Asian Dialogue on AI Governance - report

Table of Contents

Asian Dialogue on AI Governance 1

Session 1: The notion of AI Governance – industry, regulatory, and academic perspective 1

Session 2: Decoding the elements of AI Governance: Explainability, Ethics, Fairness and Transparency 3

Session 3: Human Interaction with AI 4

Session 4: The responsible use of AI: accountability and liability related issues 6

Session 5: Data privacy and data protection's role in AI Governance 7

Final Remarks 10

Session 1: The notion of AI Governance – industry, regulatory, and academic perspectives

Yeong Zee Kin, Deputy Commissioner, PDPC
Marcus Bartley Johns, Regional Director, Government Affairs and Public Policy, Microsoft
Mark Findlay, Professorial Research Fellow, SMU
Moderator: Jolyon Ford, Associate Professor, Australian National University

The first session was aimed at laying out the groundwork for a broader debate on the concept of 'AI governance' through highlighting three different perspectives: regulators, industry professionals, and academics. Zee Kin initiated the discussion with a presentation on how Singapore has been supporting AI-development in the country through encouraging consumer trust and good corporate behaviour. Too early for legislation, the approach has been led by opening up conversations across three groups: industry; workers and consumers; and researchers. Concrete steps taken so far have been the setting up of an Advisory Council on Ethical Use of AI and Data; a research programme (CAIDG); and the publication of the Model AI Governance Framework. Finally, Zee Kin introduced new initiatives that are currently taking shape: firstly, regulators are looking into training and certification programmes for professionals who will be implementing AI solutions. Secondly, in order to understand the changing nature of work, the Lee Kuan Yew Centre for Innovative Cities has been conducting research in the interaction between work, employment and AI.

Following from this, Marcus provided an industry perspective on 'AI Governance': stressing that Microsoft's perspective was shaped by its development and embedding of AI in their services. He remarked on the increasing consensus on developing principles for AI adoption – and on a convergence of principles across the field. Microsoft has issued their own set of principles, which includes fairness, reliability and safety, privacy and security, inclusiveness; with transparency and accountability being foundational principles across all four. The challenge, nonetheless, lies in putting these principles into practice: there is a growing gap between technical discussions on AI and what citizens' views of AI are – a gap that should be closed. There have already been a number of attempts to translate these principles into practice: such as legal and regulatory frameworks, governance frameworks, standards and good

practices, and technical methods and tools. Still, use-case contextualization remains essential across all frameworks. For example, considerations for an AI-assisted consumer lending programme will be wildly different from one designed to suggest convenient meeting times. We should recognise when and where existing frameworks have something to offer for the application of principles into practice – and where legal gaps have appeared to render frameworks inadequate for current technology. Contract and criminal laws, for example, have traditionally been used to assign responsibility for actions, which may in turn help us in operationalizing 'accountability'. On the other hand, when it comes to facial recognition technology there may be more legal gaps in existing frameworks. Finally, Marcus concluded his presentation by touching on the role of law, he suggested that it would be important to start with the most sensitive use-cases – in areas of significant harm – to identify existing legal and governance gaps. Pilots and iterative processes – a feature of Singapore's approach – has been useful for evaluating processes, but research still needs to be done to develop more contextual understanding across different use-cases in order for us to apply principles into effective practices.

Mark brought the academic's perspective into this discussion: he made three observations in response to the two preceding presenters. Firstly, that the idea that 'relevant research is essential' is perhaps not a unanimous consensus – researchers consciously need to make the case for it. He noted the emergence of authoritarian populist governments which are increasingly moving away from evidence-based policies. Regulators and legal researchers thus face an opposition from people in power who do not care for good evidence. Secondly, a robust legal framework is required: governance manifests itself through forms of regulation, as such we need frameworks that make sense and that people can understand. The question becomes one of making governance through regulation seem real to people who are scared and frightened. He suggested researchers talk to concerned individuals how one might govern AI, rather than to perpetuate the myth of being governed *by* AI. Nonetheless, regulatory frameworks are often sluggish, oftentimes built on a resistance to understanding technological changes. Regulators are more interested in economic innovation rather than regulatory safety or prudence, and so we have ended up in a strange landscape of regulation: with outmoded notions of legislative controls (e.g., data protection, privacy protections) that have not been sufficiently modified to come close to where AI sits. As important as the law is, he suggests that it might be becoming a mask. Thirdly, he commented that this regulatory framework ought to be pitched at the demographic poised to make the most use of AI. He argued that people in different places see things differently – an Asian approach to AI might be vastly different from a Western approach based on rights, and so we need to recognize such differences in order for regulatory frameworks to be pitched at a language that people can identify with.

