Cambridge Forum on AI: Law and Governance (2025), **1**, e12, 1–20
doi:10.1017/cfl.2024.12

Cambridge Forum

RESEARCH ARTICLE

Governing intelligence: Singapore's evolving AI governance framework

Jason Grant Allen[1], Jane Loo[2] and Jose Luis Luna Campoverde[3]

[1]SMU Yong Pung How School of Law; Director, SMU Centre for Digital Law, Singapore; [2]SMU Centre for Digital Law, Singapore and [3]SMU School of Computing and Information Systems and affiliated researcher, SMU Centre for Digital Law, Singapore
Corresponding author: Jason Grant Allen; Email: jgallen@smu.edu.sg

(Received 23 July 2024; revised 7 November 2024; accepted 8 November 2024)

Abstract
This paper provides an outline analysis of the evolving governance framework for artificial intelligence (AI) in Singapore. Across the Singapore government, AI solutions are being adopted in line with Singapore's "Smart Nation Initiative" to leverage technology to make impactful changes across society and the economy. In tandem, Singaporean authorities have been assiduous to release a growing number of governance documents, which we analyse together to chart the city-state's approach to AI governance in international comparison. The characteristics of Singapore's AI governance approach include an emphasis on consensus-building between stakeholders (particularly government and industry but also citizens) and voluntary or "quasi" regulation, lately with an emphasis on promulgating standards (*AI Standards*, n.d.) and audit-like frameworks. Singaporean regulators have also been early movers (globally, and especially in the region) in the promulgation of normative instruments on AI governance including developing the world's first AI Governance Testing Framework and Toolkit, AI Verify. The Singapore approach may be compelling for other jurisdictions in the region and around the world with an interest in a collaborative, balanced and consensual approach to governing AI outside of strict regulatory mechanisms. However, any jurisdiction adopting aspects of its evolving model would have to duly account for relevant differences in social and institutional conditions.

Keywords: AI governance; Singapore AI governance; collaborative AI governance; AI ethics and governance

1. Introduction

This paper provides an outline analysis of the evolving governance framework for artificial intelligence (AI) in Singapore. Recognised as one of the most "AI ready" jurisdictions in the world, AI has been identified as a technology cluster of strategic importance for the national economy's continued development, and the city-state is influencing the trajectory of AI as a vendor and even developer of AI solutions as well as through the typical policy levers. Across the Singapore government, AI solutions are being adopted in line with Singapore's "Smart Nation Initiative" to make impactful changes across society and the economy, and preparatory work has commenced on a large language model (LLM) competent in South-East Asian languages (Goh, 2023). Singapore's signature approach to AI governance reflects its governance culture more broadly. The characteristics of the approach are an emphasis on consensus-building between stakeholders (particularly government and industry but also citizens) and voluntary or "quasi" regulation, lately with an emphasis on promulgating standards

and audit-like frameworks. This is perceived to harness the productive energy of free-market capitalism contained within clear guardrails,[1] reinforced by a willingness (and capacity) to intervene if and when deemed necessary. Singaporean regulators have been early movers (globally, and especially in the region) in the promulgation of normative instruments on AI governance including developing the world's first AI Governance Testing Framework and Toolkit, AI Verify (*Singapore Launches World's First AI Testing Framework and Toolkit to Promote Transparency; Invites Companies to Pilot and Contribute to International Standards Development*, 2022). This "responsible optimism" approach has been reinforced by the advent of Generative AI (GenAI) as an actual, rather than future or hypothetical, economic factor in Singapore's knowledge- and services-based economy. This approach may be compelling for other jurisdictions in the region and around the world with an interest in a collaborative, balanced and consensual approach to governing AI outside of strict regulatory mechanisms. Singapore's influence can already be observed to have influenced the recent Association of South East Asian Nations (ASEAN) Guide on AI Governance and Ethics (*ASEAN Guide on AI Governance and Ethics*, n.d.) which was promulgated in early 2024. However, whether any regulatory strategy is successful can only be determined over time; further, whether the approach translates into economies and societies of a different scale, and with different regulatory environments and cultures, is an open question.

2. Background

Singapore was recently ranked highest globally in terms of "AI preparedness" by the International Monetary Fund (*AI Preparedness Index*, n.d.). The country has adopted various strategies, including establishing national AI strategies (NAISs), creating research hubs and focusing on both domestic and external AI markets to promote both AI uptake domestically and Singapore's position in the global community as a leader for responsible AI ecosystem building (Khanal, Zhang & Taeihagh, 2024).

Immediately, this makes its AI governance approach interesting from a comparative perspective – the obvious question being whether its approach could be adopted elsewhere with the same effect. The island city-state is thus an interesting study in technology governance. On the whole, we could characterise Singapore as a jurisdiction that is an early uptaker, pro-business and generally "tech optimistic." At the same time, there is a strong tradition of state regulation for public safety and, as we have seen in areas like crypto assets (Chua, Seah & Jenie, 2024), willingness to make decisive moves to protect the Singaporean public from harms as they arise (or are foreseen).

As a recent analysis noted, Singapore's journey is marked by long-term plans such as the Smart Nation Initiative and by its attempt at finding a balanced position as a small state and middle power amid the global AI rivalry between the US and China. Although a proper analysis of this aspect is beyond the immediate scope of this article, it is important to note that the soft-law and technical-driven approach described in the following analysis is in some (possibly large) measure a product of Singapore's attempt to thread the needle between great power competition, in which technology governance is a major theatre.

A number of further characteristics set the stage for understanding Singapore's unique brand of AI governance as it is evolving. First, Singapore is a highly advanced knowledge economy in South-East Asia (SEA), and a clear standout in the region in terms of economic development and digital infrastructure. While it sounds cliché, Singapore's complex and at times paradoxical characteristics are important to keep in mind: Singapore is a common law system, a financial hub, an ageing population, a democracy with remarkably stable party dominance and a plural and modern but deeply Asian society with broadly conservative, Confucian-influenced values. This puts the city-state at cross-roads of economic, cultural and political geographies, which underscore its apparent "miracle" of economic

[1]See, for example, Huff (1995), elaborating on the nuanced interplay of market and planning in Singapore economic philosophy.

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

development over the past decades. Singapore's political system and regulatory culture characteristically combine a strong commitment to free-market capitalism with an appetite (and capacity) for interventionism where necessary, undergirded by a small scale and population and a preference for consensus-driven decision-making.

Another relevant piece of the backdrop is a relatively high level of adoption of and apparent trust in emerging technologies. A 2023 survey of 210 respondents, for example, found that over 67 percent of Singaporeans had used GenAI at that time, and that 69 percent of them described their experience as positive. The conclusion drawn from the report was that, despite the current news attention (much of which was negative), a significant portion of the respondents demonstrated positive or accepting perspectives towards GenAI (*Asia Pacific AI Readiness Index*, 2023; Cazzaniga et al., 2024; Lim, 2023).

Similar data are appearing from other South-East Asian countries, with more respondents expecting benefits than drawbacks from AI deployment (Von Kameke, 2024). In contrast to Europe or the US, "doomer" narratives feature less prominently in the Asian discourse on AI.[2] (This positive attitude was recently stylised in the Sci-Fi motion picture *The Creator*, in which a conglomerate "New Asia" becomes a kind of haven for intelligent androids persecuted by reckless, interventionist American forces.) One reason for this could be that worker populations in the West have experienced the social and economic downside of automation and globalisation over the past decades, where Asian populations have experienced some upside from Western deindustrialisation. These positive perceptions towards AI (at least in Singapore) could perhaps also be attributed to government initiatives that target citizen inclusion and participation in AI development and deployment,[3] or perceptions that have been cultivated around the utility and promise of AI as positively benefiting society.[4] They may also relate to Singapore's concerns about the demographics of aging and workforce replacement.

While these points are conjectural, there do seem to be clear differences between developing and emerging economies on the issue of trust and attitudes towards AI adoption. A recent poll, for example, indicates that about 75 percent of Americans fear job losses from AI (Marken & Nicola, 2023). On the other hand, despite similar concerns over job displacement ("Singaporeans across Generations Raise Concern on Impact of AI in Job Security," 2024), Singapore provides a unique contrast as a highly developed, knowledge-based economy that is AI-optimistic and less concerned with AI's potential risks than its opportunities and benefits.

Another relevant background fact is the "communitarian" orientation of Singaporean ethics, which is often contrasted with the rights-based, individualistic orientation of Western ethical approaches.[5] It is easy to make broad claims about the different ethical orientation of Asian societies influenced by Confucianism and other non-Western traditions. What this actually means is often left at the level of vague, intuitive, truisms. However, as in all truisms, there is undoubtedly something true there: Confucianism does emphasise notions such as societal "harmony" in a manner quite different to, for example, Kantian or Utilitarian ethics.[6] This paper is not the place to explore that question, so we

[2]This has been consistently evident, for example, in the Roundtables held under the SMU-Microsoft Asian Dialogue on AI Governance since 2019.

