can meaningfully understand.

Rather than developing new AI regulations or codes, it might be more practical and achievable to ensure that AI does not transgress existing civil rights. Some legal scholars argue that AI codes are ambiguous and lack accountability. More effective would be to govern AI according to existing institutions, like the international human bill of rights; if an AI removes a person's rights, it should not be acceptable.[20]

While some AI innovations might be genuinely new (such as autonomous vehicles), experts believe we are still very far from the dawn of general AI, which would genuinely challenge our existing rules. "We are very far still from really building or knowing how to build intelligent machines—intelligent in a sense of not just better at playing chess, or golf, or driving a car," says Tomaso Poggio at **MIT**. "This is the greatest problem in science, and could be up to 50 years away." There are many other "more pressing and immediate risks to mankind, from global warming to existing nuclear weapons."

Key takeaway

A global approach is unlikely—Many countries and international organizations are forming panels and working groups to understand the risks posed by AI and create frameworks for guiding its future development. The emergence of a single framework is unlikely given the many subtle differences around the world in culture and perception of ethical rights and wrongs. Yet the most important thing is interoperability, that the frameworks being developed are flexible so that the ecosystem can thrive.

Guidelines for growth—Executives in Asia understand that AI must be developed in a way that is responsible and transparent. Communicating clearly with customers and stakeholders about how data is gathered, used and shared will continue building trust. Innovative policy approaches around data governance are also likely to emerge.

[18] *www.nature.com/articles/s41586-018-0637-6*
[19] *www.techcrunch.com/2018/03/12/report-calls-for-algorithmic-transparency-and-education-to-fight-fake-news/?guccounter=1*
[20] *www.technologyreview.com/s/612318/establishing-an-ai-code-of-ethics-will-be-harder-than-people-think/*

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

7. Conclusion

In Asia, AI is being industrialized and commercialized at scale, humming in the background everywhere from hospital diagnostic clinics to smartphones, banks, and military facilities. Concern about the potential threats to societies and individuals that this technology poses may be outweighed by confidence about the benefits, but it is not absent. This report's conclusions are:

1. **Frameworks over regulation.** Stakeholders in Asia's AI ecosystem are calling for greater pragmatism and precision in the way that the AI ethics debate is framed. Some risks are already covered by regulations, whether human rights laws or data privacy rules. Survey respondents are looking for governments to lead, but any policy, frameworks, or regulation must be created sensitively and without stifling innovation.

2. **Transparency and trust.** Governments can focus on building a trusted and socially beneficial AI ecosystem by strengthening the foundations for AI development. This includes ensuring that data is collected in an ethical way and holding companies and organizations to account for bias or discriminatory practices. Trust-building also involves public education about AI and managing data controls and privacy. Proponents of algorithmic transparency argue that data processes and the ways in which AIs make decisions should be open and accountable, preventing the erosion of human rights.

3. **The best and the worst of AI.** In the decade ahead, it is likely that Asia will see some of the fastest applications of AI in areas such as healthcare, agriculture, financial services, and many other areas that benefit people's lives. Already, countries like China are racing ahead in the use of AI to democratize access to medical services. Yet the flip side is that Asia is also already witnessing some of the most invasive and controversial uses of AI for state surveillance and social control. This is likely to continue; the establishment of supranational ethical principles guiding the use of AI by governments will be a vital step.

4. **Maintaining harmony between humans and machines.** Our survey shows that business leaders are concerned that AI will remove more jobs than it creates across the region. Private and public sector organizations must think broadly about people capabilities and invest heavily in reskilling. The businesses that are able to, for example, train displaced finance staff to become IT and cybersecurity specialists will save substantially on recruiting costs in the long term. For policymakers, the challenge is to balance the continual push toward industrial digitalization with investments in long-term human capital development.

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

About MIT Technology Review Insights

For more than 100 years MIT Technology Review has served as the world's longest-running technology magazine, the standard bearer of news and insights on how the latest technologies affect the world around us. Read by a global community of innovators, entrepreneurs, investors and executives at the highest level, it offers an unrivaled authority that is backed by the world's foremost technology institution, and features editors with a deep technical knowledge and understanding of technological advances.

MIT Technology Review Insights is the content solutions division of MIT Technology Review. It includes two main divisions: Research and Live Events. Aligned with the same stellar editorial heritage and standards as the magazine itself, we leverage our access to a wide network of subject matter experts and leading content contributors to create custom content for clients who want to reach new audiences with relevant, cogent and high-quality stories and experiences to users wherever they want it—in digital, print, online, and via unique in-person experiences.

While every effort has been taken to verify the accuracy of this information, MIT Technology Review Insights cannot accept any responsibility or liability for reliance by any person on this report or any of the information, opinions or conclusions set out in this report.

© Copyright MIT Technology Review Insights, 2019. All Rights Reserved.

MIT Technology Review Insights

insights.techreview.com

@techreview @mittr_insights

insights@technologyreview.com