The discussion following these presentations touched on a couple of their underlying themes. While part of the discussion was about whether it is essential to develop different governance systems for public and private sectors, a larger part of the discussion was dedicated to questions around trust and developing accountable systems. David highlighted that there is a growing concern about data use outside the context in which it was originally provided: simply enabled by a 'Terms of Use' clause that people do not read. Brian suggested – while acknowledging its imperfections – that there is scope to learn from established accountability mechanisms already in place in the financial sector. Questions were raised about whether there might be nuances in the Asian context that might lead to differences in governance expectations compared to Western domains: humans need to remain responsible for decisions, but the implementation of that responsibility has varied meanings and expectations. For example, in some countries people may have different expectations of AI services: if, due to institutional corruption and

structural racism, people feel that they cannot trust each other, then AI's utopian promise of neutrality may lead to a naïve faith in a machine. Another question raised was that of how to embed a culture of accountability and whether that might be engrained in curriculums in universities, particularly for software developers.

Session 2: Decoding the elements of AI Governance: Explainability, Ethics, Fairness and Transparency

Yong Lim, Associate Dean for Student Affairs, Associate Professor, Seoul National University School of law. Co-Director, SNU AI Policy Initiative.
Moderator: Su Jiang, Associate Professor of Law, Peking University

In session two, Yong Lim discussed some observations about current steps towards building trustworthy AI systems. He suggested that if we came close to having a system that adhered to common AI principles – e.g., transparency, accountability, fairness – that that may be the equivalent to creating a saint-like system, in other words: virtually impossible. The challenge is transforming principles into concrete policies and norms, which requires an articulation of shared goals. We typically have similar expectations to AI creators and developers, Yong Lim flips the question around to ask: what if developers wanted 'explainable' AI law? What if they wanted explainable and encodable standards for them to create trustworthy systems? For example, 'fairness' is both a legal term in Korea and a commonly articulated AI principle – yet when it comes to articulating an encodable concept, there are multitudes of competing and malleable conceptions of fairness, each of which have certain trade-offs. Does this mean that trustworthy systems are an unattainable goal? If we acknowledge that it is extremely problematic to put into practice, and for courts to assess whether a system has met requirements of fairness, should we then say that for certain decision processes – an AI system should not be used?

To address this question, Yong Lim suggested pulling on lessons from existing human systems. He observed that decisions made by judges or board of directors are often not fully explained: judges write decisions but often do not fully articulate their reasoning – and yet we still trust them. He suggested that despite it being not possible to fully trust these systems, we accept their outcomes because they have been designed for failures: judges need to have specific qualifications in order to be appointed, their judgements are rendered in writing, and there is an appeals processes for rooting out errors. Linking back to questions of trustworthiness, Yong Lim argued against the tendency to think of a trait-based approach to AI systems (i.e., we want systems to be fair, just, and non-discriminatory): to do that we need to accomplish and succeed in the difficult task of teaching AI systems what these concepts mean. Instead, he suggested that we should put in place mechanisms for preventing undesirable outcomes by pursuing 'acceptability': is an AI system acceptable as being trustworthy?

He suggested that we need to start accepting more probabilistic outcomes and inevitable errors are part of the package of AI systems. Part of this is also developing an understanding of the margins of error that are acceptable to people. Finally, he concluded his presentation by suggesting possible mechanisms that might increase the acceptability of a system – integrated testing and verification processes (i.e., checks and balances within the system); qualification requirements for both creators and operators; and periodic or random reassessments by auditors.

Mark commented on two overlapping questions raised by the presentation: the first was that of AI introduced into legal decision making, while the other was about the normative framework of law being introduced in AI. On the latter, he observed that the paradox of law has always

been that its normative framework rarely ever happens in practice – nonetheless, its normative framework may not impact on how AI operators. The fairness debate, he observed, occurred in the fringes. In the courts, 'fairness' is oftentimes a concept that is taken as understood and thus unquestioned. In addition, designers often think about law as a science of certainty – and so what we need to develop is an open discourse between designers and law so that the former may come to understand that the law is often deeply discretionary. As such, we need to have realistic expectations of where AI will aid in decision-makings, and where that might not be possible. Responding to these comments on discretionary systems, Yong Lim highlighted that while we may not end up with AI judges – evidence suggests the increasing use of AI-assisted judgements or factual analyses. For example, states in the US have started using algorithms to aid their bidding processes for procurement.

