

Singapore Management University Institutional Knowledge at Singapore Management University

Centre for AI & Data Governance

SMU Institutes, Centres, Labs & Initiatives

8-2019

How to address the AI Governance discussion? What can we learn from Singapore's AI strategy?


Nydia REMOLINA LEON

Singapore Management University, nydiarl@smu.edu.sg

Siew Hui SEAH

Singapore Management University, shseah@smu.edu.sg

Follow this and additional works at: <https://ink.library.smu.edu.sg/caidg>

 Part of the [Asian Studies Commons](#), and the [Science and Technology Law Commons](#)

Citation

REMOLINA LEON, Nydia and SEAH, Siew Hui. How to address the AI Governance discussion? What can we learn from Singapore's AI strategy?. (2019). 1-18. Centre for AI & Data Governance.

Available at: <https://ink.library.smu.edu.sg/caidg/1>

This Working Paper is brought to you for free and open access by the SMU Institutes, Centres, Labs & Initiatives at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Centre for AI & Data Governance by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email libIR@smu.edu.sg.

How to address the AI Governance discussion? What can we learn from Singapore's AI strategy?¹

Nydia Remolina²
Josephine Seah³

The following speech was based on a research supported by the National Research Foundation, Prime Minister's Office, Singapore under its Emerging Area Research Project Funding Initiative, and it was delivered by Yihan Goh, Dean, Professor of Law and Director of the Centre for Artificial Intelligence and Data Governance at Singapore Management University, at a conference on AI and Robot Law organised by Keio University with support from the Research Institute of Science and Technology for Society of Japan (RISTEX), in Tokyo, Japan the 6th of July, 2019.

Introduction

1. Good afternoon one and all. Thank you for inviting me here today and for the opportunity to share my perspective of Singapore's experience with AI Governance. AI, machine learning and deep learning are widely emerging to draw the interest of many companies towards the field of automation. Great opportunities for many industries and the economies come from the use of AI. AI and Data are everywhere, AI impacts all industries, changes commercialization and affects society. A report released in 2017 by research firms Accenture Research and Frontier Economics said Singapore could double the size of its economy within 13 years with successful adoption of AI – a growth rate that would take 22 years otherwise. However, as private and public sectors experiment with AI, they are also wrestling with new ethical and legal questions. Do you trust AI to select and purchase groceries for you? How about trusting it to determine the amount of your next pay raise? Or letting it decide whether you're eligible for a loan? Think tanks, research organizations, regulators and policymakers are crafting proposals with the aim to mitigate the risk of harm that these technologies might bring.
2. I will be summarizing Singapore's efforts in these regards, which led the country to win a top award at the World Summit on the Information Society Forum, a United Nations level platform, on April 2019 for its efforts in AI ethics and governance.

¹ This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its Emerging Area Research Project Funding Initiative.

² Research Associate, Singapore Management University's Centre for AI and Data Governance.

³ Research Associate, Singapore Management University's Centre for AI and Data Governance.

The initiatives that contributed to the win included: Asia’s first model AI governance framework that was released in January; an international and industry-led advisory council on the ethical use of AI and data formed last June; and a research programme on the governance of AI and data use established in partnership with the Singapore Management University last September through the Centre for Artificial Intelligence and Data Governance that I lead.

3. I will divide my presentation as it follows: first, I will broadly provide some background to the Singapore’s AI strategy; in the second part I will explain exactly what the Model AI Governance Framework is and how Singapore is addressing the AI Governance discussion; third, I will compare Singapore’s approach with other jurisdictions’ and how important it is to contribute to the international debate on AI governance; and fourth, I will mention some challenges ahead and observations on what is missing from current conversations about AI.