[3]See initiatives like Tech Kaki that welcome citizens' feedback and participation in the development of new Government digital products and services (*Tech Kaki Community*, 2024) Additionally, Singapore's focus on upskilling and enhancing citizens' digital literacy may have also contributed to enhanced trust and confidence in AI.

[4]See for example, the uses and integration of technologies like robotic dogs In Singapore for the purpose of enforcing safety and security (Haziq Mahmud, 2022).

[5]A classical statement of the "shared values" common to all of Singapore's various ethnic communities was promulgated by the government in 1991 as comprising: (1) Nation before community and society above self; (2) Family as the basic unit of society; (3) Community support and respect for the individual; (4) Consensus, not conflict; and (5) Racial and religious harmony (see *Shared Values*, 1991).

[6]"For Confucians, the difference between harmony and disharmony is like that between right and wrong, good and bad, and success and failure.… As far as the need for harmony is concerned, Confucians tend to see more consistency than distinction between the 'private' and the 'public' (as is seen in Western liberalism), between the political and the non-political, and between human society and the natural world. When persons and things are engaged in a healthy, stable interplay and each gets its due,

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

will not focus on this aspect of Singapore's approach to AI governance explicitly. However, it is fair to say that a communitarian ethos and harmony-focussed ethical orientation is evident on the face of Singapore's brand of AI governance. This ethos is succinctly encapsulated by Singapore's "whole-of-nation" approach (*Be Part of Our AI Journey: Working Together to Better Make Use of AI*, n.d.), which integrates the efforts and resources of Singaporeans, businesses, researchers and policymakers in driving effective AI governance. Additionally, Singapore's commitment to "community-building" initiatives similarly underscore the country's emphasis on the importance of collaboration and knowledge-sharing models (like the AI Verify Foundation initiative explored below).

To date, Singapore has not enacted any specific law regulating the use of AI in general, and there is no indication that legislation specifically targeting the use of AI horizontally will be introduced soon. If anything, sectoral regulation of AI use cases in areas such as mis/disinformation could ensue if the gravity of the threat is deemed sufficient. The notable absence of AI regulation in Singapore may appear at odds with Singapore's regulatory enthusiasm in other fields. It is likely an expression of the belief that technology is best harnessed by shaping conditions for it to be used in economically productive and socially acceptable ways. As mentioned above, the latent capacity for the Singapore government can act swiftly to impose any legal interventions deemed necessary is an important differentiator to consider when comparing Singapore's regulatory approach to other jurisdictions, in which legislative response time may be considerably longer and more contested.

Although the conversation is steadily moving from ethics to regulation, the recourse to general law should not be forgotten in the AI governance debate, either: where the use of AI gives rise to personal injury, property damage or financial loss, the claimant can seek remedies under the general law. Singapore does not have product liability laws; instead, remedies are available under statutes such as the Sale of Goods Act – and of course, actions in tort and contract under the common law are also "living" and adaptable to AI. The Personal Data Protection Act similarly presents itself as an important legal instrument for claimants as it imposes obligations on organisations when it comes to the collection, use and disclosure of their personal data.[7] Further, some specific legislation such as the Road Traffic Act and the Health Products Act do address specific aspects of AI use.

In the sections that follow, we trace relevant developments in terms of significant documents, initiatives and projects, framed within the city-state's NAIS, which is a core component of its bid to be a world-leading "Smart Nation."

3. Singapore's NAIS

A significant background factor to consider is that Singapore's AI journey has been charted by a national strategy formulated in 2019 (NAIS 1.0) and updated in 2023 (NAIS 2.0), formulated as part of the country's Smart Nation Initiative launched in 2014. In this section, we briefly outline the approach set out in the two versions of the NAIS and trace out some of the differences between them.

The NAIS 1.0 focussed on three key aspirations: to develop Singapore as a global hub for AI innovation; to generate new business models and deliver innovative services to improve human welfare; and to equip the workforce to adapt in the changing digital economy. NAIS 1.0 pursued a holistic (or even total) vision of digital transformation, built on efforts in certain key sectors (e.g., healthcare, education and digital trade), and a raft of national AI projects and efforts towards building out an ecosystem including the "triple helix" of researchers, industry and government (*National AI Strategy: AI for the Public Good, for Singapore and the World*, n.d.).

this is deemed harmony, and the opposite as disharmony.… The ideal of an individual is to harmonise not only with one's own person but also with other individuals. The ideal of a society is to harmonise not only with the society but also with other societies. The human ideal is to achieve harmony not only among the members of the human community but with the rest of the cosmos as well." (Li, 2006; see also Song, 2021)

[7]Data protection laws are important because data serves as the foundational building blocks for most AI innovations.

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

NAIS 2.0 continues this vision and strategy in broad terms, but the emphasis is more strongly focussed on public good (i.e., beyond that assumed within economic development), global competition (and competitiveness) and on building a more robust AI ecosystem rather than national projects. NAIS 2.0 also reiterates and strengthens the state's strategic commitment to cultivating an AI-ready populace comprising both technical experts and AI users. NAIS 2.0 elaborates three "systems," 10 "enablers" and 15 "actions." The systems and enablers are set out in the following terms:

> **System 1: Activity Drivers (Enablers: Industry, Government, Research)**. Industry, Government, and public research performers have deep technical capabilities that can be applied to deliver value. We need to orchestrate them around meaningful use cases and problem statements to transform our economy and society.
>
> **System 2: People & Communities (Enablers: Talent, Capabilities, Placemaking)**. We will attract more top-tier researchers and engineers to work with and from Singapore. More of our technology workforce should work to scale novel AI solutions, that form part of the toolkit which a confident base of enterprises and workers can use.
>
> **System 3: Infrastructure & Environment (Enablers: Compute, Data, Trusted Environment, Leader in Thought and Action)**. We will ensure that Singapore hosts the necessary infrastructure and provides a trusted environment for AI innovation. This will make us a credible leader and preferred site for AI development, deployment, and adoption. (*National AI Strategy: AI for the Public Good, for Singapore and the World*, n.d.)

In short, Singapore's official approach to AI is positioned as optimistic, cooperative, inclusive and frames AI governance and risk in terms of prerequisites to unlocking its assumed benefits. Singapore's government plays a proactive role in promoting AI by acting as a policymaker and a market participant through public procurement and sector-specific incentives. The strategy also sheds some light on Singapore's dedication to fostering a trusted AI ecosystem that is conducive to AI recipient communities, while ensuring the effective mitigation of AI risks through clearly defined responsibility structures. Significantly, Singapore's community- and ecosystem-building approach fosters a collective, participative and responsible approach to solving AI-related issues; with the government actively encouraging its citizens, industry, academia, and international partners to collaborate in the development, deployment and monitoring of emerging technologies. To elaborate, the government maintains open communication channels to solicit ideas and solutions for new National AI projects from the public (*Be Part of Our AI Journey: Working Together to Better Make Use of AI*, n.d.), while industry is encouraged to partake in beneficial data sharing partnerships and practices. (*Trusted Data Sharing Framework*, n.d.) Industry is also incentivised to digitise and create new AI solutions with the government's active support, (*Be Part of Our AI Journey: Working Together to Better Make Use of AI*, n.d.) and citizens are continually provided access to various programmes to enhance their AI literacy and skillsets through initiatives such as AI Singapore Talent Development programme, Techskills Accelerator and SkillsFuture for Digital Workplace Programme.

Throughout, the strategy stresses, the importance of maintaining societal values like fairness, non-discrimination and human-centricity as principles operating to mitigate AI risks and integrate AI technologies into societal frameworks to protect important public interests and safeguard public trust. This framing helps us to contextualise and understand the various initiatives that have been launched towards AI governance in the past five years. Of course, the devil is always in the detail, and in particular how such general and open-ended principles are actually operationalised.

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

4. An early mover (but not towards "regulation")

A recent study observed that the introduction of AI governance principles is a relatively recent phenomenon: as late as 2016, no company and only one country (the US) had issued official documents on AI governance. (Chesterman, Gao, Hahn & Valerie, 2023) By 2021, that number had grown to many dozen – and a majority in Western Europe. Subsequently, in 2023, in a paper by Kluge *et al.*, the authors noted the explosion of over 200 AI ethical guidelines and governance policies published by public bodies, academic institutions, private companies and civil society organisations worldwide (Corrêa et al., 2023).