Other participants pushed the concept of 'acceptability' further: Marcus suggested that many questions about what might be 'acceptable' look different according to who and where they are placed in the chain which might implicate what they understand about the system in the first place. Arisa highlighted the example of her work with doctors and medical applications, and observed that some doctors are themselves hesitant to use AI-assisted applications because of their own professional pride. Acceptability, she suggested, needs to be a question not only levelled at end-users but also professionals using and interacting with these systems who may be worried about their own employability. Malavika commented that fairness cannot merely be seen as a property of an AI/ML system – doing so fails to recognise the spectrum of our expectations where trust of a property of social systems and systems of justice. Referencing boyd's work on the topic, she explained that some of the things that we value are about visuals and performance – rather than technical decisions of fairness. We need to be careful about abstracting away the social context when technology is introduced. Finally, Nydia commented on the differences in 'acceptability' that might come from contrasting different sectors. Contrasting the health and financial sector, for example, she suggested that both would rely on different standards of 'acceptability'. In addition, she remarked that a common recurring theme that can be observed here is that through contrasting existing human-led systems (made by judges and board of directors) with AI systems, we might develop much needed insights into when and why we expect different – and often higher – standards from AI systems.

Session 3: Human Interaction with AI

Arisa Ema and Takaski Matsumoto, The University of Tokyo
Moderator: Brian Tang, Executive Director, Hong Kong University's LITE Lab

In the third session, Arisa and Takaski presented on their research in Japan around human interactions with AI. Innovation and research in Japan needs to be set against current societal challenges: a super-aging society and the risk of economic stagnation. The question facing innovation and research, as such, is how to accomplish these while being cognizant of and addressing ethical challenges. The government has been a major driver for pushing out AI principles in Japan, with the Ministry of Internal Affairs and Communication starting to look into AI principles in 2016 and culminating in the social principles of human-centric AI released in 2019. In addition to government initiatives, academic researchers have also expressed their interest and concerns around the developing of systems that reflect certain principles, recently releasing a 'statement of machine learning and fairness'. In addition to these, major industry players have also launched initiatives to address ethical, legal, and social considerations.

Based on their research on roughly 30 documents on AI principles, Arisa and Takaski suggested that we can conceptualize Trustworthy AI as a structure and break it down into its

constituent parts. They identified three different areas: the AI system, the service providers, and the end-users, and further broke each area into smaller sections.

