KONRAD ADENAUER STIFTUNG

Mass Data Sharing in Smart Cities

Regulation and governance of the public-private interface

A better city for tomorrow.
A better life for all.

Notice at King Albert Park MRT Station in Singapore, advertising 'World Cities Day 2023'.

SMU
SINGAPORE MANAGEMENT UNIVERSITY

Centre for
AI and Data Governance

ISBN: 978-981-18-9628-6

Published by:
Konrad-Adenauer-Stiftung Ltd
Rule of Law Programme Asia
Address: ARC 380, 380 Jalan Besar, #11-01 Singapore 209000
Phone: (+65) 6603 6171
Website: https://www.kas.de/rlpasia

Centre for AI & Data Governance
Address: 55 Armenian Street, Singapore 179943
Phone: (+65) 6828 0753
Website: caidg.smu.edu.sg

Authors: Dr. Berenika Drazewska and Prof. Mark Findlay

Designed by: Dexel Pte Ltd

Table of Contents

Note to the Reader

With the rapid convergence of technology and urbanization, Smart Cities represent the archetype of innovation and efficiency. However, as we navigate the complex landscape of urban development, public and private partnerships become increasingly critical and underlie significant matters of data governance.

The report you hold in your hand, titled 'Mass Data Sharing between Private and Public Partnerships in Smart Cities' comes from an original initiative spearheaded by the Konrad-Adenauer-Stiftung's Rule of Law Program Asia.

This report showcases the Rule of Law Program Asia's commitment to fostering an environment where cutting-edge technology can unite with robust legal and governance frameworks, ensuring its alignment with principles of justice, transparency and accountability. It offers a deep dive into the dynamics of mass data sharing in smart cities and examines the blurry relationship between private and public entities.

This publication would not have come to fruition without the rigor and expertise of the two distinguished authors, Dr. Berenika Drazewska and Prof. Mark Findlay. Their insightful analyses and research have shaped a document that presents the challenges at hand, but also paves the way for citizen-centric and informed decision-making.

We would also like to thank (in no particular order): Mabel Choo, Kasper Drazewski, Poomsiri Dumrongvute, Calum Handforth, Manoj Harjani, Shazade Jameson, Erika Legara, Yong Lim, Jane Loo, Anisa Pratita Kirana Mantovani, Aishwarya Natarajan, Sharanya Shanmugam, Nidhi Singh and PeiChin Tay, as well as the individuals who kindly agreed to support this research, but wished to remain anonymous.

Acknowledgement is also due to the Dean of the Law Faculty of the Singapore Management University and its Centre for Data and AI Governance (CAIDG) for their invaluable support and commitment in facilitating the research endeavours enclosed in this report.

As we embrace the transformative potential of Smart Cities, the Konrad Adenauer Stiftung's Rule of Law Program Asia invites policymakers, academics and practitioners to engage with the findings of this report. In doing so, we collectively contribute to the creation of urban landscapes that are not only technologically advanced, but also supported by principles of justice, accountability and the rule of law.

Stefan Samse, Director of the Rule of Law Program Asia
Olivia Schlouch, Program Manager at the Rule of Law Program Asia
Singapore, March 2024

Executive Summary: Getting Data Better Means Getting Better Data

There are at least two ways of understanding the importance of this Report and its implications. The essential research purpose was to examine the nature of mass data sharing between private and public agencies in the commerce and administration of certain smart cities. With this knowledge the research speculated on and selectively exposed the governance challenges posed by this sharing for stakeholders, citizen/residents in particular, in various data relationships and arrangements. Predicting that good data governance policy and practices can address these challenges, the Report proposes a model strategy that grows from commitments where stakeholders will employ trusted data spaces to create respectful and responsible data relationships, where the benefits of data sharing can also be achieved without compromising any stakeholder interests.

At a broader level, and in the context of an ongoing data sharing explosion worldwide, the significance of good data gathering and using environments is assuming major proportions. The Report could thus be seen as an argument for the recognition of citizen/residents and their interests as necessary determinants of good governance, when private and public data management merge. And beyond this there is cause to consider what benefits the private and public agencies ultimately want from the data sharing, and whether current sharing practices threaten the outcomes that all data stakeholders need for urban life experiences to develop positively. The Report details worrying instances of stakeholders lacking sufficient information about its extent and consequences – and they are not only the data subjects, but sometimes also the public and private agencies involved in the sharing.

During the COVID pandemic surveillance regimes, breaches in data protection led to loss of public trust. This resulted in the sapping of public confidence in the exercise of control policy relying on co-operation with the data subject, and probity of the data processing. If trust among the data subjects deteriorates, platforms that rely on extensive data access and reuse will find it more difficult to continue data harvesting even where customers and clients are required to contribute data for receiving urban services and commercial transactions. The compulsory provision of personal data without which a growing range of urban services will not be available to citizen/residents compromises the buffer of consent, particularly when assumed consent is not based on information about reuse, and will lead to the damaging of relationships of trust between the platforms and their users. By the same token, the loss of potentials for data subjects to validate their data will have negative ramifications for data quality and integrity.

Imagine a smart city without mass data sharing and digital platform service delivery – that is simply not possible. Then think about the same city operating with data of questionable quality, where data subjects know little or nothing about what happens to their data, and yet are compelled to receive services through platforms that operate with inadequate data governance. These two scenarios are far from hypothetical. The Report instances how challenges to governance posed by mass data sharing require action if the role that data can play in urban development and city life experiences is to be positive and responsible.

So, what is new about this analysis? Admittedly, public-private partnerships have been a feature of urban administration for a long time. However, this report is not concerned with the public sector divesting its obligations to private agencies through private law contracting, and the dangers this has posed for openness and accountability in general. The research is centred on how mass data shared between and beyond public and private agencies adds a new consideration when looking for accountable, transparent and responsible data governance. In smart cities, personal data fuels these public/private administrative and commercial arrangements. The influence exerted by big info-tech providers in how data drives smart cities is a variable

when examining governance questions. These large, often absentee (offshore) commercial entities are creating and operating platforms that draw together in one portal many different public administrative services that had jealously kept their data pools and pathways separate before digital convergence. Citizen/residents are being induced or compelled to use these super apps and in so doing provide all types of personal data, the reuse of which is typically obscure. Data is not like other contracted 'valuables'. It can be limitless and almost impossible to contain. Additionally, as the platform reliance on personal data and its 'on-selling' grows, the concern is that data accuracy and the protection of data integrity are becoming casualties. In turn, a voracious appetite for unruly data poses a challenge to the sustainability of these platforms both in terms of data subject confidence, as well as internal validation.

Another issue, largely the focus of Part C of the Report, is how conventional governance based on the principles of data protection, designed as it is to deal with the rights and interests of individuals against the use of identified data harvesters, is not well-positioned to impact the governance of mass data sharing. The problem of mass data sharing is compounded by the collapsing of public and private dimensions and responsibilities in the handling and control of data, and the most fundamental governance separations like public and private law become blurred. This convergence introduces novel questions, like what principles do and should inform the governance options that in Asian smart cities are heavily skewed to a reliance on voluntary compliance from the private sector, and 'soft-touch' public sector regulation reluctant to damage business interests at the heart of data trading.

Smart cities are surveillance-heavy. Much of this surveillance has justifiable purposes and could be appreciated as helping ensure the quality of urban life experience. That said, the bulk of it is pervasive, clandestine and infiltrates into even the most mundane activities of daily living. In democratic, representative political traditions, including those that may have executive delivery, the public sector owes duties to citizen/residents for the handling of surveillance data. Yet, the reality of surveillance in smart cities is that most of the tech infrastructure and digital processing of surveillance data is outsourced to the private sector, where such obligations may be more private law oriented and diffuse, as data is shared across surveillance networks and beyond.

