P03 / Governance and Policy Design Lessons for Trust Building and Responsible Use of AI, Autonomous Systems and Robotics

Topic : SCIENCE, INTERNET AND TECHNOLOGY POLICY

Chair : Araz Taeihagh (National University of Singapore)

Second Chair : Devyani Pande (National Law School of India University)

Third Chair : Shaleen Khanal (The University of Nottingham Ningbo China)

Fourth Chair : Lili Li (Auburn University)

GENERAL OBJECTIVES, RESEARCH QUESTIONS AND SCIENTIFIC RELEVANCE

Artificial intelligence (AI), Autonomous Systems (AS) and Robotics are key features of the fourth industrial revolution, and their applications are supposed to add $15 trillion to the global economy by 2030 and improve the efficiency and quality of public service delivery (Miller & Sterling, 2019). A McKinsey global survey found that over half of the organisations surveyed use AI in at least one function (McKinsey, 2020). The societal benefits of AI, AS, and Robotics have been widely acknowledged (Buchanan 2005; Taeihagh & Lim 2019; Ramchurn et al. 2012), and the acceleration of their deployment is a disruptive change impacting jobs, the economic and military power of countries, and wealth concentration in the hands of corporations (Pettigrew et al., 2018; Perry & Uuk, 2019).

However, the rapid adoption of these technologies threatens to outpace the regulatory responses of governments around the world, which must grapple with the increasing magnitude and speed of these transformations (Taeihagh 2021). Furthermore, concerns about these systems' deployment risks and unintended consequences are significant for citizens and policymakers. Potential risks include malfunctioning, malicious attacks, and objective mismatch due to software or hardware failures (Page et al., 2018; Lim and Taeihagh, 2019; Tan et al., 2022). There are also safety, liability, privacy, cybersecurity, and industry risks that are difficult to address (Taeihagh & Lim, 2019) and The opacity in AI operations has also manifested in potential bias against certain groups of individuals that lead to unfair outcomes (Lim and Taeihagh 2019; Chesterman, 2021).

These risks require appropriate governance mechanisms to be mitigated, and traditional policy instruments may be ineffective due to insufficient information on industry developments, technological and regulatory uncertainties, coordination challenges between multiple regulatory bodies and the opacity of the underlying technology (Scherer 2016; Guihot et al. 2017; Taeihagh et al. 2021), which necessitate the use of more nuanced approaches to govern these systems. Subsequently, the demand for the governance of these systems has been increasing (Danks & London, 2017; Taeihagh, 2021).

CALL FOR PAPERS

Many studies have highlighted the urgency for and the challenges of governing AI, AS and Robotics (Firlej and Taeihagh 2021; He et al. 2020; Tan and Taeihagh 2021; Tan et al. 2021; Radu 2021; Taeihagh 2021). In this panel, we are interested in governance and policy design lessons for Responsible Use and Building trust in AI, AS and Robotics by answering the following key research questions:

· What governance and policy design lessons have been learnt so far in addressing risks and unintended consequences of adopting AI, AS and Robotics in different domains and geographies?

· What are the challenges of responsible use of AI, AS and Robotics, particularly in the public sector?

· What are the emerging theoretical, conceptual and empirical approaches to understanding new and unconventional regulatory approaches, governance strategies, institutions and discourses to govern these systems?

· What lessons have been learnt so far from the public and private organisations' standard setting and

development of guidelines in managing these systems?

· How can the public and expert viewpoints be better considered for the regulation and governance of AI, AS, and Robotics to increase trust in AI?

· What is the role of governments in promoting trustworthy AI and building trust in AI?

REFERENCES

Buchanan, B. G. (2005). A (very) brief history of artificial intelligence. Ai Magazine, 26(4), 53.

Chesterman, S. (2021). Through a Glass, Darkly: Artificial Intelligence and the Problem of Opacity. The American Journal of Comparative Law, 69(2), 271-294.

Danks, D. (2019). The value of trustworthy AI. Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, 521–522.

Guihot, M., Matthew, A. F., & Suzor, N. P. (2017). Nudging Robots: Innovative Solutions to Regulate Artificial Intelligence. Vand. J. Ent. & Tech. L., 20, 385.

He, H., Gray, J., Cangelosi, A., Meng, Q., McGinnity, T. M., & Mehnen, J. (2020). The Challenges and Opportunities of Artificial Intelligence for Trustworthy Robots and Autonomous Systems. 2020 3rd International Conference on Intelligent Robotic and Control Engineering (IRCE), 68–74.