Singapore counts as an early mover, and its first foray was made with some aplomb. The Infocomm and Media Development Authority (IMDA) developed a Model Framework for AI Governance and officially launched it at the World Economic Forum (WEF) in 2019. This was a bold statement that Singapore was (i) open for business, i.e., advertising an open and pro-innovation regulatory environment while also (ii) staking a claim as an early mover in AI governance that would insist on some minimum normative AI standards. We will review the substantive provisions of the Model Framework (as updated) below.

In the wake of announcements regarding the European Union's AI Act early this year, it is easy to assume that the terminus of AI governance evolution must be "hard" regulation of a horizontal kind. But it is not a foregone conclusion that general (as opposed to sector-specific) regulation is the necessary or even desirable end point on any given jurisdiction's AI governance journey – nor, in fact, that ideal-typical hard regulation is the goal at all. In fact, as the travails of the AI Act draft over the past years have demonstrated, hard-edged legislative regulation of AI with prohibitions or conditions on its development and deployment is probably the most challenging form of regulation to get right.

Not only is it difficult to define the subject matter (AI) with adequate precision;[8] as Jerrold Soh *et al.* observe, AI "regulation" can take a number of shapes (Soh, Lim & Yeong, Forthcoming). The categories are not rigid, and most national approaches actually display some hybridity or mixing. However, if we set out a spectrum of regulatory approaches from industry self-regulation to co-regulation, quasi-regulation and direct regulation "proper," we could characterise Singapore's approach to AI governance as *quasi-regulatory*:

> Quasi-regulation describe arrangements through which government influences businesses to comply, but which do not form part of formal government regulation. It is also commonly known as "soft law", or a rule "which has no legally binding force but which is intended to influence conduct". General categories of soft law include procedural rules, interpretive guides, instructions to officials, prescriptive/evidential rules, commendatory rules, voluntary codes, rules of practice, management or operation, and consultative devices and administrative pronouncements. These may be published in the form of "internal guidelines, rule books and practice manuals, circulars, operational memoranda, directives, and codes of conduct." (Soh et al., Forthcoming)

As Soh *et al.* continue, the Model Framework is a "voluntary" and "ready-to-use" tool primarily aimed at helping organisations deploying AI to do so in an appropriate manner, with particular regard to notions of "responsibility" and "human-centricity":

> The Model Framework reflects a collaborative style of governance, where industry and government co-develop measures and frameworks in an iterative process that allows regulators to better understand the shape and constraints of [AI] technology and therefore what could eventually become a set of rule and regulations governing them. (Soh et al., Forthcoming)

[8]The debate may centre around whether AI should be defined more "technically" to contribute to better precision or whether a more human-based vocabulary should be adopted to cover advances in AI so the law can be "future proofed."

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

On its terms, the Model Framework is situated within Singapore's overall approach to AI governance that is premised on balancing the twin imperatives of (i) facilitating innovation while (ii) safeguarding consumer interests by proactively managing the risks that arise from emerging technologies. In our current geoeconomic climate (i.e., of industrial policy and direct state involvement in critical industries and technologies), it should also be stated that the Singaporean model seems to favour stimulating innovation through a mixture of soft-touch regulation (to avoid "stifling" innovation) on the one hand and significant public sector investment in critical technology from trusted vendors to catalyse the industry and investment in research and development through universities and similar institutions.

Thus, although Singaporean regulators and other stakeholders are well aware of the various risks presented by AI in general (and by GenAI in particular), the advantages of a "softer" approach have been deemed to outweigh those of a more direct form of regulation at this stage and for the foreseeable near term. In mid-2023, Singapore officials were quoted as saying that no bespoke regulation was on the cards at the current time, and that the focus was rather on learning about use cases from industry in order to determine what kind of regulation might be needed. Characteristically, the approach is one of working together with industry, research organisations and other governments (Chang, 2023).

As alluded to above, the choice not to adopt a hard regulatory approach may also allow Singapore as a small jurisdiction and a net importer of technology to avoid direct regulatory incompatibility with any of the global majors – the US, EU and China – which have each adopted a different AI regulatory strategy, and may allow Singapore to position itself as a leader in the space, situated as it is in a region with many small- and medium-sized jurisdictions that find themselves in a similar position.

Nevertheless, While Singapore's AI governance framework is recognised for its flexibility that balances innovation with risk management, an important consideration is whether this reliance on quasi-regulation will continue to provide sufficient safeguards as AI technologies become more ubiquitous. The pro-business, tech-optimistic stance Singapore adopts – focusing on voluntary compliance and industry collaboration – has enabled rapid AI development without the constraints of rigid legal frameworks (Khanal et al., 2024). However, the absence of strict regulatory mechanisms raises questions about long-term accountability, particularly as GenAI introduces novel risks like misinformation and privacy concerns. A more enforceable governance structure could ensure that as AI continues to evolve, it remains aligned with ethical standards and safeguards public trust. Singapore's current reliance on collaborative governance might also benefit from incorporating stronger accountability measures to address these challenges while maintaining its innovative edge.

With this background in mind, it is now convenient to review the Model Framework and relevant instruments issued by other Singapore government agencies.

5. The IMDA Model Framework for AI governance

The IMDA is a statutory board under the Ministry of Digital Development and Information (previously, Ministry of Communications and Information). It took its modern shape in 2016 following the merger of several predecessor bodies with remits in telecommunications, computing and media; today, IMDA has a broad remit that includes oversight of Singapore's Personal Data Protection Act by the Personal Data Protection Commission (PDPC). IMDA is not only a regulator; it is also a development agency that plays an active role in the promotion of Singapore's digital economy agenda and holds a number of statutory functions of a development rather than regulatory nature.[9]

[9]See Info-communications Media Development Authority Act 2016, s. 5(1) setting out development functions including, *inter alia*: promoting research and development into technological matters relating to the information, communications and media industry; promoting, where suitable, self-regulation in the information, communications and media industry in Singapore; promoting the use of information and communications technology in Singapore and, where necessary, to collaborate with the Government Technology Agency (established by section 3 of the Government Technology Agency Act 2016)

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

Although separate divisions operate within the body, this dual nature has probably complemented the consensus-driven, collaborative approach typical of Singaporean rule-making and oversight, embracing various government entities, industry stakeholders and civil society institutions such as universities and international organisations.

The Model Framework provides practical, voluntary guidance for private sector organizations on ethical and governance aspects of AI design and deployment. The Model Framework was released initially in January 2019, with a second version in January 2020. The initial release of the Model Framework in 2019 was widely acclaimed as being the first in Asia to provide practical guidance (*Singapore Releases Asia's First Model AI Governance Framework*, 2019) to organisations in navigating key ethical and governance challenges when designing/deploying AI solutions. The Model Framework 1.0 underscores the role of corporations in minimising AI risks to foster public trust and mirrors Singapore's collaborative, industry-led AI strategy. It is avowedly agnostic in terms of technology, algorithm, sector, scale and business model and provides a baseline set of considerations to adopt – and adapt. It is generally more concerned with deployments at scale, and less concerned with incidental deployments of AI modules in conventional software stacks (Soh et al., Forthcoming).

Both versions of the Model Framework are built around two main ethical principles. First, AI used in decision-making should be transparent, explainable and fair; second, AI solutions should prioritise "human-centric" values. These principles are translated into practical steps, addressing the challenge of implementing abstract ethical concepts.

Since its inception, the Model Framework has aspired to adaptability and agility. Described as a "living document" (*Singapore Releases Asia's First Model AI Governance Framework*, 2019), the Model Framework was designed to evolve alongside the rapidly changing AI landscape. Consistent with this vision, the second iteration of the Model Framework was published shortly after its original release. Based on industry feedback and consultation, the second version incorporated additional considerations (including an interest in robustness, reproducibility and auditability) to enhance the Framework's operability and relevance (Model Artificial Intelligence Governance Framework Second Edition, n.d., p.5). The second version also included additional industry examples to showcase how organisations can align more closely with the four outlined areas. The second version of the Model Framework also pays closer scrutiny to broader internal and external stakeholder relationships, and how organisations can determine the appropriate level of human involvement required in AI-decision-making.

Key elements of the second version of the Model Framework include the following:

- *Internal AI governance*, emphasising the importance of clear roles and accountability within organisations to manage AI-related risks effectively. This includes providing relevant personnel with proper training and resources to discharge their duties and establishing a system of risk management.
- *Human involvement in AI decisions*, advocating for a structured approach to evaluating and integrating human oversight based on potential harm and severity. This determination of human involvement ought to be a continuous process.
- *Operations management*, focusing on the integrity of data and algorithms to prevent biases and ensure fair AI outcomes.
- *Stakeholder interaction*, emphasising transparency and the importance of clear communication about AI use to build trust. Stakeholders' access to open communication avenues is also critical in the management of AI risks.

The Model Framework thus encourages industry self-regulation over legislative approaches, advocating for a balanced approach between innovation and risk management. It is worth noting also

in respect of that; and promoting the use of the Internet and electronic commerce in Singapore and to establish regulatory frameworks for that purpose. See also s. 6 setting out the powers of the IMDA.