The research focused on smart cities in Asia, recognizing how this region has driven digitized urban development across diversified economic and social terrain. Experiences of smart city data sharing in other locations were also drawn upon. While it would be impossible to cover every major instance of public/private data sharing in urban surveillance and platform delivery, those that have been selected for the empirical examples in the following Report are indicative of contextual variables and common universal characteristics when it comes to governance challenges. All of the examples chosen were significant for the administration and commercial frames of smart cities, and some reached beyond urban locations and were impactful country-wide. Again, the peculiarity of smart city data management was the extent to which digital infrastructures relied on mass data sharing.

The Report is structured in three parts, with their progression and inter-reliance depending on testing and confronting the initial analytical concerns regarding governance and mass data sharing. Part A introduces research purposes and talks in brief and simple terms about the theoretical and method tools used to unpack the assumptions driving the research mission. Working in an area of social policy research as yet underdeveloped in the governance literature required the research team to draw from many disciplines and to grapple with terms and concepts that, while often complex, require reducing down to understandings that both a general readership and that specialized in other fields can manage. This part leads onto understanding the empirical exercise in Part B.

The second part (Part B) adopts an 'orbits and snapshots' language. Due to declared constraints that required the research to have a suggestive rather than representative direction, the team selected several important broad fields of platform service delivery and/or surveillance to provide windows into how data sharing operated in action on the ground, i.e., the 'orbits'. 'Snapshots' are more specific applications of data sharing located in instances of platform data management and surveillance. They are referred to as snapshots not to imply that the analysis is static, but rather that it tended to look at selected features of an urban administrative/commercial data pathway from angles that are especially important for the interests of citizen/residents. While not intended to speak for all data sharing in its wider orbit, individual snapshots reveal important challenges particular to their community context and, when juxtaposed with other snapshots treated in similar fashion, more universal challenges orbit-to-orbit. As an example of this method, the several snapshots that deal with devices and platforms that share health-related data reveal insights into data sharing in the most specific and localized contexts, then extending to general governance issues for other snapshots in the same orbit and in other orbits where common challenges are revealed. From smartphone monitoring to the delivery of more generalised health coverage, these snapshots reveal how citizen/residents are ill-informed about what of their data is shared, how and with whom. Contributors of data are induced to provide it for different reasons, some of these changeable and influencing the nature of their initial consent. Health data is among the most valued by data subjects and yet, the policies made available to data subjects are inadequate for them to understand the consequences of participating in these platforms. Part B does not present 'findings' as such, just as part C does not make 'recommendations'. If these are to be considered important, they await more work and a much greater sensitizing across the data ecosystem of the importance of what this report can (and cannot) cover.

Those reading this review might be tempted to skip to Part C of the Report for the policy propositions emerging thereby. We would advise against approaching the model without the benefit of what it is intended to address and divorced from some of the limitations that

foundation research such as this will always work with. Part C presents a model for *getting data better and thereby getting better data*. It could only be a model for several reasons even beyond the suggestive nature of the empirical work. The Report emphasizes the importance of building trusted data relationships and the need for a dispersing of 'data power' particularly across to citizen/residents. These pre-conditions, while crucial, cannot be achieved in a one-size-fits-all mode. Different contexts of data sharing with different stakeholders operating different data arrangements will necessitate that once stakeholders are encouraged to engage and participate in a shared governance exercise then the terms and nature of that exercise will be detailed (and owned) by them. This eventuality is not some vague aspiration. Mutual data strategies as have recently emerged with open finance indicate that co-creation and co-production of trusted data relationships can be achieved even in the seemingly most intractable of environments. A model was further preferred to avoid the appearance of external governance injection or propositions that exhibit some of the same agency-oriented features that are identified in the Report, as well as of typical privacy-focused governance approaches inconsistent with a citizen-centric outcome. While Part C is a model, it is not devoid of detail. There are components presented as vital in ensuring the engagement with and sustainability of a co-production governance experience. An important aspect involves identifying the possible motivations for stakeholders coming from different power bases to get and stay involved. Conflict resolution opportunities are built in as pressure valves recognising that trust can be a fragile bond in governance between various interests and preferred outcomes. A graph is provided to indicate how the different parts interconnect. Part C concludes with a discussion of how the model can be actioned with several important dynamics set out.

What are the **10 takeaways** a reader can expect from the Report?

1. Mass data sharing between public and private agencies is happening in digitized smart cities.
2. Rarely do citizen residents have information that allows them to understand the nature and consequences of this sharing for their data and for themselves.
3. While mass data sharing has positive outcomes for city life experience, it also presents governance challenges.
4. These challenges directly affect citizen residents in terms of accountability, transparency and responsible data practices.
5. Additionally, the public and private agencies are not infrequently ill-informed about data re-use and its impact on data quality and integrity.
6. It is vital that inclusive governance strategies are developed and employed for the sake of the responsible access and use of data which is so important for commercial and administrative purposes in smart cities.
7. Existing data governance policies and practices are not designed to best address the unique challenges posed by mass data sharing.
8. For all stakeholders to benefit from the responsible access and use of data in smart cities, more information openness needs to be ensured regarding data sharing.
9. A co-creation/co-production approach to governing mass data sharing may be helpful in the establishing of trusted, responsible and respectful data relationships.
10. Once these relationships are fostered and maintained, they are key to getting data better which will in turn lead to getting better data.

Preparatory considerations

Some might hold that there is nothing new about governance concerns raised at the intersection of public and private administration. However, the specific context of SE Asian smart cities[1] in which digital reliance, the accumulation, access, management and proliferation of mass quantities of data play out, and the potential of citizen/resident marginalising in data control makes it worthwhile to reflect on the probable shortcomings of conventional public and private governance modes.

This project is not an exercise in rejecting public/private partnerships in the administration and service delivery in smart cities; neither is it disputing the benefits to administrators and residents offered by data-driven technology, or indeed, mass data sharing in such contexts. Instead, based on doctrinal research and an empirical examination of six 'snapshots' which provide limited windows into such arrangements and practices in some SE Asian smart cities, this report ('the Report'):

- Discusses the nature of personal data being shared as a result of developing digital platforms that provide portals for accessing and engaging with some fundamental service areas in modern urban life such as health, education and government information;
- Examines how input data and surveillance data is or could be shared between public and private administrative providers;[2]
- Speculates on the challenges to the integrity of data (defined by the data subject's intention to make her data available for sharing in the original context)[3] posed by such data sharing;
- Looks at these challenges in terms of governance options and responsibilities, and
- Offers a governance model that could address some of these challenges without constraining legitimate benefits from public/private digital administration and mass data sharing consequent on such engagements.[4]

The research moves on from well-examined concerns for the marketizing of personal data through service delivery and surveillance.[5] It interrogates (where possible and appropriate) instances of digitized individual and community-oriented essential service reception in smart cities where private law arrangements and market forces have become enmeshed with the responsibilities that once were clearly the domain of public law and administration. Traditional *separation of powers* governance codes (like judicial review for public law, and contractual standing in private law) are obviously less discrete, or outright insufficient for the task of providing accountability to citizens and their communities in the combined use of their life-space data by public and private agencies. The Report notes the need to explore novel governance options that confront and address the power asymmetries that can arise out of mass data sharing and pervasive surveillance in smart cities.

It is the longer-term goal of this preliminary enquiry to stimulate more expansive research and policy interest in what should be a crucial field of data governance and responsible data management in smart cities. Currently, digital and information governance debates internationally are concerned with the risks posed by AI-assisted technology and enabling algorithms, as well as consequent safety measures for users and society such as data protection and rights or consumer agency and competition law, where these are available and actioned. Responsible and respectful data access and use practices is equally, if not more important when considering the risks posed by digital technology, and the need for safety assurance. Digital technologies require personal data to learn, process and function, and thus data governance should be at the forefront of considering responsible and safe AI. This Report is directed to analysing mass data sharing rather than the technology that makes sharing possible.