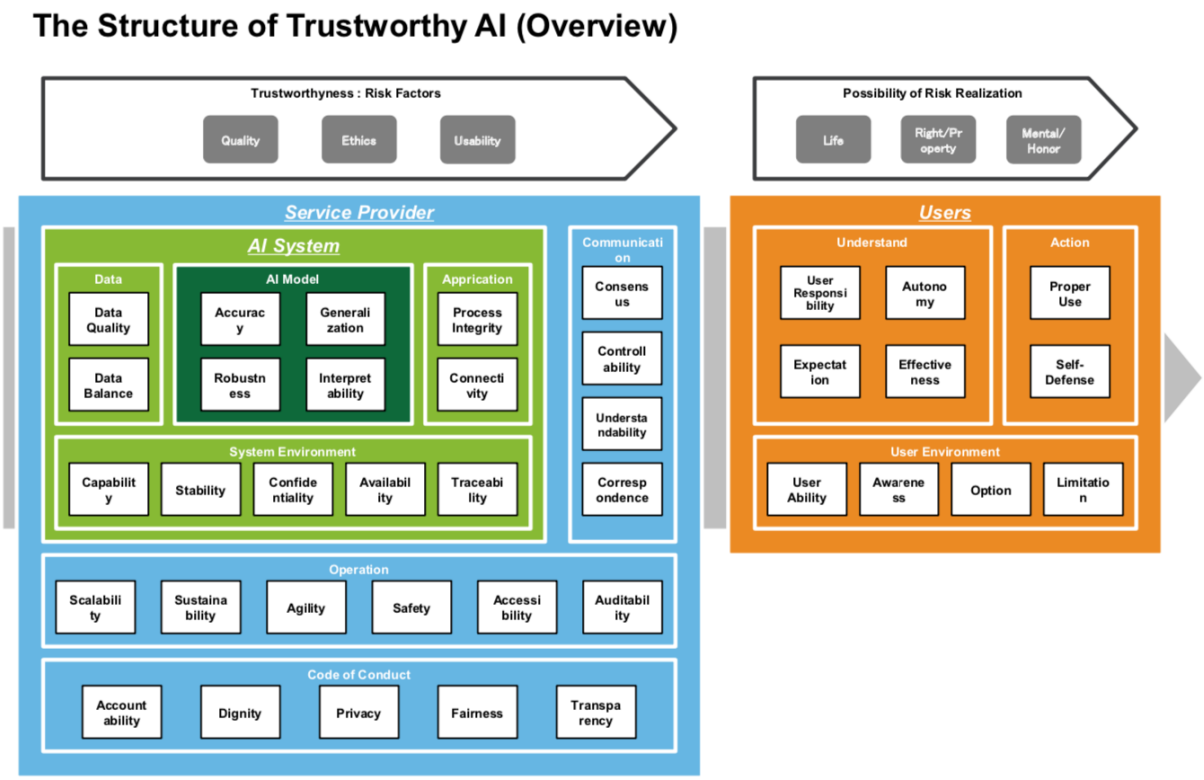

Figure 1. By Arisa Ema & Takashi Matsumoto

In the case of a loan screening AI, for example, considerations of explainability; fairness; robustness and safety; and dignity and authority may be especially salient. Looking at explainability as an example, one risk may be that users might not use the application if they are unconvinced by the reliability of an AI's results. One might then use the structure to mitigate this risk by identifying key components that need to be addressed in order to realise the concept of explainability. For example, in the first area – AI system – questions of interpretability of the model could be addressed by ensuring that information is provided about the reasons certain decisions were made by an AI. In the second area – service provider – a code of conduct might highlight the service provider's responsibility to explain the outcome of each decision; part of their responsibility also involves communication of what the system is doing to their end users. Finally, in the third area – the end-users – end-users should have the ability to understand (through digital literacy programmes, for example) an AI's outcome *and* be aware of options for action, such as protesting or challenging the outcome of an AI-enabled decision.

In the discussion that followed, Brian commented that the structure provided a way to think about the different ways in which we might think about the 'human in the loop'. He shared 6 categories and invited participants to use these as a lens to comment further on the structure presented. The 6 categories were: 1) human as AI trainers; 2) human as user and trainer; 3) human as AI-quality controller; 4) human as AI-explainer or interpreter; 5) human as AI-creator; and 6) human as the AI customer or user, who stands at an arm's length from the system.

Mark commented that the model was a good attempt to particularize the ways in which AI principles might connect with different parties in a decision-making process. The structure thus gives a good spotlight on the different relationships and different areas of responsibilities that are all part of that process. It might also be a good starting point for thinking about standardization initiatives. He also suggested that similar questions from big data being replicated in the discussion about AI systems. That is, with big data it is currently difficult to find pathways from source to use. In AI systems, a similar problem emerges between the model development and its eventual application: the structure provides good governance pathways that might help with resolving the question of validity. Responding to this, Arisa suggested that Japan has many stakeholders – government bodies, start-ups, and big companies – and that this model might assure each different stakeholder of what steps might be necessary to create trustworthy systems. She also agreed with Mark that it is remains difficult to link service providers and users together – while there is general consensus that AI systems ought to be explainable, it remains an option question of who should be explaining them, how they ought to be explained, and how bridges between people might be created.

Nydia also suggested that one might think about the regulators' role in this structure, and whether expectations of how the structure might be used may change based on the regulators' sophistication of understanding. Malavika commented that the framework would also help in understanding where things failed or where systems have broken down, potentially helping companies identify where they may need to invest more time and resources. These in turn may help in thinking about interventions to build more 'humans in the loop' mechanisms.

Session 4: The responsible use of AI: accountability and liability related issues

Guobin Cui, Associate Professor, Tsinghua University
Moderator: Chen Siyuan, Associate Professor, SMU

Next, Guobin Cui touched on liability regimes for AI systems. Current product liability law (PLL) recognizes 3 types of defects: manufacturing defects, design defects, and warning defects. Nonetheless, under current law, software is not addressed as a product under PLL. He identified a few reasons why: software is typically embedded in hardware – the latter of which falls under PLL. Often when separated from hardware, software is unlikely to cause 'physical harm' as addressed in tort law. In addition, contract law has been traditionally used to regulate software. Nonetheless, he suggests that in the future it might be the case to for software to be recognized independently as a product under PLL. Particularly for autonomous vehicles, it is currently common practice for software companies to license their driving systems to different car manufacturers. In the future, as such, it might be possible to distinguish software as a traditional product.