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

that the principles outlined in the Model Framework (as iterated) do not deviate significantly from commonly-endorsed AI ethical principles. These principles, emphasising the values of transparency, explainability, fairness, non-discrimination and retention of human involvement are similarly covered in regulatory approaches like the EU AI act. The Model Framework can be said to complement Singapore's AI governance approach by leveraging on the capacity, resources and initiative of industry participants. Industry's efforts are then strategically bolstered with significant government support encompassing both financial (to foster innovation and help businesses digitise) (*How We Can Help*, n.d.) and technical (*100 Experiments*, n.d.; *Be Part of Our AI Journey: Working Together to Better Make Use of AI*, n.d.) dimensions.

Through supporting initiatives like the "Implementation and Self-Assessment Guide" and a "Compendium of Use Cases", organisations are aided in understanding and applying these principles within their own organisational and technical environment to promote more ethical and responsible AI development and deployment. True also to the spirit of the Singapore Model Framework being a "living document", it is relevant to consider the availability of communication pathways through which industry members can feedback on Model implementation and suggest improvements[10] for future incorporation. All of these initiatives combined help to enhance AI adoption and integration in society, its economy and government, and also to establish a reciprocal trust relationship between the state and relevant industries, and trust between industry, AI and the respective communities it operates within with the aim of promoting a thriving, innovative ecosystem underpinned by ethical AI practices.

6. AI Verify and AI Verify Foundation

In 2022, a Minimum Viable Product (MVP) called "AI Verify" was launched by IMDA and the PDPC to fill the perceived gap in AI governance testing and evaluation. AI Verify comprises an integrated toolkit for evaluation and testing (one is tempted to say "audit") within an organisation's own enterprise environment. In this regard, it can be observed that alongside initiatives like the Model Framework, the deployment of toolkits like AI Verify also showcases Singapore's unique technological approach to guide industry's responsible development and deployment of AI.

In mid-2023, Minister for Communications and Information Josephine Teo announced the launch of the AI Verify Foundation at the "ATxAI" conference to take stewardship of AI Verify (*Singapore Launches AI Verify Foundation to Shape the Future of International AI Standards through Collaboration*, 2023). Although IMDA continues to play a leading role, the vision seems to be that the consortium of industry partners will play an increasingly active role. Indeed, the AI Verify Foundation played a pivotal role in developing the most recent Model AI Governance Framework for Generative AI.

The AI Verify Foundation aims to tap a global, open-source community to develop AI testing tools for the responsible use of AI and to promote best practices for AI governance as they emerge (*AI Verify Foundation (Home Page) – Building Trust Through Ethical AI*, 2024). This approach is indicative of Singaporean government leadership to precipitate a consortium of industry players around a consensus that something is needed, and that it is incumbent on those players to contribute their vision of what that something is and how it ought to be achieved. The messaging is that "industry needs to demonstrate to their stakeholders their implementation of responsible AI in an objective and verifiable way" and that AI Verify is a way for them to do so: "IMDA and PDPC have taken the first step to develop an AI Governance Testing Framework and Toolkit to enable industry to demonstrate their deployment of responsible AI" (Ministry of Communication and Information, IMDA, & PDPC, 2022).

[10]Singapore encourages organisations to share and feedback on practical examples and their experiences in implementation by emailing corporate@pdpc.gov.sg (see *Singapore's Approach to AI Governance, n.d.)*.

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

The AI Verify toolkit does not define AI ethical standards itself, but is designed to be consistent with various AI governance frameworks, of course including Singapore's Model Framework as well as those from the European Union and Organisation for Economic Cooperation and Development (OECD) (*OECD AI Principles Overview*, n.d.; *What Is AI Verify?*, n.d.). The focus on interoperability for AI Verify sits squarely with the NAIS 2.0 to safeguard Singapore's interests and position in shaping the international rules, processes and governance trends emerging around responsible and trustworthy AI. The MVP is intended to produce higher levels of public trust in businesses deploying AI in Singapore, to facilitate interoperability of different governance frameworks (especially those that incorporate broadly similar ethical principles), and to contribute to the development of international standards on AI, including the ISO/IEC JTC1/SC 42 on AI to facilitate compliance (*ISO/IEC JTC 1/SC 42 Artificial Intelligence,* 2017).

To briefly examine the area of AI standards here (as also comprising Singapore's comprehensive approach in governing AI), it is worth noting that Singapore has a dedicated AI Technical Committee (AITC) (AI Standards, n.d.) involved in the adoption and development of AI standards and technical references both internationally and domestically. Furthermore, the AITC plays an important role in the different ISO/IEC JTC1/SC 42 working groups, contributing to ongoing discussion about current and future ISO standards related to AI. Beyond the significance of AI standards in enhancing trade connectivity and cultivating trust and reliability in AI recipient communities (Cheong & Liew, 2023), Singapore's participation in key international standard-setting organisations has underscored Singapore's relevance and autonomy in shaping current norms related to responsible and trustworthy AI on the international stage.

7. Initiatives of the Monetary Authority of Singapore

Given the importance of the finance industry to the Singaporean economy, and the scope for AI deployment in finance, the Monetary Authority of Singapore (MAS), which is the central bank and main financial regulator, has also been active in promulgating vertical AI governance instruments as well. In late 2018, MAS published "*Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector* (2019). This document was, in turn, updated in early 2019 to reflect the Model AI Governance Framework set out above.

MAS' activities in the AI governance space are also illustrative of the approach of using industry consortia with trusted private sector counterparts to develop and implement AI governance frameworks and tools. The "Veritas Toolkit" was released in February 2022, providing an assessment methodology for the FEAT principle of Fairness within financial institutions' model development lifecycle. The toolkit allows a user to create a project, and generate an assessment questionnaire and assessment report. Veritas Toolkit version 2.0 was released in 2023 with an improved Fairness assessment methodology and new assessment methodologies for the remaining principles (Ethics, Accountability and Transparency) (*MAS-Led Industry Consortium Releases Toolkit for Responsible Use of AI in the Financial Sector*, 2023). The toolkit is available on a github repository for organisations to test, and a number of whitepapers have been published providing use-cases and experiences with the methodology and toolkit (*Veritas Initiative: What It Is*, 2023; *Veritas-Toolkit*, n.d.).

The industry consortium comprised 31 financial institutions, including many multinationals. Again explaining the initiative in the context of Singapore's evolving NAIS, MAS Chief FinTech Officer Sopnendu Mohanty has described the next steps in the following terms, which are fairly indicative of the tone, position and direction of travel more widely:

> Our next step is to build an open-source ecosystem for [Responsible AI (RAI)] that is self-sustaining and vibrant. This ecosystem will facilitate the pervasive and consistent adoption of RAI principles by FIs and tech companies, as well as support a new generation of RAI stewards

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

who are skilled in using RAI tools. Additionally, the ecosystem will foster an active community of developers who can implement the FEAT Principles through the Veritas open-source ecosystem. Through this collaborative approach, we aim to promote RAI practices and ensure that the benefits of AI are shared by all. (*Veritas, Trust in AI - Document 5: From Methodologies to Integration*, n.d.).

8. Model Framework for GenAI

The subsequent sections move to discuss Singapore's evolving (and future) AI governance and policy frameworks in the context of responsible GenAI development and deployment. GenAI dominated the headlines in 2023, and the unprecedented level of public adoption of GenAI products stimulated increased regulatory attention to address a unique set of risks that affect not only the advancement of AI models but its implications for the general public. This widespread integration of GenAI across various industries has transformed workflows and reshaped professional practices.

In early 2024, the AI Verify Foundation and IMDA released a draft Model AI Governance Framework for Generative AI for public comment (AI Verify Foundation, 2024) and the version (GenAI Framework) was later finalised in May 2024 (AI Verify Foundation, & IMDA, 2024). This framework expands on the existing Model Governance Framework that covers "Traditional AI."

The framework builds on ideas from an earlier discussion paper (IMDA, 2023a) and earlier technical work to provide an initial catalogue and guidance on suggested practices for safety evaluation of GenAI models as well as insights from a Generative AI Evaluation Sandbox that was launched by IMDA and AI Verify Foundation in October 2023 to drive the development of evaluation benchmarks in GenAI (IMDA, 2023b). In substantive terms, the framework set outs nine proposed "dimensions" to support a comprehensive and trusted AI ecosystem. Predictably, the core elements rest on concepts that are stock-in-trade for AI governance: Accountability, Data, Trusted Development and Deployment, Incident Reporting, Testing and Assurance, Security, Content Provenance, Safety and Alignment Research and Development, and AI for Public Good.