Content

I.	Singapore’s AI Strategy	2
II.	Singapore Model AI Governance Framework.....	5
	II-I. Principles Based Approach.....	5
	II-II. Collaborative Approach.....	7
	II-III. Practical Approach.....	8
	II-IV. Global and international approach.....	8
	II-V. Business and Objective Approach.....	9
III.	Comparison of Singapore’s approach with others	9
IV.	Opportunities and Challenges for the road ahead	11
	IV-I Can our data be secure?	12
	IV-II Who participates in shaping the data-driven economy, and how?.....	14
	IV-III What is missing from current conversations about AI?.....	15
V.	What is AI Governance in practice? (optional)	16

I. Singapore’s AI Strategy

4. Over the past decades, Singapore has become one of the leading jurisdictions pushing for a more digital, more connected, and inclusive society. Digital inclusion is about ensuring all Singaporeans have access to technology that can enhance our everyday lives, and equipping people with the skills and know-how to use technology safely and confidently. This is the vision that underpins Singapore's Smart Nation initiative. Launched in November 2014, Singapore's Smart Nation is envisioned as a Singapore where people will be more empowered to live meaningful and fulfilled lives, enabled seamlessly by technology, and offering exciting opportunities for all. Jumping onto the opportunities afforded by 'Big Data', the initiative committed the Government to making more data available to public and private actors in hopes of collaborative innovations that would propel Singapore's digital transformation.
5. The Smart Nation focuses on 5 key domains: transport, home and environment, business productivity, health and enabled ageing, and finally, public-sector services. Since its launch, the government has been doing its best to stay ahead of the curve. In 2017, Dr Yaacob Ibrahim, the Minister for Communications and Information, identified four frontier technologies that would further the groundwork infrastructure that underpins the country's ambitions for its Digital Economy. Artificial Intelligence (AI) was identified as one of these technologies, along with cybersecurity, immersive media, and the Internet of Things.
6. The Smart Nation Initiative thus guides much of what the country has been doing around AI and robotics. We can see that the goal has been to become a trusted hub for test-bedding, deploying, and scaling up AI solutions in the context of a highly urbanised city. Good governance of AI has become a key issue for the government in ensuring that the technology can be responsibly applied across industries.
7. Singapore's AI Governance strategy, as such, can be seen as one borne from the country's ambitions to be a globally recognised smart city, harnessing its public data into meaningful action for the betterment of public services and infrastructure. The necessity of building a strong, trusted, and collaborative environment thoroughly shapes the country's efforts to boost responsible development of AI.
8. In order to take AI to the next level, it is important to be prepared in terms of capabilities, technology investments, as well as regulatory requirements. That is why Singapore has been working on several AI initiatives, such as:

- a. AI Singapore: set up in May 2017 by a government-wide partnership comprising National Research Foundation (NRF), the Smart Nation and Digital Government Office (SNDGO), the Economic Development Board (EDB), the Infocomm Media Development Authority (IMDA), SGINnovate, and the Integrated Health Information Systems (IHiS). With an up to S\$150 million fund, this initiative is focused on catalysing, synergising and boosting Singapore's AI capabilities in an effort to work with industry to apply AI to industry problem statements and see how to create solutions. In other words, practical AI applications.
- b. An international and industry-led Advisory Council on the Ethical use of AI and Data: Amongst other things, the Advisory Council assists IMDA in engaging stakeholders on issues that support the development of AI governance capabilities and frameworks. The Advisory Council will be able to tap on technical, legal, ethics and international experts from different fields and representatives for consumers and civil society to guide its work.
- c. A Research Programme to explore ethical and legal implications of AI: This was established in September 2018 and is run by the Singapore Management University School of Law. SMU's Centre of AI and Data Governance was formed from a research grant from the NRF and IMDA. was set up to conduct scholarly research on policy, legal, regulatory, governance, ethics and other issues relating to AI and data use. It will support the Advisory Council and inform Government and industry discussion on AI challenges through its research. SMU was awarded a major grant of S\$4.5 million from the National Research Foundation (NRF) and the IMDA to helm this five-year Research Programme.
- d. A Digital Services Laboratory (DSL) to accelerate AI technologies: The DSL team, comprising coders, engineers, data scientists, data architects, technical business development officers and other specialists, will work with government agencies, Institutes of Higher Learning and Research Institutes to leverage their capabilities, and translate them for deployment into industry.
- e. A National Speech Corpus which enables AI speech solutions to work with Singapore's lexicon. It contains 2,000 hours of locally accented audio and corresponding text transcriptions. There are more than 40,000 unique words within the text transcriptions comprising local words such as "Tanjong

Pagar”, “ice kacang”, or “nasi lemak”. The data is made available via the Singapore Open Data Licence.