Guobin recognized that some challenges would arise if this were to happen. One, for example, might be determining design defects under PLL. While one method of addressing this has been the consumer expectation test. Nonetheless, it is also possible that the newness of autonomous vehicles makes it difficult to know what these expectations are: such as how to know how safe a vehicle should be. An alternative is to utilize the risk utility analysis, but here particular challenges rise around information costs and the courts ability to understand the technologies well enough to make educated decisions.

Recognising that there will be many difficulties in proving software defects, Guobin suggested that these are not necessarily wholly new challenges: all high tech products (new drugs, aircrafts, medical devices, etc) face similar problems. The challenge we face, he suggests, is

not so new that an overhaul of the system is required. Since it is impossible for consumers to fully understand defects, he thus suggests moving towards a strict liability system where consumers do not have to prove defects in the event of an accident. A strict liability regime might end up saving administrative costs, spread losses, reduce accidents, and ensure victim compensation. On the other hand, Guobin also recognized that a strict liability regime that doesn't require proof of design defects is difficult in reality where majority of car accidents continued to be caused by human error rather than manufacturing or design defects. Rather than the advantages listed above, possible results of shifting to such a regime might mean more careless drivers and accidents and higher prices for products. One approach towards addressing these challenges has been Bryan Choi's suggestion of creating a standard for "Crashworthy Code". In addition, tort law systems are only one tool – other areas worth looking into include government regulation, insurance policies, contract law, or market reputation mechanisms. Finally, Goubin concluded that these liability regimes would ultimately differ based on levels of risk: for example, a high-risk system (e.g., airplanes and cars) might require strict liability; while such a regime for a computer OS might be less helpful.

Siyuan kicked off the discussion that followed by commenting on areas of contrast in the discussion on regulating liability in AVs. New Zealand, for example, are extending their current no fault liability regime to include AVs. Another jurisdiction of interest might also be the United Kingdom where the Automated and Electric Vehicles Act of 2018 has left insurers in charge of suing parties who might be at fault for the accident. Both jurisdictions are thus preempting evidential difficulties in AVs, lessening the questions that need to be answered before compensation can be offered to affected parties. Nonetheless, he also noted that if one has an impulse to know *why* an accident has occurred, both schemes make it difficult to prioritize answers to that question. Shifting from AVs to killer robots, on the other hand, Siyhan noted that thinking about liability might not fully address the challenges at hand. If a robot ends up killing someone, will we be happy so long as it was not a wrongful death? Is that acceptable? There seems to be a different expectation for responsibility and liability based on the context in which these software are used and deployed. Warren noted that a similar issue occurs when an AI system that might be capable of making highly accurate decisions compared to a human continues to be rejected by human beings. One question that was raised here is whether or not this may be a generational issue – we may not be willing to accept farming out responsibilities to an AI, but younger generations who grow up with AI/ML systems may come to think less of these choices and might be happy to delegate presently contentious decisions to a software. Warren suggested that this was a question that is evolving beyond the law. Other participants also discussed other mechanisms for thinking about liability: Yong Lim suggested mandatory insurance for users of AVs, as well as manufacturers and operators. Brian commented on the open-source nature of AI models, to which Guobin noted that the regulation of liability between intermediaries was also another open question.

Session 5: Data privacy and data protection's role in AI Governance

Smitha Prasad, Centre for Communication Governance at National Law University Delhi
Moderator: Warren Chik, Associate Professor, SMU. Deputy Director, SMU Centre for AI and Data Governance

Smitha Prasad rounded out the presentations with a discussion the intersection between data protection regimes and AI governance initiatives. She noted that data protection laws already have a set of basic principles (lawfulness, fairness, transparency, purpose limitation, data minimization) that have become foundational in international efforts of AI governance models and frameworks. In practice, she suggested that we have seen this in the efforts towards

developing 'privacy by design'; in the institutionalization of data protection impact assessments that need to be completed before the implementation of new processing systems; and in the development of audits. From these observations, she raised three challenges facing AI governance: firstly, there remains an over-reliance on the idea of a competent regulator. Noting that this is rarely the case, the challenge becomes who has the capacity to apply these principles in the governance of AI. For some countries, that job has been the task of data protection regulators, but for countries that do not have these regulators, or regulators that are skilled enough to do this adequately, it remains open question. Secondly, data protection laws currently are weakened by the assumption of meaningful consent or control by the data subject – a problem that becomes exacerbated by the introduction of AI/ML models. Thirdly, it remains another open question of who gets to define harms, and who gets to expand existing definitions of harm to account for algorithmic uses of personal data, or consequences that are not necessarily directly related to personal data.