[1] For our understanding of smart cities, see p. 17.
[2] As will be developed later, the distinction between public and private in data access and use is no longer hard and fast, but for the purposes of appreciating the distinct governance responsibilities triggered by each of them, it is here retained.
[3] This analysis uses the concept of 'data integrity' rather than privacy or data rights (or even H. Nissenbaum's proposal of 'contextual integrity' as a theory of privacy) as a more comprehensive value for the protection of data. Data integrity here is seen not so much as a quality of data, but rather respect for the original intentions and purposes of data, as communicated by the data subject. In this use context may indeed be relevant to locate the original purpose for the conveying of data, as suggested by Nissenbaum. She premises her theory on the existence of distinct social norms prescribing the proper flow of personal information across different contexts (based on factors such as who is the sender and the recipient, and what is the principle under which the information is sent or transmitted), with a contravention of these norms being qualified as a violation of contextual integrity, and therefore privacy (Nissenbaum, 2009). At the same time, this analysis agrees with N. Couldry and U. Mejias who observe that in modern day data relations the 'contexts' for sharing data are becoming increasingly muddied, and thus an approach based on contextual integrity of data flows may not suffice to address the underlying power asymmetries (Couldry & Mejias, 2019, p. 177).
[4] For an explanation of the modelling methodology, see p. 24.
[5] Whether it be via personal data protection regulation, data rights frames, best practice regimes or ethics principles, governance strategies already exist in many jurisdictions concerned about personal data. The Report will explain why governance concerns about mass data sharing may not be as well met by these prevailing governance strategies, and why additional governance initiatives may be appropriate. In so doing, and due to the focus on positive additions to governance policy, a critique of current data protection approaches that do not sufficiently address mass data sharing is referred to selectively and a supplementary model is provided.

The primary focus of our analysis are certain smart city contexts in Southeast Asia. With wise counsel and creative advice from a group of regional experts who comprised the core of our conversation methodology,[6] we have been able to test some of our hypotheses linked to the challenges that arise in urban settings against the interest of data subjects (citizen/residents) and their communities[7] in other smart cities experiences. Throughout the Report the research team confirms the importance of contextual specificity, balanced against more universal notions of good data practices. In our limited comparative exercise there should be no inference drawn of one data use practice in one smart city being more or less challenging than another. The Report tries to present context-specific data sharing instances which have been influenced by the particular nature of smart city development in that location.[8] From this approach, a more general representation of governance challenges can be observed.

With this understood, it is anticipated that readers beyond Asia will find issues of interest in the Report as it relates to mass data sharing, governance, public-private collaborations and digital urban administration.

[6] For an explanation of the conversation methodology, see p. 23 ff.

[7] Throughout the Report we refer to data subjects, i.e. citizen-residents, and their communities. Two observations are in order here. Firstly, there is always some contention around using what could be seen as a hierarchical notion of 'subjects'. But here it designates nothing more than individuals who originate the purposes and communication of personal data and the communities among which the data is conveyed. Secondly, the category of 'citizen-residents' is not meant to exclude individuals without a given citizenship, or not domiciled within a given city, nor even be geographically limited to urban dwellers. Instead, 'citizen-residents' are here understood in a broad sense as users and beneficiaries of 'smart' infrastructures and technologies in public service delivery and who will often, though not always, be based in urban centres.

[8] For example, the rollout of smart city initiatives in Indonesia has often been hindered by the challenges faced by the nation in its digital transformation. These challenges include a nascent data protection law (Law No. 27 of 2022 on Personal Data Protection ('PDP Law') was enacted into law on 17 October 2022, and is currently in a transition period until October 2024); a general lack of information about the government's 'smart' initiatives among the public; a digital divide which is matched by socio-economic, cultural and geographic divides among the regions, and the limited technological capacity within the government in addition to the lack of necessary infrastructure, which has influenced the government to rely on private companies in the rollout of smart city initiatives. On this see Drazewska (2023b).

Introduction

It could be said that smart cities are 'smart' because of the way digital transformation and mass data sharing have transformed city administration and urban lifestyles. The physical infrastructure of smart cities and the delivery of public and private services now depend on AI-assisted technology and the application of data which is largely generated by citizen/residents (data subjects) going about their daily lives. In this sense 'surveillance' is not only to be understood as using technology to chart and order human behaviour. It also involves accumulating, transacting, and negotiating personal data on a scale hitherto not possible in the operation of urban space.

In conventional city governance, public bodies are subject to administrative regulation (public law), and private agencies respond to the market and its arrangements (private law). However, if this governance distinction was ever clear, it is now becoming blurred by:

- The reliance of public administration on the private sector for the provision of digital technology, and the operation of data surveillance at many levels;
- The confluence of public-private partnerships in the delivery of digitized infrastructure in urban environments;
- Consequent mass data sharing in these partnerships and the data markets they generate;
- Divesting the responsibilities for urban services delivery and the data management that accompanies such developments from state to private sector agencies, as well as
- The fact that private companies which support public agencies through smart technology, ideas and services do not cease their activities as private, profit-driven operators benefitting from the almost limitless potential for data re-use in a smart city context.

Even without these specific data-driven partnerships, the private sector, through digitized commerce and financial transaction platforms, is taking over many of the domains of interest that might traditionally have been viewed as public administration responsibilities, such as resident payments of rates for the provision of institutional services. All this is perhaps an expected by-product of digitized urban living, so why the need for detailed research consideration? The question that the Report interrogates recognises that it is not simply about the merger of state and private sector service delivery, or the strains this puts on conventional governance frames, but that digital transformation and AI-assisted technology enables mass data sharing at scales not previously imagined by conventional governance requirements. The challenges canvassed in Part B reveal that digitized mass data sharing has significant impacts on personal data integrity, citizen inclusion (participation) in city administration, as well as consequent surveillance potentials. Each of these suggests a level of importance to be reflected in governance strategies, even if it may be obscured by the more salient aspects of data access and sharing.

Yet, with the merging of public and private data in such partnerships, the conventional separation between public and private governance domains is challenged, and now likely specious. This novel reality cannot but affect the applicability of key regulatory frameworks and governance strategies, revealing the insufficiency of an approach that focuses on 'public' versus 'private' sector, and 'public' versus 'private' data. To give some examples, the Singapore 'Model Framework', first released in 2019, as updated in 2020 (IMDA & PDPC, 2020) is geared to the private sector, assuming market governance still holds as the active responsibility of private sector data users. On the other hand, the personal data protection (PDP) laws both in Singapore and Malaysia[9] are supposed to target private sector data practices, and not directly address public sector data use. Thus, they do not account for the welding together of governance dimensions, and confluence of data through mass sharing. Additionally, both these models are consensual, top-down and tend to minimise active engagement from data subjects, even though they are the ones who bear the abov ementioned impacts of data sharing and merging.