- f. A multi-pronged training courses to boost AI talent in Singapore: the TechSkills Accelerator (TeSA) programmes will be scaled up to develop more local talent and equip existing IT professionals with in-demand skills. TeSA is a partnership between the Government, industry and the National Trades Union Congress (NTUC) to build and develop a skilled ICT workforce here.
- g. An inter-agency taskforce this year to study how Singapore will develop AI as a strategic capability and become a trusted global hub for test-bedding, deploying, and scaling AI solutions. The Government will also take steps to build know-how in AI to enable workers to use AI tools. This means teaching computational thinking and data literacy in schools, and training adults in data science and AI skills.

II. Singapore Model AI Governance Framework

- 9. The genesis of the Model AI Governance Framework can be traced back to efforts by policy makers and regulators in Singapore to articulate a common AI governance approach and a set of consistent definitions and principles relating to the responsible use of AI, so as to provide greater certainty to industry players and to promote the adoption of AI while ensuring that regulatory imperatives are met. The Singaporean approach seeks to strike a balance between letting innovation happen and identifying, evaluating and proactively managing the risks that arise.
- 10. Launched earlier this year after a 6-month consultation process that kicked off in June 2018, the first edition of this accountability-based Model Framework aims to frame the discussions around the challenges and possible solutions to harnessing AI in a responsible way.
- 11. Now, I would like to mention some of the main characteristics of the Model Framework:

II-1. Principles Based Approach

- 12. The Model Framework aims to collect a set of principles, organise them around key unifying themes, and translate them into some very practical measures in four

specific domains. This is to help guide businesses in deploying responsible AI based on four pillars:

- 13.(i) Internal AI governance structures and measures: The Model Framework espouses a proper governance structure with clear responsibility assigned to individuals when deploying A.I. A clear definition of roles and accountability — be they an algorithm engineer or a businessperson — is fundamental to ensure responsible deployment of A.I. use cases. The organisation's existing internal governance structures can be adapted, and/or new structures can be implemented if necessary. For example, risks associated with the use of AI can be managed within the enterprise risk management structure; ethical considerations can be introduced as corporate values and managed through ethics review boards or similar structures. Organisations should also determine the appropriate features in their internal governance structures. For example, when relying completely on a centralised governance mechanism is not optimal, a de-centralised one could be considered to incorporate ethical considerations into day-to-day decision-making at operational level, if necessary. The sponsorship, support and participation of the organisation's top management and its Board in the organisation's AI governance are crucial.
- 14.(ii) Risk management in autonomous decision-making: identifying consequences and potential impacts of autonomous decision making is key. According to the Model Framework, once the risk assessment is conducted, it will be possible to determine the degree of human involvement needed in the autonomous decision-making process. If the harm caused by a wrong decision is severe, there should be more human involvement and safeguards in the decision-making process. For instance, life-and-death situations should always require a human to make the ultimate call. Inaccuracies and bias in data sets and models could lead to unintended outcomes. Researchers have found that A.I. algorithms working with Western-centric data sets often fail at cross-cultural identification. Worse, they might perpetuate stereotypes and prejudices — for instance, an A.I. may not be able to comprehend the notion that a female president is possible if the data set is comprised exclusively of U.S. presidents. Organizations should strive to address this in their safety measures and deployment strategies. This could be done by building more diverse and representative data sets or tweaking the algorithms to make them more inclusive. This allows organizations to mitigate the risks associated with the use of A.I., which benefits the end users of the technology.
- 15.(iii) Operations management: The Model Framework uses a generalised AI adoption process to describe phases in the deployment of an AI solution by an

organisation. Organisations should be aware that the AI adoption process is not always uni-directional; it is a continuous process of learning. This adoption process includes three stages. First, data preparation; second, application of algorithms; and third, choosing the model. Data is the lifeblood of A.I., and as such data management is a prerequisite in the Model Framework. This means knowing where the data originally came from, how it was collected, curated and moved within the organization, and how its accuracy is maintained over time. On the one hand, organisations should consider measures to enhance the transparency of algorithms found in AI models through concepts of explainability, repeatability and traceability. An algorithm deployed in an AI solution is said to be explainable if how it functions and how it arrives at a particular prediction can be explained. The purpose of being able to explain predictions made by AI is to build understanding and trust. Organisations deploying AI solutions should also incorporate descriptions of the solutions' design and expected behaviour into their product or service description and system technical specifications documentation to demonstrate accountability to individuals and/or regulators. This could also include design decisions in relation to why certain features, attributes or models are selected in place of others.