For its part, the GenAI Sandbox encourages the adoption of GenAI by small- and medium-sized enterprises (SMEs). It is less a "regulatory sandbox" in which attenuated regulations apply and more of a "technology sandbox" to encourage experimentation by Singaporean businesses. Launched by IMDA and Enterprise Singapore (a statutory board under the Ministry of Trade and Industry that is the successor of multiple agencies including the former Trade Development Board) in February 2024, the Sandbox allows SMEs to access GenAI enterprise solutions over a three-month period that could drive their marketing and customer engagement activities (IMDA, & EnterpriseSG, 2024). While this is an activity more squarely in IMDA's development authority portfolio than its regulatory role, it is further illustrative of the active and hands-on approach the agency is taking towards "responsible" AI adoption.

The GenAI Framework offers valuable insights into understanding Singapore's ongoing (and perhaps future) approach to governing GenAI and other emerging technologies. An initial review of the GenAI Framework sheds some light on the urgency of appropriate AI governance structures both in Singapore and globally. The threat posed by GenAI, particularly to public trust, encompassing issues of misinformation, manipulation and deception is stressed in the Framework prompting vigilance and action from all relevant stakeholders in the AI ecosystem. Echoing the "whole of society" approach found in the NAIS 2.0 and the "traditional" Model Framework, the GenAI Framework similarly delineates the critical roles and responsibilities of citizens, industry, research, policymakers and international partners, urging the swift development of a safe and trusted AI ecosystem.

Nonetheless, acknowledging that not all risks posed by GenAI can be adequately addressed or anticipated by the Framework, the GenAI Framework mentions the need to examine current legislative instruments (such as existing personal data laws) to appropriately safeguard citizens' interests and rights (AI Verify Foundation, & IMDA, 2024, p.10). This references the point above on how the

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

general law is still an important mechanism in the overall AI governance debate and why a horizontal all-encompassing regulation (like the EU AI Act) should not be seen as the exclusive pathway to effective AI governance. Consequently, it can be said that Singapore is not solely reliant on voluntary and "soft" approaches in governing AI but is prepared and well-equipped to leverage the strong arm of the broader law to shape AI's impact on society. Interestingly, the GenAI Framework also considers the utility of "no fault insurance" as another potential safety net to ensure adequate compensation and redress for individuals affected by adverse AI outcomes (AI Verify Foundation, & IMDA, 2024, p.8).

Like the Model Framework for traditional AI, the GenAI Framework reiterates the need for good data practices and robust security safeguards to protect individuals' personal data and other rights. However, what is noteworthy in the Gen AI Framework is its explicit endorsement of various technological tools (AI Verify Foundation, & IMDA, 2024, p.22) including privacy enhancing technologies, input moderation tools and digital forensic tools. The GenAI Framework elaborates and validates the significance of these tools in safeguarding data confidentiality, security and privacy in the age of GenAI.

In addition, the "Trusted Development and Deployment" dimension from the Framework highlights the importance of standardising safety evaluations by recommending a set of initial model safety evaluations for LLMs, as outlined in the "Cataloguing LLM Evaluations" paper (Infocomm Media Development Authority & AI Verify Foundation, 2023) This paper identifies five attributes for evaluations of LLMs, namely: robustness, factuality, bias, toxicity generation and data governance. It then proposes recommended evaluation and testing approaches for each of these attributes. Thus, it seeks to provide different stakeholders with practical guidance for effective safety evaluations before deploying their models.

Again, in line with earlier comments on how AI Verify confirms Singapore's technological approach in guiding industry's responsible development and deployment of AI, the clear nod to the utility of these tools similarly reflects the country's interest in leveraging technology to promote more responsible and trustworthy AI systems.

These aspects as revealed by the GenAI Framework, including the reliance on general law, the openness to alternative non-regulatory solutions, and the leveraging of technology to manage the development and deployment of GenAI, affirm Singapore's multi-layered approach in the management of AI risks and its promised benefits. As acknowledged in the GenAI Framework itself, "no single intervention will be a silver bullet" (AI Verify Foundation, & IMDA, 2024, p.3).

The Model Framework's adaptability as a "living document" can also be examined against the GenAI Framework. One might ask why a specific framework for GenAI was necessary, rather than updating the Model Framework, and, given the approach that was taken, how the specific framework sits within the more general one. To put a point on it, the need for a GenAI-specific framework somewhat undermines the generality of the Model Framework, similar to the difficulties of accommodating protean foundation models into the European framework that is conceptualised around systems with (specific) purposes. Of particular interest is the evolution of the ethical principle of "transparency "and how it is interpreted and reflected in the GenAI framework to deal more directly with GenAI-specific threats, including AI "hallucinations," misinformation and copyright infringement. This examination provides critical insights into the concept of what "meaningful transparency" entails, as well as the practices the state determines as necessary for industry to align with the objective of "meaningful transparency".

While the principle of transparency already features in the second version of the traditional Model Framework, the GenAI Framework appears to have expanded, shaped and contextualised this principle to deal with GenAI-specific risks and benefits. First, while the GenAI Framework acknowledges that the level of transparency will need to be calibrated against an organisation's legitimate interests (including their proprietary interest); it also calls for the development of some "baseline transparency" to allow downstream users to make more informed decisions surrounding their AI use (AI

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

Verify Foundation, & IMDA, 2024, p.14). Second, the GenAI Framework urges that all disclosure of information (including data used, training infrastructure, evaluation results, mitigation measures, etc.) should be standardised to promote comparability across AI models (AI Verify Foundation, & IMDA, 2024, p.14). Third, that transparency requirements should vary depending on the risk level of the model in question. Naturally, models that have national security or societal implications will have to satisfy stricter transparency requirements (AI Verify Foundation, & IMDA, 2024, p.14).

The substantive content of the GenAI Framework also reflects Singapore's capacity to consider novel AI risks as it emerges and the desideratum of regulatory agility. The GenAI Framework explicitly considers the current and potential challenges posed by GenAI, including copyright infringement, its interaction with fair use policies, the dangers of poor-quality datasets, synthetic content and deepfakes. Of course, as those challenges raise an open-ended raft of issues, it is highly unlikely that the GenAI Framework does so once and for all. In the context of copyright infringement, it bears noting that Singaporean copyright law has an extremely broad Text and Data Mining (TDM) exemption, which was probably enacted with discriminative rather than generative systems in mind, and may well have taken a different shape had controversies about GenAI copyright infringement been current at the time of drafting. However, the systematic presentation of these issues in the GenAI Framework perhaps reveal something about Singapore's ability to stay ahead of the curve in dealing with emerging AI risks and the country's commitment to building a trusted AI ecosystem that produces social and economic benefits for society.

Consistent with Singapore's overall interest in striking a balance between the management of AI risks against its potential benefits, the GenAI Framework contains clear references to initiatives that facilitate AI-driven innovation and its responsible adoption. For instance, Singapore's GenAI Sandbox is referenced as providing SMEs with the necessary tools and training on GenAI enterprise solutions (AI Verify Foundation, & IMDA, 2024, p.29). Additionally, the GenAI Framework underscores the value of investing more research in model safety and alignment including the setting up of AI safety R&D institutes (AI Verify Foundation, & IMDA, 2024, p.27). The importance of expanding the pool of trusted quality data sets is also acknowledged(AI Verify Foundation, & IMDA, 2024, p.11), as well as the benefits of curating a repository of representative training data sets for their specific context (AI Verify Foundation, & IMDA, 2024, p.11).

Finally, given that technology transcends geographical boundaries, the proposal for a unified framework that outlines the "minimum" governance dimensions for adopting or deploying GenAI models has also emerged as a compelling concept. The "Harmonized GenAI Governance Framework" (H-GenAIGF) outlined in a recent paper by (Luna, Tan, Xie & Jiang, 2024), offers a comparative lens through which global governance approaches can be evaluated. The framework facilitates an analysis of how well different regions address the governance processes identified by the authors.

The H-GenAIGF identifies 15 governance processes, 25 subprocesses and 9 guiding principles, all of which were derived from an extensive review of governmental documents such as regulations, acts and bills from six distinct regions. These processes are categorised into four key constituents – Data, Model, Content Generation and Ethics – providing a lifecycle perspective on GenAI governance. This lifecycle perspective encompasses the input, inner functionality, output and ethical alignment of GenAI systems, thereby enabling a comprehensive evaluation of governance approaches across different countries. Singapore's GenAI governance approach, for instance, has been identified as a mixed model (risk-principle-based) while China employs a rule-based approach, the EU and the US adopt risk-based approaches, Canada adheres to a principle-based approach and the UK follows an outcome-based approach. The authors demonstrated that "risk-based" approaches, such as those followed by the EU and US, appear to provide more comprehensive coverage. Singapore's risk-principle-based approach governance focuses heavily on practical and structured oversight, combining elements of risk assessment and management with principle-driven governance, which focuses on aligning AI systems with social and ethical values (Luna et al., 2024).