But the current reality of mass data sharing in smart cities exposes further significant limitations and insufficiencies of data protection and privacy as governance strategies. Personal data protection regulations are generally designed to *limit access* to data so that individual privacy interests are ensured, and legitimate authorised access. This approach is incompatible with smart cities' insatiable appetite for data, and its use and reuse in mass quantities (Custers & Ursic, 2016). Therefore, in our view, *responsible* rather than *limited* data access can offer a better alternative to personal data protection in terms of governing mass data sharing in a smart city context. Therefore, a governance strategy for mass data sharing should facilitate data access, provided this does not undermine the data integrity of data subjects (citizens) and their communities. Equally, the personal data protection approaches typically rely on the rarely sustainable *limiting of data sharing to a single purpose*, beyond which the data is usually required to be deleted like in sec. 10(2) of the Malaysian Personal Data Protection Act 2010 (PDPA). Additionally, the basic principle of informed consent is something of a mystification in an era of pervasive and 'cascade-like' sharing; its limits are perhaps best exemplified in the split-second, multiparty sharing of personal data with advertisers involved in the practice of Real-Time Bidding (RTB) which underlies much of advertising on the internet (Veale & Zuiderveen Borgesius, 2022). But perhaps most importantly, the focus on 'personal data' rather than monitoring user behaviour for the creation and use of algorithmic identities by powerful players may be detracting from the real issues at stake (Lindh & Nolin, 2016). This caution underpins the

[9] The Personal Data Protection Act 2010 (Act 709) in Malaysia, and the Personal Data Protection Act 2012 (No. 26 of 2012) in Singapore.

Report's preference not to engage in distinguishing between personal and business data but instead to focus on the origination of what might be seen as personal data (from citizen/residents) and concentrate on ensuring data integrity in the face of mass data sharing.

Separately, PDP laws in the region often have limited scope and application through the existence of exceptions such as the limitation of scope to 'commercial transactions' (Malaysia), and exemptions from the reach of such legislation for the public sector (in Singapore and Malaysia). This raises doubts about the extent of their applicability to partnerships involving data sharing such as are of interest to the Report, especially if, for example, it is argued that the private partners act on with delegated public authority either in operating service delivery applications or providing the technology and know-how for processing shared data. Additionally, the PDP authorities in all three focus states (Malaysia, Indonesia and Singapore) are government agencies no doubt mindful of public/private data sharing interests.

The second consideration is that of corporate compliance regimes with a top-down data management preference, and a heavy reliance on consumer trust. These regimes rarely encourage data subject participation, and discourage requirements for the notification of data use and reuse. Indeed, most privacy policies analysed in Part B do not discuss exhaustively and clearly the matter of how citizen data might be reused beyond the primary purpose of enabling access to service. Thus, the data subject is caught up in disabling information deficits and this can have a fundamentally negative impact on data integrity, considering that the data subject has no role in verifying the accuracy and quality of data about them in a smart city context.

The appeal of PDP and privacy approaches may be additionally circumscribed in some Asian contexts where privacy, a key notion in personal and computer information ethics, is observed as a largely introduced concept, with the word itself having been loaned from English, as is true for *privade* in Thai (Kitiyadisai, 2005), *puraibashii* in Japanese (Horibe, 2017) and it equally rings true also for the word *privasi* in Malay and Bahasa Indonesia. That said, the importance of the concept of privacy for Islamic culture and Sharia law, in particular the prohibition of 'spying on another' in the Quran (Hayat, 2007) may be a factor appreciating relationships to privacy in Malaysia and Indonesia (Walters, Trakman & Zeller, 2019, p. 7). As for understanding private domain and private personhood, in many countries where living is more communal, the distinction between private and public space might be more permeable than in Western jurisdictions, as exemplified by the uses of five-foot walkways underneath historic shophouses in SE Asia, including Singapore (Harding, 2018).[10] Additionally, like in Singapore, Malaysia[11] and Indonesia,[12] privacy may not be a constitutional right, and personal data may not be protected in the same actionable way that it is with, say, the GDPR. Therefore, governance challenges will be viewed against more organic notions of community/neighbourhood identity, and citizen/resident life-space interests.[13]

Beyond the realm of individual privacy claims and personal data, community identities may come under particular risk from the platformisation of areas of public service delivery. Dan Cohen (2022) describes how edu tech products and services rely on severing the links between students, communities, teachers and schools to serve 'bare' education decoupled from a specific location which would have been a factor in forging those social connections. Similarly, self-tracking usually implicates other people and can affect relationships with them (Neff & Nafus, 2016, pp. 3, 9), yet these social dimensions seemingly disappear as tracking services and devices are usually marketed as tools of self-knowledge and self-care. If it is through commodifying intracommunity bonds that new datafied social orders are pursued as Couldry and Mejias (2019) suggest, alternative governance strategies must emphasize and protect these bonds – perhaps even more so in communitarian cultures. Therefore, the Report and the analysis on which it draws are not primarily geared to consider data sharing from the point of view of the challenges to privacy and personal data protection it poses. Instead, the Report addresses the good governance of public/private data sharing as requiring a repositioning of the data interests of citizen/residents and their communities, working to ensure they have first say in the control of accessing and using such data. The priority is proactive data subject engagement rather than increased protections from an external agency.

[10] In making this observation caution should be raised at careless generalisations about 'Asian cultures' or Asian practices. While attitudes may be culturally influenced when it comes to individual and communal liberties, there is not one prevailing position on such across the vastly diverse region.
[11] Although the Constitution of Malaysia (first enacted in 1957 as the Constitution of the Federation of Malaya; amended in 1963 to become the Constitution of Malaysia) does not recognize a right to privacy per se, Malaysian courts have sometimes referred to it as a fundamental human right, or even a constitutional right. See Chin Jhin Thien & Anor v Chin Huat Yean @ Chin Chun Yean & Anor [2020] 4 MLJ 581 (FC), para 90, and Genting Malaysia Bhd v Pesuruhjaya Perlindungan Data Peribadi & Ors [2022] 11 MLJ 898 (HC), para 154, respectively.
[12] The Constitution of Indonesia (1945) does not explicitly envision a right to privacy; at the same time, Article 28G para 1 refers to concepts which are often linked to privacy. See fn 74.
[13] Life-space interests is a generalizing of those interests that would be considered the responsibility of urban administrations to offer and ensure in a well-functioning urban space.

Further, the research works from the acceptance that some forms of surveillance,[14] certain situations of data sharing and responsible data access can enhance the resident life experience in smart cities. The core interest of the research was, firstly, to learn more about the role and function of data sharing in certain Asian smart cities and to look at the challenges to data integrity posed to data subjects and their communities in such sharing. Secondly, the research has endeavoured to model an appropriate governance frame that recognises the merging of public and private interests while holding true to the notion of 'citizen-centricity'. This notion reflects the belief that the participation of citizen/residents in the sharing of their data as smart cities evolve is central to any good governance model for responsible data management (Drazewska 2023a). Indeed, citizen-centricity has been identified in policy planning for certain cities as a core aim, as discussed further below.[15]

But pursuing good governance of data sharing in a citizen-centric commitment is no easy task. Observations from our roundtables with regional experts confirm that, firstly, citizen/residents are largely ignorant of how their data is used and reused as a consequence of living in smart cities. Secondly, some countries are so internally diversified when it comes to digital transformation, digital readiness and even common languages of communication, that massive information deficits and lack of citizen awareness would be fundamental challenges to a governance approach resting on transparency and accountability.[16] Adding to this, the private sector stakeholders may see themselves as bound by contractual limitations in discussing their data practices, or view excessive openness as possibly harmful to their position in the market. Public sector agencies, on the other hand, sometimes are confused about their responsibilities to citizen/residents in disclosing data use while respecting state confidentiality and legal constraints. Due to the black box, or opacity problem associated with how data is processed by info tech algorithms, they also can have limited understanding of the technology they license and its data processing potential.[17]

Further, it is not common, beyond very broad statements of intent, to be presented with detailed urban data use policy from smart city planners when investigating associated data use and dependencies. There will be instances in the analysis to follow where the barriers to 'knowing' the nature of presently-exercised urban data governance are replete and confounding. Even so, it is possible to identify governance 'flashpoints' through examining specific orbits within which data is shared (i.e., large areas of public service delivery)[18] and their particular lived applications in the experience of citizen/residents, and from there to craft governance best practice in this age of surveillance and data confluence through mass sharing.