16. (iv) Customer Relationship Management: As new technologies can be intimidating and alienating to the unacquainted, companies are encouraged to follow two guiding principles. First, decisions made by or with the assistance of A.I. should be explainable, transparent and fair to consumers. Second, A.I. solutions should be human-centric. As A.I. is used to amplify human capabilities, the protection of the interests of human beings, including their well-being and safety, should be the primary considerations in the design, development and deployment of A.I.
17. In sum, the Model Framework seeks to equip its users with the tools to anticipate and eventually overcome AI potential challenges in a practical way. The Model Framework entails tailoring the measures to address the risks identified for the implementing organisation.
18. AI technology is too new and too nascent to have firm rules. The government is tackling AI governance challenges adequately by not setting a prescriptive rule. Surely, the collaborative approach will help regulators to assert when shaping the principles that will govern AI and position Singapore as a leading Digital Economy and Smart Nation.

II-II. Collaborative Approach

19. Singapore's Model AI Governance Framework is fully aligned with the country's Smart Nation ambition, its development emblematic of Singapore's collaborative style of governance. Along these lines, for Singapore it is indispensable to work together with the industry and with a global perspective. In other words, the industry and the government are working together as partners to come up with some of these measures and frameworks. The industry then applies it during the pilot stage and provides feedback to the Government. This is an iterative process that allows regulators to better understand the shape and constraints of this technology and therefore what could eventually become a set of rules and regulations governing them.

II-III. Practical Approach

20. From the advent of the internet to the proliferation of A.I. and big data, legislation has historically lagged behind new technologies. The Model Framework focuses a lot on implementation. Reflecting a rethink of how regulation should look in the face of tech disruption, practicality is the hallmark of the Model Framework.

21. Adopting Singapore Model AI Governance Framework will assist organisations in the complex world of regulatory compliance. The Model Framework does not absolve organisations from compliance with current laws and regulations, but as it is an accountability-based framework, adopting it will assist in demonstrating that organisations had implemented accountability-based practices in data management and protection, as the Personal Data Protection Act and OECD Privacy Principles. This creates an incentive for organisations to participate in the pilot and provide feedback in order to improve the Model.

II-IV. Global and international approach

22. The Model Framework is Singapore's contribution to the global discussion on the ethics of AI by providing a framework that helps translate ethical principles into pragmatic measures that businesses can adopt. The Model Framework was in fact launched during this year's World Economic Forum annual meeting. Davos is a gathering of global industry leaders and leaders from the government. By announcing the Model Framework in Davos, invites global feedback on what Singapore is doing.

23. Additionally, the WEF's Centre for the Fourth Industrial Revolution (C4IR) is a key partner in terms of the development of the framework. More importantly, this is a

way we can propagate the Model Framework to the business community and get feedback on a global basis.

II-V. Business and Objective Approach

24. Singapore is a pro-business jurisdiction. We are also keen to engender a rules-based, norms-based trading and economic environment globally. Therefore, when we propose some of these ideas, they tend to be seen in that context. This philosophy allows policy makers and regulators to think about a more objective cost-benefit in policy making rather than focusing on a particular agenda as opposed to some certain other jurisdictions.
25. This pro-business approach has helped Singapore to design a Model Framework that puts customers first and, simultaneously provides real incentives for businesses to be open. It builds a relationship of trust. Business will have to explain how A.I. technology can successfully deliver results to clients. If not, clients won't use those products. An ecosystem of trust creates a virtuous circle for A.I. innovation. Educated customers are more likely to be receptive, which leads to greater adoption rates, which in turn generates useful information on which companies can build and fine-tune new technologies.