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

The authors identified that Singapore's approach demonstrates promising results, particularly in its coverage of essential GenAI governance processes. Strong coverage is observed in processes such as "Model Validation & Testing," "Ethical Alignment & Human Rights," and "Ethical Design & Deployment." However, despite these strengths, Singapore's approach requires refinement, as there are areas with partial coverage.

Five out of the fifteen key governance processes, including "Model Development" and "Managing Distribution & Access Control" exhibit partial coverage. Furthermore, significant governance gaps persist in five other critical processes, such as "Upholding User Rights & Control" and "Managing Distribution & Access Control." These processes are particularly vital, as they ensure that individuals retain autonomy over their personal data, decisions generated by GenAI models and the responsible and equitable distribution of content, with clear controls over access and the conditions under which generated content is distributed. A number of other important governance concerns such as data processes, transparency and accountability of certain subprocesses and audit procedures, and responsiveness through feedback mechanisms to issues as they arise were also observed to be tackled only to a limited degree.

Notably, Singapore's approach places a strong emphasis on the evaluation of GenAI models, a critical aspect for mitigating risks associated with their deployment. One standout feature of Singapore's framework is its comprehensive coverage of the "Assessing & Mitigating Toxicity" subprocess. It is the only approach, among the six compared, to fully address this, which is increasingly important as the outputs of GenAI models are now accessible to a broad spectrum of users, many of whom may lack deep technical expertise yet, able to create and share generated content.

9. Project SEA-LION

Building upon the preceding discussion, there are also other ongoing and notable initiatives worth mentioning. For instance, the "Generative AI × Digital Leaders" initiative which aims to support digitally mature enterprises with access to GenAI expertise and resources (*IMDA Launches Generative AI × Digital Leaders Initiative to Provide Businesses Access to GenAI*, 2024) and the SEA-LION ("Southeast Asian Languages In One Network") project (*Sea-Lion*, n.d.), the region's first LLM (Tan, 2024).

Given its increasing economic importance (and increasing geopolitical sensitivity), AI is a focal point for subsidies and other stimulatory measures. Singapore has identified global leadership in the field of AI and AI governance as a matter of national significance and this has translated into ambitious funding for research and development in fields such as AI and advanced computing.

In late 2023, IMDA announced a partnership with AI Singapore and the Agency for Science, Technology, and Research (A*STAR) to launch a LLM trained with SEA languages with significant funding from the Singapore National Research Foundation, the major public research funding body. SEA is a linguistically, socially and economically diverse region with a multitude of languages that are underrepresented in pre-training data used for current LLMs. Project SEA-LION is a family of models that employ a unique approach to tokenising language (i.e., breaking sentences and words into data units) tailored to SEA languages. Embedded within the agenda set by NAIS 2.0 to "nurture Singapore as a global leader in AI solutions," the perceived benefit is GenAI capabilities better suited to Singapore and the region's cultural landscape:

> As technology evolves rapidly, there is a strategic need to develop sovereign capabilities in LLMs. Singapore and the region's local and regional cultures, values and norms differ from those of Western countries, where most large language models originate. A cornerstone of this initiative is the development of multimodal and localised LLMs for Singapore and the region to understand context and values related to the diverse cultures and languages of Southeast

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

> Asia, for example, managing context-switching between languages in multilingual Singapore. (IMDA, 2023c)

In the first instance, it will focus on major languages including local variants of English, Bahasa Indonesia, Malay, Thai and Vietnamese, and will be extended to include other regional languages in due course (AI Singapore Brings Inclusive Generative AI Models to Southeast Asia with AWS, 2024).

10. Guide to job redesign in the age of AI

Another topic of growing global concern acutely felt in Singapore is the potential impact of AI on the future of work. Although successive generations of AI, as indeed automation technology more broadly, has impacted the economics and labour dynamics of various industries in the past, the advent of GenAI in 2023 brought home the potential for AI to impact Singapore's knowledge economy – for good and for ill.

Singapore's perspective on the issue is likely framed by the demographic reality of an aging population and low birth-rate, combined with historical and present reliance on migrant labour in areas such as domestic work, aged care, construction, and food and beverage, as well as in skilled knowledge work. In the medium to long term, automation is perceived not only as beneficial but perhaps critical to the viability of the Singaporean economy and society. At the same time, there are deep concerns about the short- to medium-term impacts of automation on various segments of the workforce – in addition to the familiar patterns of workforce displacement through automation in manufacturing, transport, etc. GenAI is providing inroads into tasks performed by white collar knowledge workers, as well (Lung, 2024). Whether for enhanced productivity or knowledge worker replacement, the developments of the past two years bear on Singapore's self-conception as a small state whose people are its only significant 'natural" resource. While there has been no regulatory action on the point to date, conversations are being held at various fora in Singaporean society (William, 2024).

Notably, in 2020, the IMDA and the PDPC collaborated with the Lee Kuan Yew Centre for Innovative Cities (LKYCIC) at the Singapore University of Technology and Design to publish the first "Guide to Job Redesign in the Age of AI" (IMDA et al., 2020). The Guide focuses on a human-centric approach to AI and provides guidance to organisations and employees to unlock the potential of AI, so as to increase the value of "work." The four areas of job redesign include: transforming jobs, charting clear pathways between jobs, clearing barriers to digital transformation and enabling effective communication between employers and employees.

The future of work in the context of GenAI is also addressed (albeit briefly) in the GenAI Framework. Concerning the dimension of AI for public good, the responsible use of AI in the workforce is underscored with a focus on empowering workers through upskilling and appropriate job redesign (AI Verify Foundation, & IMDA, 2024, p.30), in particular. It remains to be seen how adequate these early efforts are; that will be determined by the actual impacts of GenAI-driven automation on labour dynamics, which is unknown, as well as on broader dynamics in Singapore's fairly unique labour market.

11. ASEAN AI guidelines

Finally, an examination into Singapore's AI governance approach would be incomplete without also factoring Singapore's significant influence on AI governance in the region. In early 2024, the ASEAN published its Guide on AI Governance and Ethics (*ASEAN Guide on AI Governance and Ethics*, n.d.). The Guide is intended to serve as a practical guide for organisations that wish to design, develop and deploy traditional AI technologies in commercial and non-military or dual-use applications. It focuses on encouraging alignment within ASEAN member states and fostering the interoperability of AI frameworks across those jurisdictions. It also includes recommendations on national-level

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

and regional-level initiatives that governments in the region can consider implementing to design, develop and deploy AI systems responsibly.

In broad terms, the Guide is consistent with the Model Framework, being guided by principles such as: Transparency and Explainability; Fairness and Equity; Security and Safety; Robustness and Reliability; Human-centricity; Privacy and Data Governance; and Accountability and Integrity. Both frameworks adopt a risk-principle-based approach, which not only emphasises the identification and mitigation of specific risks associated with AI deployment but also focuses on the guiding principles that govern these risks, particularly in relation to their impact on users. This approach ensures that organisations can calibrate the level of human oversight required based on the potential consequences for end users, balancing risk management with ethical alignment. Additionally, both frameworks allow for a high degree of flexibility, enabling organisations to comply in a manner that is feasible and appropriate given their specific contexts – an important factor in the ASEAN region, where operational environments vary widely.

ASEAN is a diverse and heterogeneous association in terms of legal system and economic development, and Singapore is clearly among the leading jurisdictions (if not the leading jurisdiction) in terms of AI governance initiatives. It is perhaps unsurprising that the Singaporean approach has been influential. As Kristina Fong explains:

> Singapore will be the Chair of the ASEAN Digital Ministers' Meeting (ADGMIN) and Related Meetings in 2024. It will be in a good position to lead ASEAN in the development of the ASEAN AI Guide on Governance and Ethics. With Singapore's position at the European Commission's High Level Expert Group on AI, it can guide the bloc's AI strategy to better align it with international standards including good regulatory principles found in the EU's AI Act. Singapore has an important role to play in bridging gaps within ASEAN on AI and at the same time, steering ASEAN's AI guidelines toward international standards. It will be important for ASEAN to have a strategy that is not just symbolic in form but also able to safeguard the ASEAN economies in this dynamic arena. (Siew Leng, 2023)

This said, one should still be mindful that aspects of the Singaporean approach to AI governance may not translate well (or at all) into other ASEAN jurisdictions, which have a very different economy, society, regulatory culture, resources, capabilities and institutional setting.

12. Conclusion: The future of AI governance in Singapore

This paper has outlined how the Singapore AI governance approach strives for balance between facilitating innovation, employing AI for wealth creation, managing AI's potential risks, safeguarding consumer interests and (at least behind the scenes) navigating geopolitical risk.