Lim, H. S. M., & Taeihagh, A. (2019). Algorithmic decision-making in AVs: Understanding ethical and technical concerns for smart cities. Sustainability, 11(20), 5791.

McKinsey. (2020, November 17). Global survey: The state of AI in 2020 | McKinsey.https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/global-survey-the-state-o

Miller, H., & Stirling, R. (2019) The Government Artificial Intelligence (AI) Readiness Index Report 2019.https://www.oxfordinsights.com/ai-readiness2019

Page, J., Bain, M., & Mukhlish, F. (2018, August). The risks of low level narrow artificial intelligence. In 2018 IEEE International Conference on Intelligence and Safety for Robotics (ISR) (pp. 1-6). IEEE.

Perry, B., & Uuk, R. (2019). AI governance and the policymaking process: key considerations for reducing AI risk. Big Data and Cognitive Computing, 3(2), 26.

Pettigrew, S., Fritschi, L., & Norman, R. (2018). The potential implications of autonomous vehicles in and around the workplace. International Journal of Environmental Research and Public Health, 15(9), 1876.

Radu, R. (2021). Steering the governance of artificial intelligence: national strategies in perspective. Policy and Society, 40(2), 178-193.

Ramchurn, S. D., Vytelingum, P., Rogers, A., & Jennings, N. R. (2012). Putting the 'smarts' into the smart grid: a grand challenge for artificial intelligence. Communications of the ACM, 55(4), 86-97.

Scherer, M. U. (2015). Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. Harv. JL & Tech., 29, 353.

Taeihagh, A. (2021). Governance of artificial intelligence. Policy and Society, 40(2), 137-157.

Taeihagh, A., & Lim, H. S. M. (2019). Governing autonomous vehicles: emerging responses for safety, liability, privacy, cybersecurity, and industry risks. Transport reviews, 39(1), 103-128.

Taeihagh, A., Ramesh, M., & Howlett, M. (2021). Assessing the regulatory challenges of emerging disruptive technologies. Regulation & Governance, 15(4), 1009-1019.

Tan, S.Y., & Taeihagh, A. (2021). Adaptive governance of autonomous vehicles: Accelerating the adoption of disruptive technologies in Singapore. Government Information Quarterly, 38(2), 101546.

Tan, S.Y., Taeihagh, A., & Tripathi, A. (2021). Tensions and antagonistic interactions of risks and ethics of using robotics and autonomous systems in long-term care. Technological Forecasting and Social Change, 167, 120686.

Tan, S., Taeihagh, A., & Baxter, K. (2022). The Risks of Machine Learning Systems. arXiv preprint arXiv:2204.09852.

P03 / Governance and Policy Design Lessons for Trust Building and Responsible Use of AI, Autonomous Systems and Robotics

Chair : Araz Taeihagh (National University of Singapore)

Second Chair : Devyani Pande (National Law School of India University)

Third Chair : Shaleen Khanal (The University of Nottingham Ningbo China)

Fourth Chair : Lili Li (Auburn University)

Session 1

Wednesday, June 28th 10:20 to 12:20 (KHE118A)

Discussants

Shaleen Khanal (The University of Nottingham Ningbo China)

Kritika Sha (National University of Singapore)

(Virtual) Artificial Intelligence (AI) for Peace and Stability

Yihan CAI

Xun Wu

Geopolitical conflicts have been a source of tension and instability in recent years as the world order is being reshaped by emerging global powers. While the advances in Artificial Intelligence (AI) can potentially contribute to peace and stability by harnessing decades, if not centuries, of accumulated knowledge, such benefits cannot be taken for granted. In fact, the use of AI may reinforce the misperception and misunderstanding among nation states, exacerbating tensions and increase the risk of cyber-espionage and cyber-warfare. In this paper, we use Chatgpt, a natural language processing tool driven by AI technology, to test hypotheses on ideological positions revealed through dialogue on key issues in geopolitical conflicts, such as the rise of China, Russia-Ukraine war, the origin of COVID-19 pandemic. In particular, we conduct analysis based on responses of Chatgpt by varying languages and localities to examine the extent to which AI tools would react to different scenarios. Our results show that the role of AI in this context is complex and multifaceted, as it could potentially be used to mitigate conflict and promote cooperation among participating countries, but at the same time, it can also create a gap we call ‘AI-augmented bias’ to people around the globe. Our study points to a new dimension in which policy and technical responses can be critical in shaping the power of AI.