III. Comparison of Singapore's approach with others

26. Countries around the world are in different stages when it comes to AI governance, but we can see from the proliferation of national AI strategies in the past five years, that it is clearly gaining momentum. In the past fifteen months, Canada, China, Denmark, the EU Commission, Finland, France, India, Italy, Japan, Mexico, the Nordic-Baltic region, Singapore, South Korea, Sweden, Taiwan, the United States, and the United Kingdom have all released strategies to promote the use and development of AI. Actually, Japan was the second country to develop a national AI strategy.
27. As the technology becomes more pervasive, so too will the efforts to put governance frameworks around the world. Singapore was the first country in Asia to put together an AI governance framework. However, other jurisdictions are currently discussing AI governance issues and developing AI governance frameworks. For example, the European Commission announced its final set of AI and ethics guidelines by March 2019, an approach likely to complement the EU General Data Protection Regulations. This initiative also includes a Pilot

programme in which companies can implement the guidelines and provide some feedback to the European Commission through an online survey.

28. Other initiatives in the same area have emerged recently. For example, China's Next Generation Artificial Intelligence Development Plan, released in 2017, states the necessity of developing an ethical code of conduct and R&D design for AI products. In December 2018, France and Canada jointly established an International Panel on Artificial Intelligence, seeking to co-lead the AI policymaking scene among G7 countries. Also, On February 2019, the President of the United States announced the American AI Initiative, which is a concerted effort to promote and protect national AI technology and innovation. The Initiative implements a whole-of-government strategy in collaboration and engagement with the private sector, academia, the public, and like-minded international partners. On a more international scale, the OECD presented on May 2019 a set of principles on AI to promote the innovative and trustworthy use of AI that respects human rights and democratic values.

29. Even though Singapore's Model Framework has unique characteristics, as I mentioned already, it also certainly shares some similarities with other frameworks and models that have been released.

- a. First, all these models are part of a broader push towards having a national strategy on AI.
- b. Second, all frameworks converge on similar principles of *human-centricity*, recognising that decisions made or assisted by AI should be explainable, transparent, and fair to users.
- c. Third, all models share an interest with regards to pushing towards a global discussion of AI. All jurisdictions are aware that AI governance issues need a multi-jurisdictional approach, even if it is not possible to create a global standard. The models share a common interest in engaging in the international debate on AI governance and deepening international cooperation in the field.
- d. Fourth, all the initiatives are functionally algorithm-agnostic, technology-agnostic, and sector-agnostic. Singapore's Model Framework, for example, does not focus on specific AI or data analytics methodology, neither does it focus on specific systems, software or technology. It applies regardless of development language and data storage method and is meant to be a

baseline set of considerations and measures for organisations operating in any sector to adopt.

- e. Fifth, all jurisdictions try to integrate efforts from public sectors, policymakers, academia, and the industry. In Singapore, the Model Framework joins other AI initiatives that follow such a strategy of pulling in experts from across public sectors and industry experts. These same patterns of collaboration can similarly be found, at different levels, in the EU, US, and China.
- f. Sixth, as of now, all models and initiatives spearheaded by national governments are neither rules nor regulations. Although this is likely to change in the years ahead as we are better able to understand the long-term social consequences of using AI.

IV. Opportunities and Challenges for the road ahead

30. Putting these initiatives into practice, I would like to highlight some ways in which AI is being folded into the lives of everyday Singaporeans. For that, we can look at the country's efforts towards improving the public transport system, headed by the Land Transport Authority.

31. Public Transport is one of the 5 domains of Singapore's Smart Nation initiative. Here, data analytics, smart systems, and autonomous vehicles have been highlighted as key solutions for the future of transport planning and operations. The plan is for roads and transport systems to be optimised: traffic will be smoother, public transport will be more comfortable and reliable, and in turn, with less reliance on private cars, our air will be cleaner.

32. To accomplish these goals, LTA and A*Star, the government's R&D agency, launched a joint partnership to oversee and manage autonomous vehicle research, test-bedding, and development of applications by industry partners and stakeholders. Since 2015, LTA has been facilitating AV trials by various technology developers in designated locations around the Island, with the target of serving residents in 3 neighborhoods by 2022.

33. On top of their own efforts, LTA is also working closely with industry leaders to develop transport solutions. The organisation has partnered with IBM to implement a programme that combines video data at selected train stations, along with telecommunications travel data from a local telco, with the aim of crowd

management at train stations across the country. With the predictions developed by IBM's programme, LTA's traffic controllers were able to anticipate and better manage the flow of traffic to prevent the build-up of congestion.