The introduction of AI complicates these questions even further. In the case of AI-related policy-making, for example, government policies have already started incorporating AI for welfare, digital governance, and smart cities. These actions have been propelled by the thinking that big tech might solve all problems of governance. This optimism is also driving the push for applying technological solutions to nation building politics. In India, for example, the push towards supporting local businesses and start-ups has led to a strong push for data localization, forcing data transfers from big tech to start-ups in India. Individual data, as such, has become a natural resource in a larger infrastructural project to build the Indian economy. Issues here reflect similar challenges that have propelled or shaped data protection regimes: such as data processing, surveillance, and who get to have access to data collected by private and public organisations. Finally, Smitha observed that that there is currently a race towards defining one's place in the international AI ecosystem. Countries like Singapore, the European Union, and the United Kingdom all having figured out what they want to accomplish in the foreseeable future. Nonetheless, the rush to take a stand has also led to a number of questionable policies; and she concluded by noting that while India has yet to define their role in this ecosystem, the larger question might be whether such a definition is even necessary to begin with.

The discussion that followed delved into the differences in experiences across the Asian countries represented at the roundtable. Warren initiating it by talking about Singapore's experience with its data protection regime and its overlap with its AI governance framework. Noting that both discussions come from the same source – The Infocomm Media Development Authority (IMDA) – he observed that Singapore had a national ecosystem in place to develop a conversation around these issues. He offered three further comments, firstly, that unlike the GDPR, Singapore's data protection laws are not premised on human rights, but about managing data in order to grow the economy while ensuring security and trust. Human rights, as such, may be of less discursive relevance in the country and thus, might need to be de-linked from the discussion. Secondly, he suggested that we ought to be looking at AI from two perspectives: 1) looking at what businesses are doing with AI to collect information, and how the sophistication of AI has expanded the definition of what personal information is; and 2) understanding how AI might be harnessed as a tool for ensuring compliance. Finally, he suggested that one way to bring the conversation forward was to delve into specific sectoral experiences. In Singapore, for example, the financial sector has taken the lead in crafting up policies – yet there are fundamentally differences between that sector and medical applications. Ethical questions, it follows, have evolved differently and will need to be addressed more specifically in each sector. Expanding on this, Nydia commented that we also have to take into account regulatory objectives. For example, in the financial sector a major objective is financial

inclusion. As such, regulators might prioritise lax data protection laws or privacy considerations in order to include more people, alternatively, they might limit the uses of certain algorithms so as to not exclude certain demographics. There is, as such, a constant balance in objectives that lead to tensions between data protection regimes and the possible applications of AI.

Brian commented that in Hong Kong, AI governance has largely fallen within the data protection commission's remit because other organisations still lack the relevant expertise. Compared to Singapore, however, he thought that the latter continues to have better coordination across governments.

Expanding on the thread of government coordination, Smitha shared that India is struggling to maintain coherency across its policies. In addition to multiple national level AI policies, sectoral regulators may also be prioritising other objectives. She observed that once a data protection authority is finally in place, it might be only a matter of time before contradictions in priorities and actions between them and sectoral regulators (e.g., financial regulators) emerge. Finally, some states are further along this conversation – talking about using facial recognition for surveillance, for example – while other states have yet to be fully digitalised.

Yong Lim offered some insights on developments on South Korea, which recently had a major overhaul of their data protection regime that had implications for privacy regulations, telecommunications legislation, and uses of credit information. Their original data protection regime was known for being extremely strict and barring any innovation that used data. The impetus for this change was to boost the country's economy in the fourth industrial revolution through increasing assess to data and allowing for the combination of different datasets across firms, while still allowing for consumer/personal control. He also commented that data protection continues to have particular challenges – replicated in AI Governance – because in many cases the issue is not that data is collected, but how it is used. Many people are locked out of knowing how it is being used and who is using it – and privacy laws have been insufficient at addressing these shortcomings. How, then, can we learn from these shortcomings in order to do a better job for AI Governance? Smitha agreed and noted the increasing consensus that data protection laws are important but insufficient. More issue will arise from AI-use, but connected to these questions are also open questions about the ways in which platforms need to be governed, and how social media networks work. She suggested that we need to stop seeing data protection as a solution in itself, and think about bigger ways of regulating big tech.