Local approaches to governance of AI systems in China: Challenges of adopting national AI policy

Shaleen Khanal (The University of Nottingham Ningbo China)

Hongzhou Zhang (The S. Rajaratnam School of International Studies)

Araz Taeihagh (National University of Singapore)

How do local governments respond to vertically imposed policies on emerging technologies? This paper answers this question by examining vertical policy diffusion of artificial intelligence in China. We will use the case of China’s New Generation Artificial Intelligence Development Plan (AIDP) to build a unique dataset of China’s local and provincial AI-related policies that cite the AIDP and examine the factors that affect the speed and the nature of vertical policy diffusion. To understand the determinants of speed of policy diffusion, we will conduct a dyadic regression analysis that will identify the factors that shape the spatial transmission of policy across various provinces. Additionally, to understand the nature of policy transmission, we will examine the width and depth of policy design and policy mixes of individual provinces.

Our paper will contribute to the emerging field of policy design literature by examining the factors that shape local governments' policy mixes to the govern an emerging technology. Our paper will also contribute to the literature in policy diffusion, which has mostly analyzed horizontal forms of diffusion across political spaces but has paid inadequate attention to vertical diffusion, by examining local governments' response to a central government's mandate.

The Impact of Cultural Factors on Citizen Acceptance of Public Sector Applications and Governance of Artificial Intelligence

Ashley Khor (University of Pittsburgh)

The rise of automation, big data and artificial intelligence (AI) has resulted in public sector agencies experimenting with the new technologies, from piloting predictive policing programs and facial recognition technologies to implementing digital contact tracing applications during the COVID-19 pandemic. In spite of substantial investment in research and development, societal acceptance of such technologies has been mixed. In the case of digital contact tracing, country-specific effects have also been observed with countries such as Qatar and Singapore reporting much higher rates of downloads and adoption of digital contact tracing applications as compared to other countries. At the same time, evidence on citizens' perception of public sector use of artificial intelligence is sparse and disparate (Konig, 2022). There is also a gap in research on how national culture influences public acceptance or the adoption of privacy protections (Thompson, McGill, Bunn and Alexander, 2020).

As the public sector continues to embrace the potential of AI and digital technologies to transform the way it provisions citizen services, governance and new policy solutions are becoming increasingly imperative. Studying public opinion of public sector AI applications as well as their motivations to coproduce with government is therefore important to understand how government responses may impact on social acceptance. Several studies have begun to examine citizens' AI perceptions which have primarily centered on European Union members (European Commission, 2017 and 2020; Grzymek and Puntschuh, 2019; Araujo et al., 2020; Starke and Lünich, 2020) and the United States (Smith, 2018; Konig, 2022), as well as predominantly focus on private sector applications. There remains major gaps regarding what is known about social acceptance for AI in government (Konig, 2022). Some evidence is emerging that transparency can lead to an increase in social acceptance (Aoki 2020; Grimmelikhuijsen 2022; Kennedy, Waggoner, and Ward 2022; Miller and Keiser 2021; Schiff, Schiff, and Pierson 2021). On the other hand, Konig (2022) found that in Germany, citizens are willing to trade away transparency and stakeholder involvement for small effectiveness gains in the case of policing and healthcare.

This paper proposes to contribute to the existing literature by conducting a cross-country analysis of citizen perceptions towards public sector applications of AI as well as proposed governance and policy solutions via a mixed methods approach. Firstly, it seeks to analyze the relationship between cultural factors and social acceptance, such as by testing whether Hofstede's model of national culture predicts citizen approval. Secondly, it proposes to extend existing studies to the lesser explored region of Southeast Asia, including Singapore, Thailand and Malaysia. One approach could be to leverage conjunct analysis to study trade-off decisions such as effectiveness and safety and security vis-à-vis transparency and privacy. The insights from this study are intended to contribute to the understanding of what governance and policy design lessons have been learnt so far in addressing risks and unintended consequences of adopting AI as well as how public and expert viewpoints be better considered for the regulation and governance of AI to increase citizen trust and societal acceptance.

Comparative Analysis of Governance of Biometrics in AI systems

Lynnette Hui Xian Ng (Carnegie Mellon University)

Abigail Lim (National University of Singapore)

Araz Taeihagh (National University of Singapore)

Biometric data refers to data derived from human anatomical features, such as fingerprints, facial features, and iris features. Over the years, using these biometric features in Artificial Intelligence (AI) applications has grown precipitously. Such applications include facial recognition systems, phone voice recognition technology, and even face enhancement technologies on social media videos. Technologies harnessing biometric data have been used in government organisations for identification, personal authentication devices, by gaming and social media technologies for entertainment. Since biometric data encompasses physiological and behavioural characteristics, there must be governance for the responsible use of these systems.