34. Similarly, LTA has also been able to leverage the wisdom of the crowd through crowd-sourcing techniques. That is, commuters are able to indicate through an app their preferred commuting routes, and bus operators respond by providing the necessary shuttle bus services. The app thus enables commuters to be able to participate in the route designing process, and bus operators are able to tailor demands to meet these needs.
35. Thus, we can see not just how the Model Framework might be expected to fit into the larger process of developing the Smart Nation, but also Singapore's distinctive style of governance through collaboration, experimentation, and innovation. Such collaboration ensures that technologies like AI, robotics, and data analytics are used for social, public goods. The government improves key and core public services, businesses develop systems that are human-centric and accountable, and underneath all these advancements, members of the public are able to build trust in these systems.
36. There is no doubt of the benefits that the use of AI could bring to society. However, the greatness of these benefits will depend on how we tackle AI challenges within a framework of responsible use of AI. According to a survey conducted by the MIT Technology Review, 37% of AI industry participants believe that the Asia Pacific will lead in the development of ethics and government frameworks (compared to 36% who identified Europe as the leader, and 22% who say it will be the US). Singapore took a first step in the region with the Model Framework. However, is this enough?
37. As academics we should contribute to the global development of AI Governance, resolving—in some cases preventing—the unintended consequences and social harms that these emerging technologies are likely to generate. I want to finish my presentation with three challenges - ranging from technical to social - that we should be prepared to confront:

IV-I Can our data be secure?

38. The first challenge raises a technical challenge that is not solely confined to Singapore's shores: can our data really be secure?

39. By and large, research indicates that cybercrime is on the rise — news headlines support these findings as major companies like Marriott, Equifax, Yahoo, and Facebook find themselves in the crosshairs of cyber-attacks. The cybercrime economy has grown to enjoy at least \$1.5 trillion in profits each year. The value of the cyber security market is anticipated to reach \$300 billion by 2024, according to a 2019 press release by Global Market Insights, Inc. The frequency in which Cybersecurity Ventures predicts that a business will fall victim to a ransomware attack this year in its 2019 Official Annual Cybercrime Report (ACR). The company also estimates that number will increase to every 11 seconds by 2021.
40. Recently, Singapore suffered two high profile data breaches in its healthcare sector. In one of them, hackers infiltrated the databases of SingHealth, Singapore's largest group of healthcare institutions that collects data from polyclinics, public hospitals, and outpatient clinics. The personal particulars of 1.5 million patients, including the outpatient prescriptions of our Prime Minister, were compromised.
41. In the aftermath of the attack, a national inquiry revealed that the hackers were particularly skilled and most likely state-sponsored, with a clear goal of gaining the personal and outpatient medication data of the Prime Minister. Their tactics bore the characteristics of an Advanced Persistent Threat group, as seen from the suite of advanced, customised, and stealthy malware used; as well as their ability to find and exploit vulnerabilities in the IT networks and systems in question.
42. On the other hand, the inquiry also found a number of vulnerabilities, weaknesses, and misconfigurations in the IT network and system in question that contributed to the hackers' success in obtaining the data, many of which might have been identified and remedied before the attack. For example, the inquiry revealed a lapse in key personnel, where relevant staffers did not have the adequate levels of cybersecurity awareness, training, and resources to understand the security implications of their findings, or to respond effectively to the attack.
43. In the coming years, we will need to be prepared for other, even more highly sophisticated forms of attack on systems that are in the progress of being even more thoroughly linked together.
44. Singapore has been taking steps towards acknowledging the importance of these developments. This year, 'digital defence' was introduced as the sixth pillar in the nation's Total Defense Framework, highlighting the importance of being aware of and prepared for cyber-attacks. The pillar joins 5 others: military, civil, economic, social, and psychological defences. In addition, the Cybersecurity Act came into

force last year, establishing a legal framework for the oversight and maintenance of national cybersecurity in Singapore.

45. It remains to be seen whether these will be sufficient. As the saying goes - and particularly true for the security of a network - we are only as strong as our weakest link.

IV-II Who participates in shaping the data-driven economy, and how?

46. The second challenge is a social question at the heart of unpacking what it means to participate in shaping a data-driven economy.

47. Crowdsourcing mechanisms, such as those employed by the LTA, offer a simple example of a virtuous and constructive feedback loop enabled by a vastly improved communications infrastructure. People were able to indicate a demand for a particular route, and bus operators were able to respond accordingly.