Mark suggested that a rights-based focus is insufficient in this arena: this cannot be resolved by questions about ownership. Rather, people have certain expectations of data use and continuously have those expectations violated. What is missing so far has been knowledge across people and systems that are using data. If there is a commercial enterprise out of using people's data, we still need to require an openness in that use. Large, private firms are currently making money from data that others have produced – without the latter's knowledge. We cannot expect to have a successful protection programme until people are aware of that these uses are. These are also complicated by the fact that data producers themselves are careless, and so we need some form of dual process: firstly, having transparency as a key theme in secondary data use; and secondly, educational mechanisms for responsible data use and production.

Jolyon raised a question about culturally different expectations around issues like privacy. The anxiety in Australia about the government having access to accumulated data sets comes from a set of expectations about what the government knows. In contrast, in China, people are much less surprised and correspondingly much less anxious about the government having that information. If these cultural concepts are different, how might one approach these differences? Warren agreed that in Singapore we generally see less anxiety as well, observing that there isn't a consistent approach in the country and businesses have dealt with privacy issues in their own ways, although he notes the existence of the APAC cross-border privacy rules system. Convergence, however, is emerging around certain other topics like cybersecurity. Malavika noted that in India it was only a recent victory that 'consent' was required, and yet this fails to capture the issues raised by AI, where ambient technologies are capturing data all the time. How might that square into consent? It is also the case that in India, the government is exempted from the need to attain consent for data collection and uses. Smitha also noted that in India consent is focused on the *notification* aspect, rather than actual consent. Mark acknowledged the importance of cultural relativity in the discussion around expectations, suggesting that much of the question revolves around citizen expectations of the state and private sector. What issues/challenges arise in the discussion of AI Governance, as such, becomes extremely context specific.

Arisa then made some observations about Japan's data protection regime, noting that the country revises its law every three years. In contrast to the EU's 'data portability' discussion, her own research found that the public might not be ready to engage in that discussion – and so rather than introducing data portability, more interest is currently being shown in the 'information bank' concept. In relation to this, she also noted that there is currently discussions about more virtual assistants who might end up making decisions for how one's personal data ends up being used.

Final Remarks

Malavika offered some closing thoughts and questions: one theme that came out of the discussion was the concept of what AI cannot accomplish – a concept that is as important as where AI will have the most impact. She also noted that we often fail to discuss the invisible human labour that powers these technologies, and raised the question of why we expect more from technology than from other human beings. What can we borrow from other fields – finance or IP, for example – about what has worked, so that we won't have to start from scratch?

Mark reflected on six ideas that came out of the discussion:

1. The concept of cultural differences – the idea of 'Asianness' and governance. How do we get away from tokenism and avoid the suggestion that we might be mimicking unhelpful stereotypes. How do we make 'Asianness' something that is a valuable research device?
2. The human-machine interface. If we are going to have governance that is about creating relationships in which AI exists and has an enriching potential for human decision-making, how do we do that? Is there an Asian function?
3. The idea of governance, trust, and appropriateness. He suggested that the idea of trust is a dangerous device in some respects because it might mask confusion and distrust; rather, we want to discuss the generation of a trusted framework in which AI can sit and develop. Majority of people don't naturally trust machines, AI, and its proponents: we want to see ourselves as researching trust positively.

4. Pathways of principles. How do we get a handle from principle to applied processes? How are stakeholders in Asia different and unique in their demands? How do we move from models of pathways to practice?
5. Governance Failure: risk, liability and damages. It has been governance failures that has allowed for the generation of perceived fears – this is what is underpinning the reluctance to engage with AI, the perception that we are being lied to and not told the truth (e.g., in the myths around job displacements). How do we look at other solutions beyond monetary compensation?
6. Data protection and AI enhancement. We need to move beyond seeing these two things are separate entities – AI cannot exist without data, but it is also producing data. Data governance means talking about AI *and* data governance in the same breath.