Significant assumptions underlying the research questions for this project are that:

- Mass surveillance and mass public/private data sharing in smart cities are built around intentions for social ordering through data management and technological intrusion, and thus can challenge traditional governance and the Rule of Law;
- These intentions for 'ordering' the lives of citizen/residents need to be inclusive of their participation and interests, and as such should be 'governed' by principles and practices that are fit for this new urban setting, where public and private governance modes are no longer independent or even clear; and
- While surveillance and mass data sharing in smart city administration may have legitimate social ordering purposes, they too should not be absolved from their own social ordering (governance for social ordering).

In this vein, ethical principles like transparency, accountability, fairness and non-feasance need to be given a renewed governance vitality, reflecting the intentions of Rule of Law in environments where data transactions are neither simple nor easily identifiable.[19] The research will offer the frame around which important urban data governance policy can be realised in atmospheres of real-life benefit for citizen/residents, in line with smart urbanism's primary promise of improving their current and future existence. This frame will be indicative, awaiting further empirical application and testing. In addition, it appears from our conversations with the project's research network that a universal governance model in any detail, across diverse ranges of digital transformation and data sharing would at this stage be unrealistic and indeed against the spirit of citizen-centricity and contextual relevance. That said, the model in Part C does have a generalisable expression, even as it draws from the different contextually-located challenges in Part B. It is a difficult balance to strike, as there is danger in a universal approach that does not recognise the importance of context, just as there is in overemphasising subjectivity. The model is expressed in a way that envisages that it will be applied and adapted by stakeholders in line with their negotiated relationships and mutualised interests.

[14] By surveillance we mean the ongoing and often pervasive accumulation of personal data by technology directed to observing human behaviour, or the accumulation of personal data from input sources that is used to chart human subjects and their behaviour.
[15] For example, Digital Society – one of the three pillars of the Smart Nation policy in Singapore – is premised on the ideas of citizen empowerment and inclusivity, and, among other things, involves the co-creation of solutions with citizens (MCI, 2024). Also note the statement by the Indonesian Minister for Education on user-centric design in platform education (GEM, 2023), which resonates with the co-creation model in Part C.
[16] These observations have influenced the research model adopted in Part B, which endeavours to recognize contextual particularity and balance universal governance principles.
[17] It is beyond the scope of the Report to adequately deal with data dependencies between the public and the private sector, and resident/citizens to both. The Report will, however, comment on how unrecognized data dependencies can present governance challenges.
[18] Part B will explain the method applications of 'orbits' and 'snapshots' in detail.
[19] These concepts appear prominently in ethics governance modes for technology development and application. They also feature within the core values of Rule of Law governance.

True to its ambitions of exploring co-creation and co-production,[20] the Report does not conclude with recommendations which may appear to be too declaratory, but offers observations on which policy reflection can build, with genuine stakeholder ownership underpinning the governance enterprise.

Locating the research interest

The project's substantive area of interest investigates the public/private governance challenges posed by mass data sharing in selected Asian smart cities, with a particular focus on comparing examples from South-East Asia.[21] South-East Asian cities have evidenced in recent decades a considerable commitment to pursuing digital transformation as solutions to endemic urban problems, in addition to pressures from rapid urbanization, pollution and traffic congestion,[22] based on AI-assisted surveillance technologies. This commitment has emerged through a confluence of public and private partnerships in the rollout of smart city programmes, a consequent dissolving of more conventional divides between governance dimensions, and a consolidation of potentially unchecked power of the private tech sector in the public sphere.[23]

As mentioned earlier, Asian jurisdictions do not always have a history of detailed and comprehensive personal data protection regulation, or privacy rights embedded in constitutions. In fact, justifications of data protection in parts of the region have commonly referred to the necessity to ensure the development of e-commerce rather than upholding privacy as an intrinsic value (Ess, 2005, p. 2). Additionally, with free-market, neoliberal commercial leanings, the business interests of profitability and cost efficiency in data sharing often drive much of current urban policy in Asian city planning (Chen & Shin, 2019; Park, Child Hill & Saito, 2012). Existing regimes of personal data protection also tend to reflect the states' expressed concerns that regulation should not slow urban economic growth – for example, by including business interest exemptions, such as found in India's recent Digital Personal Data Protection Act, 2023. That said, most if not all the jurisdictions with which the project is concerned have long-standing ascriptions to Rule of Law governance, be it in an Asian interpretation of state responsibility, citizen obligation, and business compliance such as the Singapore 'Model Framework', first released in 2019, as updated in 2020 (IMDA & PDPC 2020). Compliance is usually directed here to private sector best practice informed by principles set by the public sector. In this sense, the inclusion of citizen/resident data subjects in compliance requirements or monitoring is not typically ensured through trusted third parties such as data stewards.

The dominant position of private technology providers in such configurations is thought-provoking when reflecting on citizen autonomy and opportunities for their involvement in, and contestation over data use (e.g., when their data becomes repurposed for commercial ends without their knowledge and authorization), the impact on public values such as transparency and accountability, and overall legitimacy of private tech companies to act in the public sphere outside of contractual confines. The extent to which the Rule of Law has purchase on conventional private law arrangements and good urban governance is a relevant theoretical and practical consideration when examining public/private data confluence in smart city administration, because sharing takes private governance into the realm of the public. An example of this transition is where claims of public interest can infiltrate the exclusion of contractual standing, reflected in the French idea of public contracting.

The application of the Rule of Law has perhaps particular potential as a regulatory mechanism, considering that even self-regulatory modes require a clear normative constitution against which governance can be measured. As the targeted regulation and governance of data use and sharing across smart cities in the region tends to lag behind the development of data-based technology and new modalities of voracious data access and consumption, we posit that questions about the protection of the data interests of smart city inhabitants require answering through a recognition of the interoperability of Rule of Law standards. If private law arrangements accept the relevance of Rule of Law ideals, then the next issue is whether such ideals are translated into private law regulation (and private law as regulation). By this we mean to pose a (rhetorical) question: as the public and private governance modalities draw together and public governance claims a Rule of Law reflection, is it not inconsistent to see private law governance as outside the Rule of Law? Before data was shared on this scale, the private actor might have resorted to the market as the sole regulatory frame, but now the public/private governance challenges and domains are common, as Part B will reveal.

[20] See fn 135 and corresponding text.
[21] The project engages with snapshots as examples of mass data sharing contexts from Singapore, Indonesia and Malaysia. However, we occasionally refer to examples from Thailand, India and Korea. In this sense we are covering Southern and Eastern Asia beyond ASEAN states. These snapshots could equally be seen as case-studies, but because none of them can be a complete representation of the issues in question we have avoided using the term.
[22] While these are important purposes and locations of mass data sharing, the snapshots in Part B do not cover these areas of administration.
[23] Governance approaches that rely on data segregation, distinction and transgression (understood as breach of data segregation) will be poorly suited to mass data sharing phenomena, considering their sheer scale and ongoing character.

In light of these reflections, the project has encompassed 3 pillars (three overarching themes/points of interest for all three parts of the Report):

- The governance of data sharing in health, education and government tech, with a particular focus on public/private sharing, and the requirement that public/private confluences in surveillance and data access/use be better understood. Are these data practices operating accountably to ethical/ Rule of Law principles so as to enhance citizen trust and state legitimacy? In terms of legitimacy in the eyes of the citizen/resident (directed at public or private administrative agencies), ethical compliance will require a reflection on citizen/resident participation, in keeping with citizen-centricity.
- The importance of including data-subjects (residents and their communities) in the sustainable regulation and governance of data sharing in smart cities. The purpose of such inclusion in the form of co-creation/co-production (see Part C) is to address broad challenges to data sustainability and its social good in creating and maintaining humane 'smart' urban environments. Any such inclusion/participation here needs to be more than nominal, and requires motivations to be provided by powerful stakeholders in the data ecosystem for less powerful players to become engaged; and
- The measure against which powers of surveillance and data sharing in smart cities are recognised as compliant with Rule of Law principles such as transparency, accountability and equality. Again, these measures need to be actual, real and genuinely participatory and inclusive.