48. But what about other forms of participation? Singapore's data portal, launched in 2011, hosts publicly available datasets from across 70 public agencies. The portal invites citizens to use these datasets to "hack the city" - that is, to create applications for urban problems.

49. At a foundational level, a question can be raised about who has the skills and capacities to make use of such a portal. Often, we see that the technical literacy levels and financial resources required to meet these expectations are much smaller among individual citizens or civil society groups, as compared to larger corporations. It is insufficient to merely *have* an open data portal, if we fail to similarly democratise the skills required to leverage such data. Educational opportunities thus need to be built into our institutions, ensuring that everyone, across their lifetime, can have the opportunity to pick up these skills and use these datasets to answer questions they wish to ask.

50. In a world of rapid technological advancements, urban problems will be best addressed by individuals who are representative of the diversity of life in the city. The challenge lies not only in education, but also in ensuring that this diversity is cultivated, encouraged, and respected. Singapore, for example, has a near equal proportion of males and females in its population, yet gender representation in STEM fields remains largely disproportionate.

51. The premise of participation in the smart city and the deployment of AI is one that relies on both explicit participation - like developing hard and software to be used in the city - and participation that is more implicit: drawn from citizen's willingness to live in a sensor-enhanced environment. In Singapore, a recent 'smart home' trial was conducted where sensors were installing on doors in homes so as to monitor the activities of elderly residents. Still, residents disliked being seen as vulnerable and in need of aid - some even covering their sensors with towels. In the years ahead, then, both these questions of explicit and implicit participation - so crucial to achieving the ambitions of Singapore's smart city - must be handled with care and wisdom.

IV-III What is missing from current conversations about AI?

52. These observations and speculations thus bring me to my final question: what is missing from our current discussions about AI?

53. Undoubtedly, there has been a rush towards adopting AI and embracing the promises of a cutting-edge technology. Still, we are only beginning to understand the potential harms being built into these systems. It is now widely acknowledged, for example, that facial recognition technologies can discriminate due to choices made in the selection of their training data. The truth is that no matter how good governance frameworks may be, no matter how well intentioned the development of technology, there remains opportunities for its misuse and abuse.

54. In addition, good AI Governance may not be enough to solve some deeper, more entrenched issues. The country's embrace of AI is not solely motivated to remain ahead of the curve as a knowledge and information hub, but also in response to projected labor shortages. Like Japan, Singapore also has an ageing population. By one estimate, by 2030, one out of every four Singaporeans will be 65 years and older. In another, by 2050, Singapore will be the fifth oldest and least fertile country in the world. These demographic trends make up the spectre haunting our labour force and market projections, leading to much anxiety of the future of the country and its economy. One response to these anxieties can be seen in the country's enthusiastic embrace of autonomous vehicles: In addition to LTA's move to enhance public transport, strides are similarly being made towards driverless campus busses, truck platooning systems, and unmanned road sweepers.

55. But there are larger questions on the horizon. With these trends in mind, we can go a step further and ask: how might AI transform life? As transport improves with autonomous buses, fleet management efforts respond to traffic congestion,

telemedicine gains public trust and favour, and AI-enabled precision medicine and personalised healthcare takes off, the population appears poised towards becoming even healthier and ever more mobile while work and leisure adapts to automation and digitisation. In the near future then, patterns of life in Singapore and in the world are going to change. Cultural and generational expectations of work and retirement and changing social values, may mix in a potpourri of complicated, perhaps contradictory, expectations. Good AI governance may hold the promise of living *longer*, but it cannot answer the question of what it means to live more *meaningfully*.

56. Questions like these are rarely broached about technology policy, much less our discussions about where AI may be headed, but nonetheless are present in the very decisions that shape our deployment of AI technologies, not only in Singapore, but also around the world.

57. We are definitely living exciting times in which academia is expected to play an important role. We have a duty to contribute to the international debate on AI governance and to tackle the challenges associated with the fast-evolving world of AI, in the hope that we can start to provide answers and solutions to some of the difficult ethical, legal and societal questions the technology poses.

V. What is AI Governance in practice? (optional)

58. It is not easy to put AI governance in practice. However, Singapore's Model Framework provides an example of a company that already implemented it so business can see in practice what AI Governance should look like. The use case is from the healthcare industry.