With the caveat that swift regulatory action in a vertical or sectoral fashion could well follow a scandal or shock, the governance trajectory of GenAI is unlikely to depart radically from Singapore's current AI governance approach in the near future. The government is expected to maintain a "soft touch" that is collaborative and industry-led, while delineating clear responsibility structures across the whole of society. In particular, national standards and/or technical references could also emerge, and these would enhance Singapore's AI governance ecosystem in a systematic and practical manner. Thus, individuals in their personal capacities will also be expected to do their part in navigating AI's potential risks, to take accountability for their own digital literacy and upskilling, and to contribute to the strengthening of Singapore's digital economy and digital society. In other words, all stakeholders will reap the benefits of AI (to solve societal needs and challenges) but are also expected to share in the management of AI's potential downsides and drawbacks.

Inspiring citizens' trust in AI will likely continue to feature as a national priority. Citizens' perception of AI will ultimately influence overall AI uptake and adoption, which will in turn promote Singapore's Smart Nation ambitions. The NAIS 2.0 explicitly identified AI as an indispensable force

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

essential to the continued relevance and prosperity of the nation (*National AI Strategy: AI for the Public Good, for Singapore and the World*, n.d). Embedding and reception of AI into society cannot be attained without first cultivating (and, where appropriate, restoring) public trust and confidence in the technology. The country acknowledged as early as 2018 in the MAS FEAT principles (*Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector*, 2019) that establishing a trusted ecosystem requires the proper negotiation of AI's risks and its benefits.

Current AI governance initiatives and government efforts also provide strong indication that the future of AI governance in Singapore will remain distinctive, characterised by collaborative government-industry consortia and efforts towards robust technological capabilities. Current initiatives demonstrate that Singapore is keen to be a first (and fast) mover. This proactive approach is perceived as crucial in securing the country's regional and international standing in the global AI race and discourse surrounding the responsible and trustworthy development and deployment of AI systems. Again, the efficacy of this approach will only be evident with the benefit of experience. Singapore recognises also that it cannot forge ahead into future alone and this sentiment is thoroughly reflected in its repeated emphasis for more regional and international collaborations. Currently, Singapore is an active participant in numerous multi-stakeholder platforms such as the Global Partnership on AI, the WEF, the AI Governance Alliance and the United Nations High-Level Advisory Body on AI (*National AI Strategy: AI for the Public Good, for Singapore and the World*, n.d.) With rapid developments in GenAI, the need for collaboration becomes more pronounced, particularly with a view to enhancing Singapore's research capabilities, to attract international AI talent, to share AI practices including appropriate governance strategies and to develop AI-related capacity building initiatives.

Acknowledgements. This article is an expanded version of a chapter in the forthcoming title by Calo, Ebert, Poncibo and Zou (eds.), *Handbook on Generative AI and the Law* (Cambridge University Press 2024) republished with permission of the editors and publisher, and with additional contributions from Jose LUNA. A version of this paper was presented at the RAILS/University of Turin Workshop "Generative AI and the Law" (Berlin/Online 6 July 2023) and at the Yonsei University Barun ICT Research Centre Asia Privacy Bridge Forum (Seoul, 17 October 2024), and the authors would like to thank participants at both events for their constructive feedback. Thanks to Jerrold SOH, LIM How Khang and Josh LEE Kok Thong for their comments on an earlier draft. The usual caveat applies that all errors and opinions remain our own.

Funding statement. This research was supported by the Singapore Ministry of Education (MOE) Academic Research Fund (AcRF) Tier 1 Category C Award No. 23-SOL-SMU-004 "Operationalising Responsible AI (RAI) in Public Administration."

Competing interests. The authors declare no competing interests.

References

100 Experiments. (n.d.). AI Singapore. Retrieved 1 April 2024, from https://aisingapore.org/innovation/100e/

AI Preparedness Index. (n.d.). International Monetary Fund. Retrieved 2 September 2024, from https://www.imf.org/external/datamapper/AI_PI@AIPI/ADVEC/EME/LIC?year=2023

AI Singapore brings inclusive Generative AI models to Southeast Asia with AWS. (2024, January 24). *Amazon*. https://www.aboutamazon.sg/news/aws/ai-singapore-brings-inclusive-generative-ai-models-to-southeast-asia-with-aws

AI Standards. (n.d.). AI Singapore. Retrieved 29 April 2024, from https://aisingapore.org/innovation/ai-standards/

AI Verify Foundation. (2024). Proposed Model AI Governance Framework For Generative AI: Fostering a Trusted Ecosystem. *AI Verify Foundation*. Retrieved 1 May 2024, from https://aiverifyfoundation.sg/downloads/Proposed_MGF_Gen_AI_2024.pdf

AI Verify Foundation (Home Page)—Building Trust Through Ethical AI. (2024). AI Verify Foundation. Retrieved 23 July 2024, from https://aiverifyfoundation.sg/ai-verify-foundation/

AI Verify Foundation, & IMDA. (2024). Model AI Governance Framework for Generative AI - Fostering a Trusted Ecosystem. *AI Verify Foundation*. Retrieved 23 July 2024, from https://aiverifyfoundation.sg/wp-content/uploads/2024/05/Model-AI-Governance-Framework-for-Generative-AI-May-2024-1-1.pdf

ASEAN Guide on AI Governance and Ethics. (n.d.). Asean.Org. Retrieved 26 April 2024, from https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics_beautified_201223_v2.pdf

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

Asia Pacific AI Readiness Index. (2023). Salesforce. Retrieved 13 May 2024, from https://www.salesforce.com/content/dam/web/en_sg/www/documents/pdf/salesforce_ai_readiness_index_2023.pdf

Be part of our AI journey: Working together to better make use of AI. (n.d.). Smart Nation Singapore. Retrieved 29 April 2024. from https://www.smartnation.gov.sg/nais/contribute/

Cazzaniga, M., Jaumotte, F., Li, L. et al. (2024). *Gen-AI: Artificial Intelligence and the Future of Work*. International Monetary Fund. Staff Discussion Notes No. 2024/001. Retrieved 13 May 2024, from https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379?cid=bl-com-SDNEA2024001

Chang, S. (2023, June 19). Singapore is not looking to regulate AI just yet, says the city-state's authority. *CNBC*. Retrieved 26 April 2024, from https://www.cnbc.com/2023/06/19/singapore-is-not-looking-to-regulate-ai-just-yet-says-the-city-state.html

Cheong, T. L., & Liew, L. (2023, June 12). *Why Standards Matter for AI and Big Data*. Retrieved 18 April 2024, from https://www.enterprisesg.gov.sg/resources/blog/quality-and-standards/why-standards-matter-for-ai-and-big-data

Chesterman, S., Gao, Y., Hahn, J. et al. (2023). The Evolution of AI Governance. *TechRxiv*. doi:10.36227/techrxiv.24681063.v1

Chua, P., Seah, A., & Jenie, T. (2024, May 20). *Singapore: Digital Assets Regulation State Of Play*. IFC Review. Retrieved 2 September 2024, from https://www.ifcreview.com/articles/2024/may/singapore-digital-assets-regulation-state-of-play/

Corrêa, N. K., Galvão, C., Santos, J. W. et al. (2023). Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance. *Patterns*, *4*(10), 100857. doi:10.1016/j.patter.2023.100857

Goh, Y. H. (2023, December 4). $70m S'pore AI initiative to develop first large language model with South-east Asian context. *The Straits Times*. Retrieved 26 April 2024, from https://www.straitstimes.com/singapore/70m-s-pore-ai-initiative-to-develop-first-large-language-model-with-south-east-asian-context

Haziq Mahmud, A. (2022, January 20).*Home Team's robotic dog could respond to chemical incidents, support police patrols when it joins frontline operations*. Channel News Asia.