Overview of the Project

Mass data sharing in smart cities: Regulation and governance of the public-private interface (the Project) is interested in mapping the ways through which the Rule of Law (ideals and practice) and good governance based on citizen/resident inclusion can be understood and consolidated in the smart city data governance landscape, recognising the challenges posed by pervasive data production and sharing between public and private actors and agencies in surveillance and the delivery of administrative services. The Covid-19 pandemic control has normalized an extensive dependence on AI-assisted community surveillance, and a reliance on private tech companies in the delivery of public services. The legitimacy of such companies to act in the public sphere, assist in surveillance and assume roles in coordinating mass data through service delivery will depend on their position vis-à-vis considerations of good governance in smart cities. Public/private arrangements in these activities involve significant crossovers between the public and private governance spheres, accompanied by pervasive data-sharing. One example is where the state contracts with private info tech providers and platform operators to develop and operate public service access and delivery portals. Such public/private partnerships necessitate conveying public data through private processing channels to business data dimensions.[24]

Relevant examples from SE Asia include the public sector's reliance on:

- creating and developing nation-wide, multi-purpose health data integration platforms as a more advanced version of Covid-tracing apps (employing surveillance and tracking), indicating the extent of mass public/private data sharing, as shown in snapshots A (SatuSehat, Indonesia) and B (MySejahtera, Malaysia);
- private tracking technology in the development and rollout of nation-wide health programmes as shown in Snapshot C (LumiHealth, Singapore);
- private educational tech arrangements for new teaching and administrative platforms in public schooling, as shown in Snapshot E (Ruangguru, Indonesia), and Snapshot F (DELIMa, Malaysia);
- private technology providers in public service delivery functions and in the performance of key government functions using generative AI;
- chatbots and derivative technologies both in the front- and back-office of government agencies, as described in Snapshot D (Pair, Singapore);
- private communication platforms such as Whatsapp or Telegram for communicating with the public;[25]
- transfer of the private tech sector's talent and know-how to the public sector, as evidenced in the practice of 'revolving doors', involving Indonesian tech start-up founders and executives assuming positions linked with the exercise of public power (Jibiki, 2022);[26]
- super apps such as Grab or Gojek to distribute welfare benefits, during the pandemic and beyond.

The examples discussed in the Report were selected not as comprehensive, but rather as representative of essential

[24] As observed earlier regarding the position of the Report on the public/private distinction, it is currently equally difficult demarking between public and private data. In the context of mass data sharing, the data that is shared can be distinguished based on whether it has a public or a private source, and an original, public or private purpose, private here meaning 'private sector' or 'market'.
[25] The most recent WhatsApp communications from the Singapore government in the second half of 2023 addressed online scams prevention advisory and rollout of personalized health plans for all Singapore residents. It may be worth noting that France has recently announced plans to replace the use of WhatsApp, Signal and Telegram in the public sector with a French instant messaging platform due to fears for security of confidential information (Pineau & Hummel, 2023).
[26] Such public-private transfers can also follow opposite trajectories, as recently demonstrated by the hiring of an MP by Grab in Singapore (Elangovan, 2023).

urban lifestyle experiences of service delivery through public/private convergence, and subsequent surveillance. Further examples include new super apps for service delivery as well as private sector information tech providers making centralised applications (gateways through which citizen/residents can communicate their data to a variety of urban administrative services) available to the state.

This confluence of personal data use gives rise to multiple Rule of Law and governance questions not least because of pervasive privately-operated surveillance systems on which some of such public and private collaborations rely for collecting citizen data. For instance, the Covid-tracing apps were justified as an essential application of technology to ensure health and safety during a public health emergency, which cannot easily be supplanted by ongoing considerations of efficiency in the post-Covid era, impacting the legitimacy of private actors in managing public sector-generated data.

Analytical concerns

The analytical and policy concerns underpinning this research are concentrated around the governance impacts of such new data-driven collaborations. This includes the issue of how the involvement of private partners, with their opaque, IP-protected technologies contextualised within profit-oriented business models are amenable to good public governance, accountability and transparency requirements. Within market-centred private governance regimes, the traditionally transactional approaches to customer interactions, adopted through private sector service delivery and shareholder-oriented accountability mechanisms, as well as commercial/contractual relations with their contract-party parameters, will impact public values and governance principles such as accountability, inclusion, participation (including participatory data governance), transparency and fairness. These principles are also strong compliance expectations for Rule of Law governance considerations in public administration. Another related research interest is in how a concentration of power in the operational hands of tech companies, with little to no legal/regulatory constraints when it comes to data use and management, may impact smart city governance, especially in Southeast Asia where the citizen participation sphere is often qualified by paternalist or authoritarian administrative policies as a vestige of precolonial traditions (Haque, 2007).[27]

The research anticipates as its central analytical and comparative concern, therefore, that the confluence of data originating from the public and private partnerships in smart cities challenges traditional modes of urban governance, their principles, practices and modelling. As these challenges are novel in the smart city context, governance responses working from conventional public/private dualities may be inadequate to address citizen/resident interests, and current data protection and policy research is not adequately interrogating these issues.

Structure of the Report

The Report adopts the following structure:
Part A: Research Tools. This part offers a review of the research underpinnings, i.e. the theory and method considerations on which the analysis is based. Because of the multi-disciplinary nature of the central research questions, it was essential to have clear theory/method foundations which substantiate the research exercise. In particular, Part A identifies the preferred approach to governance (focusing on co-creation and co-production to produce respectful, responsible and trusted data relationships) adopted throughout the research. It also addresses the comparative approach, elaborates on the significance of the regional networking exercises, and concludes with an overview of the structure, intentions and outputs from the networking roundtables.

Part B: Orbits and Case-study Snapshots. Inspired by the preparatory work done before and shortly after the commencement of the project (Drazewska, 2023b; Drazewska, 2023a), the empirical part of the Report undertakes an analysis of examples from sharing of citizen data in the context of public services across three orbits: educational technology, government technology and health technology in Singapore, Malaysia and Indonesia. The snapshots present a variety of recent and ongoing public and private partnerships (involving single or multiple private partners from foreign as well as regional or international Big Tech companies) to roll out nation-wide initiatives as well as specialized programmes and technologies addressed to a sector of the society. Part B explains the three orbits, and provides elaborations of work done through case-study snapshots and the universal and specific concerns they raise to draw out governance and Rule of Law challenges subsequently addressed in the model presented in Part C.

Part C: The Governance Model. The final part offers a model for governance of the challenges posed in Part B and summarises the main deliverables of the research, in terms of reflections and advices. Part C should not be read as recommending definitive policy directions but rather as developing a model for governing the principal challenges presented by mass data sharing which can be translated into an agenda for policy action in various locations. A governance model is sufficiently general to offer applications across challenges that are common in various smart

[27] The Report makes no comment on the nature of different state or private sector administrative styles beyond how these impact on citizen/resident data integrity through data sharing.

city settings, while respecting the different starting points and stages of digital transformation raised in the research network conversations.

The SE Asian regional focus is recognised throughout, however, it is anticipated that the agenda will be of a broader interest to stakeholders responsible for the governance of urban planning policy wherever smart cities are advocated and already operationalized.