59. UCARE.AI deploys a suite of AI and machine learning algorithms, including proprietary deep learning and neural network algorithms, built on a cloud-based microservices architecture to provide sustainable and customisable healthcare solutions for doctors, hospitals, patients, insurers and pharmaceutical companies. A successful use case is the recent implementation of AI-Powered Pre-Admission Cost of Hospitalization Estimation (APACHE) for four major hospitals in Singapore. APACHE provides more accurate estimates, with a fourfold improvement in accuracy over Parkway Pantai's previous bill estimation system. This is done with the intent of achieving standardisation of healthcare cost estimation and provision of greater price transparency to facilitate the building and maintenance of trust between payers, providers, and patients. Before APACHE, healthcare cost estimation methods involve traditional techniques such as simple statistical

aggregations. Challenges of these traditional method include relatively high error rates, high financial and human cost of updates, and low frequency of updates due to these high costs. UCARE.AI worked with Parkway to resolve these issues with a multi-step process involving: (i) data exploration, (ii) data cleaning, (iii) feasibility assessment (iv) feature engineering, (v) machine learning, and (vi) presentation of results. With satisfactory results from the proof of concept, APACHE was then put into production. But, what happens if it fails? What is the probability of a failure and its potential impact? How to ensure a good government of this AI use case?

60. UCARE.AI adopts a proactive approach that aligns with Singapore's Model AI Governance Framework, as follows:

61. Trustworthy and Verifiable: The proposed AI governing framework acknowledges that neural networks are inscrutable, and verification of the results provided by such networks is required prior to putting them to use in human applications. UCARE.AI circumvents this problem by continuously validating the accuracy of its algorithms against the ground truth. Weekly check-ins with participating partners and domain experts are also employed to ensure quicker and more reliable iterations. Automated re-training of the data models ensure that the algorithms remain up to date. This methodology of continuous validation of its AI models with the help of experts from the hospitals will help to boost confidence in the accuracy of its predictive insights and will help train algorithms to become even more precise with each amount of data inputted.

62. Accountability and Transparency: Prior to data collection, informed consent from stakeholders would have been obtained and approval of the use of data sought via open communication channels. The careful curating and conversion of data into usable format prior to building the models ensures the AI algorithm is kept accountable and coherent to users; this is done in conjunction with the hospitals. The proper storage and repair of previously broken or missing data also serve to provide greater transparency and safety to users by minimising the influence of data gaps in the projection of the result. Careful monitoring of data is key in ensuring service reliability, and therefore detailed and consistent logging across the multiple components involved is also employed in APACHE, collected in a secure, centralised log storage that is made easily accessible to the development and operations team when required, allowing for prompt debugging and uptime tracking if necessary.

63. Fairness: The automated prediction of hospitalisation costs reduces the likelihood of human biases affecting the ultimate judgement of the data and provides an element of consistency across all predictions. Discrimination based on income

levels and insurance coverage, for instance, would be effectively negated. Although there would be concerns about the use of a 'human out-of-the-loop' system, the algorithm in question is designed to be human-centric.

64. Human-Centric: This use case highlights how artificial intelligence may be used in augmenting decision-making capabilities in a human-centric manner whilst minimising the potential risks of harm to involved parties. The automated process of bill estimation negates the need for tedious statistical calculations, thereby freeing up man-hours and effort to allow for the channeling of these into more creative pursuits. Furthermore, the information provided would serve to benefit patients and payers by allowing for more accurate cost forecasting, efficient allocation and distribution of healthcare resources, and guidance on new policy initiatives. Patients would be conferred greater peace of mind over their healthcare expenditure such that they may focus their energies on recovery instead. To minimise the risk of harm, rigorous feasibility studies are conducted prior to using the data to focus on creating a valid and robust validation framework. This will be done in conjunction with partners and their feedback on the proposed framework obtained before proceeding. A human feedback loop with inputs from the client organisation (all four hospitals, which are owned by the same company) is also in-built into each algorithm to enhance sophistication, while a manual override protocol is also included to ensure that these algorithms can be safely terminated if deemed necessary. This ensures that the algorithm remains under human control and in line with the medical field's well-established ethical principles of beneficence, non-maleficence, and social justice.