How we can help. (n.d.). Infocomm Media Development Authority. Retrieved 29 April 2024, from https://www.imda.gov.sg/how-we-can-help?page=1&keyword=&category=all&support=organisations&orgType=all&industry=all&persona=all

Huff, W. G. (1995). What is the Singapore model of economic development? *Cambridge Journal of Economics*, *19*(6), 735–759. https://www.jstor.org/stable/23600207

IMDA, & AICADIUM. (2023a, June 7). *Generative AI: Implications for Trust and Governance*. AI Verify Foundation. Retrieved 2 May 2024, from https://aiverifyfoundation.sg/downloads/Discussion_Paper.pdf

IMDA. (2023b, October 31). *First of its kind Generative AI Evaluation Sandbox for Trusted AI by AI Verify Foundation and IMDA*. Infocomm Media Development Authority. Retrieved 2 May 2024, from https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2023/generative-ai-evaluation-sandbox

IMDA. (2023c, December 4). *Singapore pioneers S$70m flagship AI initiative to develop Southeast Asia's first large language model ecosystem catering to the region's diverse culture and languages*. Infocomm Media Development Authority. Retrieved 2 May 2024, from https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2023/sg-to-develop-southeast-asias-first-llm-ecosystem

IMDA launches Generative AI x Digital Leaders initiative to provide businesses access to GenAI. (2024, March 1). https://www.mci.gov.sg/files/Press%20Releases%202024/AI__Annex_D.pdf

IMDA, P. D. P. C., Lee Kuan Yew Centre for Innovative Cities. (2020, December 4). *A Guide to Job Redesign in the Age of AI*. Infocomm Media Development Authority. Retrieved 1 May 2024, from https://file.go.gov.sg/ai-guide-to-jobredesign.pdf

IMDA, & EnterpriseSG. (2024, February 7). *Generative Artificial Intelligence Sandbox for SMEs*. Infocomm Media Development Authority. Retrieved 2 May 2024, from https://services2.imda.gov.sg/CTOaaS/Highlight/42/generative-artificial-intelligence-sandbox-for-smes

ISO/IEC JTC 1/SC 42 Artificial Intelligence. (2017). International Organization for Standardization. Retrieved 2 May 2024, from https://www.iso.org/committee/6794475.html

Khanal, S., Zhang, H., & Taeihagh, A. (2024). Building an AI ecosystem in a small nation: Lessons from Singapore's journey to the forefront of AI. *Humanities and Social Sciences Communications*, *11*(1), 866. doi:10.1057/s41599-024-03289-7

Li, C. (2006). The Confucian Ideal of Harmony. *Philosophy East and West*, *56*(4), 583–603. doi:10.1353/pew.2006.0055

Lim, R. (2023, May 23). *Perspectives on Generative AI*. Blackdot Research. Retrieved 1 May 2024, from https://blackdotresearch.sg/generative-ai/

Luna, J., Tan, I., Xie, X. et al. (2024). Navigating Governance Paradigms: A Cross-Regional Comparative Study of Generative AI Governance Processes & Principles. (Version 1). arXiv. doi:10.48550/ARXIV.2408.16771.

Lung, T. (2024, April 17). *Artificial Intelligence and the Future of Singapore's Foreign Workforce*. Centre for Strategic and International Studies. Retrieved 2 May 2024, from https://www.csis.org/blogs/new-perspectives-asia/artificial-intelligence-and-future-singapores-foreign-workforce#:~:text=At%20the%20same%20time%2C%20Singapore,join%20its%20workforce%20by%202025

Marken, S., & Nicola, T. (2023, September 13). Three in Four Americans Believe AI Will Reduce Jobs. *Gallup*. Retrieved 2 May 2024, from https://news.gallup.com/opinion/gallup/510635/three-four-americans-believe-reduce-jobs.aspx,%20accessed%2001/27/2023.

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

MAS-led Industry Consortium Releases Toolkit for Responsible Use of AI in the Financial Sector. (2023, June 26). Monetary Authority of Singapore. https://www.mas.gov.sg/news/media-releases/2023/toolkit-for-responsible-use-of-ai-in-the-financial-sector

Ministry of Communication and Information, IMDA, & PDPC. (2022, May). *Annex B: Background on Singapore's AI Governance Work*. Infocomm Media Development Authority. Retrieved 2 May 2024, from https://www.imda.gov.sg/-/media/Imda/Files/News-and-Events/Media-Room/Media-Releases/2022/05/Annex-B—Background-on-SG-AI-Governance-Work.pdf

Model Artificial Intelligence Governance Framework Second Edition. (n.d.). Personal Data Protection Commission Singapore. Retrieved 6 October 2022, from https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf

National AI Strategy: AI for the Public Good, for Singapore and the World. (n.d.). Smart Nation Singapore. Retrieved 2 May 2024. from https://www.smartnation.gov.sg/nais/

OECD AI Principles overview. (n.d.). OCED.AI Policy Observatory. Retrieved 2 May 2024, from https://oecd.ai/en/ai-principles

Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector. (2019, February 7). Monetary Authority of Singapore. https://www.mas.gov.sg/~/media/MAS/News%20and%20Publications/Monographs%20and%20Information%20Papers/FEAT%20Principles%20Final.pdf

Sea-Lion. (n.d.). AI Singapore. Retrieved 2 May 2024, from https://aisingapore.org/aiproducts/sea-lion/

Shared Values. (1991). National Library Board Singapore (Singapore, Parliament) 1 of 1991. Retrieved 30 September 2024, from https://catalogue.nlb.gov.sg/search/card?recordId=5809358

Leng, S., & Fong, K. (2023, August 7). *2023/65 "ASEAN's New Dilemma: Managing the Artificial Intelligence Space."* ISEAS Yusof Ishak Institute. https://www.iseas.edu.sg/articles-commentaries/iseas-perspective/2023-65-aseans-new-dilemma-managing-the-artificial-intelligence-by-kristina-fong-siew-leng/

Singapore launches AI Verify Foundation to shape the future of international AI standards through collaboration. (2023, June 7). Infocomm Media Development Authority. https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2023/singapore-launches-ai-verify-foundation

Singapore launches world's first AI testing framework and toolkit to promote transparency; Invites companies to pilot and contribute to international standards development. (2022, May 25). Infocomm Media Development Authority. https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2022/sg-launches-worlds-first-ai-testing-framework-and-toolkit-to-promote-transparency

Singapore Releases Asia's First Model AI Governance Framework. (2019, January 23). https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/archived/imda/press-releases/2019/singapore-releases-asias-first-model-ai-governance-framework

Singapore's Approach to AI Governance. (n.d.). Personal Data Protection Commission Singapore. Retrieved 5 July 2023, from https://www.pdpc.gov.sg/help-and-resources/2020/01/model-ai-governance-framework

Singaporeans across generations raise concern on impact of AI in job security. (2024, February 21). *Singapore Business Review*. https://sbr.com.sg/hr-education/news/singaporeans-across-generations-raise-concern-impact-ai-in-job-security

Soh, J., Lim, L., & Yeong, Z. K. (Forthcoming). *Artificial Intelligence in the Regulatory Wonderland (Draft on file with the authors)*.

Song, B., Ed. 2021. *Intelligence and Wisdom: Artificial Intelligence Meets Chinese Philosophers*. Singapore: Springer Nature. doi:10.1007/978-981-16-2309-7

Tan, A. (2024, February 5). *Sea-Lion explained: Southeast Asia's first large language model*. Computer Weekly. Retrieved 1 May 2024, from https://www.computerweekly.com/feature/Sea-Lion-explained-Southeast-Asias-first-large-language-model

Tech Kaki Community. (2024, April 18). GovTech Singapore. https://www.tech.gov.sg/products-and-services/tech-kaki-community/

Trusted Data Sharing Framework. (n.d.). Infocomm Media Development Authority. Retrieved 28 April 2024, from https://www.imda.gov.sg/how-we-can-help/data-innovation/trusted-data-sharing-framework

Infocomm Media Development Authority. **Verify Foundation, A. I.** (2023). *Cataloguing LLM Evaluations*. AI Verify Foundation. https://aiverifyfoundation.sg/downloads/Cataloguing_LLM_Evaluations.pdf

Veritas Initiative: What it is. (2023, October 26). Monetary Authority of Singapore. https://www.mas.gov.sg/schemes-and-initiatives/veritas

Veritas, Trust in AI - Document 5: From Methodologies to Integration. (n.d.). Retrieved 2 May 2024, from https://www.mas.gov.sg/-/media/mas/news/media-releases/veritas-document-5—from-methodologies-to-integration.pdf

Veritas-Toolkit. (n.d.). https://github.com/veritas-toolkit/

Von Kameke, L. (2024, January 9). *Share of consumers seeing mostly the benefits of artificial intelligence (AI) in the Asia-Pacific region in 2023, by country*. Statista. Retrieved 1 May 2024, from https://www.statista.com/statistics/1427977/apac-attitude-towards-benefits-of-ai-by-country/

https://doi.org/10.1017/cfl.2024.12 Published online by Cambridge University Press

What is AI Verify? (n.d.). AI Verify Foundation. Retrieved 17 April 2024, from https://aiverifyfoundation.sg/what-is-ai-verify/

William, W. (2024, March 9). AI-powered humanoid robot Sophia wows crowd at ST Education Forum. *The Straits Times*. Retrieved 2 May 2024, from https://www.straitstimes.com/singapore/world-s-first-humanoid-robot-sophia-wows-crowd-at-st-education-forum

Cite this article: Allen J.G, Loo J and Campoverde J.L.L. (2025). Governing intelligence: Singapore's evolving AI governance framework. *Cambridge Forum on AI: Law and Governance* **1**, e12, 1–20. https://doi.org/10.1017/cfl.2024.12