Part A: Research Tools

Overview

This Part will set out in brief what theoretical directions have been employed and how they have informed the development of inventive research methods, as well as explain the key terms and research assumptions employed throughout the Report.[28] The Report is based on original research carried out in pursuit of what we genuinely believe to be a ground-breaking research enterprise. The early and ongoing review of policy research literature fields that might be expected to reveal similar research has spotted only patchy connections with the central research interests of the Report, within and beyond the SE Asian region. The available literature varies in breadth and focus. The paucity of complementary research materials (and by extension, publicly available data) should not be seen as diminishing the importance of this work, however.[29] Rather, it can be taken as revealing the challenges associated with a research trajectory that needs access to information which in many circumstances is largely guarded behind the barriers of confidentiality. Examples include commercial arrangements with information platforms and tech companies in the provision of public services in smart cities, and their regulatory frameworks to ensure data integrity and preventing vulnerability/discrimination in data ecosystems.

In the process of identifying applied social theory to assist in explaining urban governance arrangements, the research team worked back from two core analytical concerns:

- What is meant by governance and, especially, 'good governance' when it comes to the contexts that are central for the analysis?
- How does 'social ordering' play out in the way data use can order urban communities, and how does this data use produce consequences that in turn require ordering?

It is now clear that these concerns overlap to an extent, and grow out of prevailing power asymmetries between stakeholders in various data ecosystems. To understand the dynamics of these differentials some broad theories of power are required. *In its most general consideration governance is a process of social ordering based on structures of social consensus, or hierarchies of power.* A simple example is governing a school classroom dynamic. The teacher requires some accepted status of authority that gives her power to set rules concerning classroom behaviour. She needs some physical positioning to monitor her students, some communication capacity to convey these rules, some inducements for them to be followed and some sanction power in case of a transgression. However, the initial governance authority is unlikely to be maintained unless relationships of trust are negotiated between teacher and student, and this process will require some *dispersal of power*. Power analysis is convincing and revealing for explaining the power asymmetries in data ecosystems, but it is not on its own sufficient for understanding all the reasons for, and the process of communication between those with interests in data. Why would the data-powerful be willing to relinquish control over data to the data subjects? This may not be about power ongoing, but a negotiation based on mutualized interest and as such, theories of consensus could be relevant. In the smart city urban space and with data as both the object for governance and the tool for social ordering, it is particularly important to see governance not so much as a political exercise, but a communitarian endeavour – making life-spaces best for the citizen/resident in terms of data use and technological intervention. In the present project, *governance* is therefore understood as 'governability' – how people are socially ordered to their benefit.

Prior to explaining such theoretical perspectives in detail and honing these to applied purposes, the research needs to declare its vision of the communities that comprise city living, smart or otherwise. In doing this the Project draws on the extensive work already established by the Centre for AI and Data Governance at the SMU on the importance of trust in machine/human relations, and the concept of *AI in community* (CAIDG, 2023).[30] CAIDG's work in understanding where AI and data use should fit in healthy and sustainable social spaces has drawn on Roger Cotterrell's socio-legal approach to communities, which suggests that communities are not so much things or places, but rather they exhibit and rely on networks of trust (Cotterrell, 2018, p. 2). Mark Granovetter's work on how social bonding functions also offers helpful understandings about the interplay between trust and power (Granovetter, 1973). Mark Findlay's paper on AI dependencies (Findlay, 2023) further assisted in narrowing down the area of power theorizing.

[28] In very brief terms, theory is the structure chosen for explaining a problem, while method specifies how that explanation is applied in analytical work.

[29] For the future of ensuring good data governance in smart cities, it is important that private and public data holders see and understand the value of research into data relations and good governance. Research and modelling can have significant practical value considering that in a world of Big Data sharing, it is becoming increasingly more difficult to ensure data accuracy and integrity. One simple, but profound potential in the governance model offered in Part C is that with its data-subject focus there are significant individual and communal possibilities to test and confirm (or not) data accuracy against original data sources.

[30] 'AI in Community' is a concept developed by the interdisciplinary research team at CAIDG SMU, which explores the roles of AI technologies embedded into distinct environments in community spaces. The main assumption of AIC as a critical framework is that AI technologies and the data on which they rely should be located within specific communities, and reflect and correspond with community preferences and priorities for digital life-experience. AI systems ought equally to be steered towards the public benefit of the people and communities interacting with such technologies, because they provide the data used to develop and build AI systems.

Insofar as the location of this research in smart cities in concerned, the Report views smart city as a general context against which the key transformations of in the area of digitization and platformization of public services, and of societies, play out. In that sense, the Report's understanding of the smart city context emphasizes access to digital technologies, specifically data infrastructures for accessing such services by the general public, which have made possible 'lower cost, more dynamic, and more competitive alternatives to governmental or quasi-governmental monopoly infrastructures, in exchange for a transfer of wealth and responsibility to private enterprises' (Plantin, Lagoze, Edwards & Sandvig, 2018, p. 306). As these services are becoming increasingly remote, an 'urban' setting will not always be strictly essential for accessing them.

Normative core – citizen-centricity

Prior to identifying a theory of governance, this research sought to reflect on what ought to constitute its normative core, or the *leading principle* that governance will recognise and try to achieve. ***Citizen-centricity*** emerged early from policy pronouncements advancing smart city urban development. Aside from the compatibility of a citizen-centric frame with policy aspirations such as inclusion, representation and accountability, this analysis relies on citizen-centricity in data governance because Singapore, Malaysia and Indonesia have directly or indirectly expressed a commitment to it through law and policy.[31] Singapore's Smart Nation policy has a clear citizen-centric core, whereas in the case of Indonesia a similar intention can arguably be gleaned from its adoption of data localization policies (even if they have been circumscribed over time).[32] Such policies are usually introduced citing reasons which refer to various aspects of national interest (especially economic opportunities) and citizen benefit, such as better privacy.[33] Article 17.2 of the Indonesian GR no. 82/2012 declares that purpose to be 'law enforcement, protection, and enforcement of national sovereignty to the data of its citizens'.[34] While data localization laws may not necessarily be capable of successfully fulfilling some of these aims (Burman and Sharma, 2021), and the jury is still out on whether their benefits outweigh their drawbacks (Paska Darmawan, 2016), we take the introduction of data localization measures in Indonesia to be a sign of an intention to pursue a citizen-centric approach in data governance. The importance of a user-centric philosophy in platform-based public (educational) services was recently confirmed by the Minister of Education and Culture in the context of educational technologies (GEM, 2023). Further, according to the Ministry of Communications & Informatics (Kominfo), the benefits of smart cities include improved transparency and citizen participation (Devita, 2017). Somewhat similarly, the Malaysia Smart City Framework (MSCF) uses the language of citizen empowerment and participation, referring to them as some of the key urban challenges smart cities are meant to address, and treating them as criteria that can influence the success of smart city implementation (MHLG, 2018, pp 6. and 11). The MSCF also references using data/information collected in accordance with its owners' wishes (MHLG, 2018, p. 6). This legitimizes an expectation of some degree of respectful appreciation of the importance of the data subject's inclusion and participation in the decision-making.

Indeed, while being contestable and somewhat open-ended as a notion,[35] citizen-centricity offers support for a grassroots, inclusive and empathetic governance project. As with any governance terminology, a 'citizen-centric' character of a policy may mean different things to different stakeholders in a data sharing ecosystem such as a smart city. What remains common in any interpretation of the concept is the prioritizing of legitimate citizen interest, and the pivotal positioning of the citizen in governance processes and outcomes. Indeed, if citizen/resident is central for the data,[36] then they should also be central in its management. Citizen-centricity equally implies a degree of autonomy, although the manner in which autonomy is understood can be contextually specific and somewhat culturally relative. In certain situations, it can veer towards individualist egoism, and in others it will acquire meaning from the web of communal obligations and dependencies in which the individual is located.