Although there is an increasing use of biometric data within AI systems, there is no overarching framework to govern the use and protection of biometric data at the international level. Different countries have enacted their own laws for the regulation of personal biometric data. In response to the expansion of the range of purposes of biometric data, the United States and the European Union have enacted biometric-related laws, in particular, the European Union Artificial Intelligence Act, the Illinois Biometric Information Privacy Act (BIPA) and the European Union's General Data Protection Regulation (GDPR). These three most prominent biometric regulatory legislations provide a framework for impositions of the definitions of "biometric data" and the interpretations of responsible use of biometric data within AI systems.

In our work, we present a comparative analysis of the three laws. We evaluate the regulations for their strength and weaknesses in terms of the following issues: (i) the comprehensiveness of the definition of biometric data; (ii) the purposes to which biometric data is being tied to in the legislation; (iii) the rights and fundamental freedoms of biometric data subjects as a result of their usage and legislations; and (iv) the suitability of existing legislations in addressing evolving issues such as artificial intelligence, social media and disinformation. Finally, we propose recommendations on the governance of the collection, use, processing, and disclosure of biometric data in AI systems. We hope that our analysis and recommendations towards using and protecting biometric data will steer responsible use of biometric data within AI applications, building trust in organisations that are harnessing the data to create new technologies.

Facing Frankenstein: bringing back democratic governance in AI in electricity systems

Irene Niet (University of Technology Eindhoven)

Various actors in the energy sector are increasingly integrating artificial intelligence (AI) in their systems. The digital technology can support system management, more accurately predict and match energy production and consumption, and assist in infrastructural decisions. The governance of the integration of AI in the electricity system is, however, complex.

This research analyses how non-traditional governance approaches could navigate the growing complexity of AI governance in the energy sector. Specifically, in this article, the frameworks of governance ecosystems and polycentric governance are combined.

The governance ecosystem framework refers to the idea that for effective governance, four domains need to come together in achieving public goods and services: the science and technology domain, the civil society domain, the laws and regulation domain, and the politics and administration domain (Arentsen & Est, 2023; Kool, Timmer, Royakkers, & Est, 2017). Each of these domains, which can include actors from both the public and private sector, carries a specific requirement of effective democratic decision-making. The science and technology domain brings forward scientific knowledge and technological feasibility. The civil society domain can supply social desirability and acceptance, the laws and regulations domain grants legal admissibility, and, finally, the politics and administration domain can provide political legitimacy and political acceptance (Arentsen & Est, 2023).

In line with this, polycentric governance strongly argues for the importance of having multiple centres of decision making. These centres should be relatively autonomous, aimed at self-governance, and competing or collaborating with each other under an overarching shared system of rules (Zeben & Bobi?, 2019). These centres should thus be supported in their governing capacity, for example by having access to information to make decisions, increase their governance skills, and learn from other decision-making centres.

I argue that supporting a governance process in which multiple centres of decision-making act together, strengthens the democratic character of the governance of AI in the energy sector, as multiple stakeholders are more concretely involved in those decisions directly related to them and problems and solutions are based on multiple centres of knowledge and experience. At the same time, this governance approach avoids the bureaucratic inertia often connected to traditional governance institutions. It also allows the traditional, international governance institutions to focus on their expertise: ensuring there is a shared framework of norms and values, with enough space for actors to create their own set of regulations, specified for their sector or local context.

To test this argument, this research focuses on the case of the integration of AI in the Dutch electricity sector. The integration of AI in (parts of) the electricity system is technically feasible, but its impact on society as well as the broader electricity system is unclear. Governance has proven difficult, actors from different levels and sectors of governance are involved. In a combined governance ecosystem framework and polycentric governance approach, these actors could be involved in a balanced way, without losing the aim of guiding AI integration in the electricity system towards creating and improving public goods and services.

References

Arentsen, M., & Est, R. V. (Eds.). (2023). The Future of Radioactive Waste Governance: Lessons from Europe (1st ed. 2023 edition). New York: Springer VS.

Kool, L., Timmer, J., Royakkers, L., & Est, R. van. (2017). Opwaarderen—Borgen van publieke waarden in de digitale samenleving. Den Haag: Rathenau Instituut.

Zeben, J. van, & Bobi?, A. (Eds.). (2019). Introduction: The Potential of a Polycentric European Union. In Polycentricity in the European Union (1st ed., pp. 1–6). Cambridge: Cambridge University Press.