---

[31] The analysis also adopts Rule of Law normative indicators of good governance, several of which emphasise features of governance that are compatible with or require citizen-centricity.

[32] Data localization was first introduced in 2016 with Government Regulation (GR) No. 82/2012 (PP No. 82/2012) on the Implementation of Electronic Systems and Transactions, whose Article 17.2 requires ESOs (Electronic Service Operators) for the public service to build a data center in Indonesia ('Electronic System Operator for the public service is obligated to put the data center and disaster recovery center in Indonesian territory for the purpose of law enforcement, protection, and enforcement of national sovereignty to the data of its citizens.') Source of the translation: http://www.flevin.com/id/lgso/translations/JICA%20Mirror/english/4902_PP_82_2012_e.html.
Subsequently, the localization requirements were limited in GR 71/2019 (unless the technology is unavailable in Indonesia). Both private and public-scope ESOs however remain subject to registration in Indonesia, and to Indonesian laws with respect to their activity on Indonesian territory.

[33] For example, in the case of India, the reasons to include data localization provisions in the in the discussion around the Personal Data Protection bill included:
- securing faster and better access to personal data for law enforcement, helping to avoid delays in solving crimes;
- spurring economic growth and employment;
- preventing foreign surveillance of citizen data, and
- better enforcement of data protection laws (Burman and Sharma, 2021, p. 11). Other commentators have pointed to the correlation establishing new data centers and infrastructural improvements which also lead to a better life quality of the community and an increase in the national economy (Paska Darmawan, 2016).

[34] For more detail on that regulation, see fn 32.

[35] To say the citizen must be central implies several essential questions. Central in what? Central for what? How much is the citizen actively involved in their central positioning?

[36] 'Resident' is attached to citizen throughout in recognition of that significant part of the population that makes many smart cities function, but do not enjoy formal citizenship.

In the data sharing ecosystem that is the smart city, citizen-centricity can be reasonably expected to be central within networks and relationships of data access and use. However, it should be observed that data creation is rarely a singular or unidirectional process. Data is generated through communication between data communities, and the pathways across which it travels are social. Thus, there is no singular individual that can claim to be 'central' beyond a broad acceptance that the original data subject has an important place in data hierarchies. What this also means is that *as data is a social fact* in a Durkheimian sense,[37] obligations for influencing and respecting the 'centricity' of other citizens should rest on those who gain advantage through data access and use. On the one hand, the 'central for what?' question cannot, and indeed should not be disengaged from the creation and maintenance of respectful data relationships between citizens and stakeholders in the ecosystem. But on the other hand, in delimiting the extent of responsibilities citizen-centricity involves, it is important to realize what is the power arrangement across and throughout data ecosystems. There are many realities in any such ecosystem, such as information deficit, that will result in data asymmetries that disempower individual data subjects and their communities, discourage them from participating, and even capturing any efforts by the citizen to claim meaningful participation. Therefore, consistent with equality as the cornerstone of Rule of Law, the burden rests on powerful stakeholders in a data ecosystem to actively ensure the possibility of meaningful citizen participation.

The focus of our analysis is how mass data sharing between public and private data holders (and the subsequent surveillance possibilities) create challenges requiring a different approach to data governance than that which has been traditionally offered in public and private paradigms. The merging of public and private administrative activities in urban development has been an evolving feature of service delivery for decades now. What sharpens the argument for a new governance model based on the recognition of citizen interests are the apparent challenges to citizen-centricity arising out of mass data sharing. As such, a core direction for identifying and exploring these challenges is to evaluate how individual data 'citizens' and their communities are becoming marginalized from the management of their data without sufficient recourse to governance frames consistent with Rule of Law. Citizen-centricity is not just an outcome of good governance, but also a trigger for reviewing governance adequacy.

The public/private confluence that is mass data sharing will have different dimensions in different contexts. In the EU, for example, where the GDPR has demanded a strong individual rights discourse, the influence of public regulators over the private data market and over governance considerations is dominant. In the Southeast Asian region which is the focus of this study, public governance frames and languages are somewhat likely to be influenced by the thinking and priorities of the private sector. Private sector influence may be exacerbated by dependencies on imported private sector know-how, and compounded by close relationships between Big Tech and governance elites. This was recently illustrated by Vietnam's reliance on Facebook's help for initiatives to accelerate the nation's transition to a digital economy (MIC, 2020) and to censor anti-government online content (Nguyen An Luong, 2020). Further evidence of this can be found in how personal data protection instruments and entities are qualified in favour of business interests, as already mentioned above. Equally, it remains valuable to reflect on how the influence of either side in the public/private confluence will impact on understandings and operationalizing of citizen-centricity.[38]

Conceding that governance languages and their nuances will impact on the understanding and activation of citizen-centricity does not diminish its importance as an indicium of good governance. As with Rule of Law and the sometime-apparent slippage between its expressive function and governance practice, citizen-centricity is not denied its place in good governance because some stakeholders say it can only be achieved by taking the current power asymmetries in the ecosystem as a given. And assuming that power dispersal is possible within data power relations, citizen apathy or reluctance to claim centrality also does not undermine the importance of governance policy providing support for active citizen engagement.

[37] According to Durkheim, a social fact is 'any way of acting, whether fixed or not, capable of exerting over the individual an external constraint', or 'which is general over the whole of a given society whilst having an existence of its own, independent of its individual manifestations' (Durkheim, 1982 [1895], p. 59). On qualifying data as social facts see also Risse (2023).

[38] For a view of 'citizen-centricity' in service delivery that is aligned with the current trend within the private tech sector to 'personalize' digital products and experiences see Teng, R. (2022).

Explaining governance, ordering and data

Governance

In the Report, we transcend theories of governance which primarily focus on the role of political authority, and view the act of governing as a primarily political endeavour. As an example of such an approach, according to Y. Keping, 'in terms of political science thinking, governance refers to the process of political administration, including the normative foundation of political authority, approaches to dealing with political affairs and the management of public resources. It particularly focuses on the role of political authority in maintaining social order and the exercise of administrative power in a defined sphere' (Keping, 2018, p. 3). In this interpretation of governance the importance of a strong normative foundation as a basis for legitimate authority can be observed. Rather than on political authority, the governance model discussed in Part C relies on 'buy-in' from citizens and public and private data users to ground the authority of governance. Equally, the Report advances the utility of Rule of Law normative frames that constitute good governance in practice, and therefore a general ascription to Rule of Law forms the normative backbone of the model. This can be connected to another view of governance as offered by L. Lynn, C. Heinrich and C. Hill (2000, p. 235), whereby it can also be understood as 'regimes of laws, rules, judicial rulings, and practices that constrain, prescribe, and enable the production and delivery of publicly supported goods and services.'

The research in the Report works with a participatory style of governance consistent with much smart city policy in the region which speaks of citizen-centricity, the localising of data, and the recognition of community interests as fundamental. In the same vein and in contrast to traditional, top-down theories of government, the idea of governance may be thought to emphasize co-ordination and 'various forms of formal and informal types of public-private interaction' (Pierre, 2000, p. 3) whether by a government, network or market. In this sense, governance may be perceived as an alternative to state control (Hirst, 2000, p. 13). The model presented in Part C is a co-creation/co-production project where governance is prioritised and actioned by the three major stakeholder groups (regulatory recipients of a good governance strategy), i.e. public agencies, tech companies and citizen/residents), and in that sense embodies bottom-up engagement. Our overarching assumption is that in the context of in smart cities, data-subject participation is important in data governance, as individuals ought to have a choice in deciding who can see their data and how that data is used across different contexts as a matter of their digital self-determination (Remolina & Findlay, 2021).[39]

Rather than on the state and its institutions, the emphasis for the authority of governance in the present analysis configures different kinds of activities and processes of governing, often involving more diverse actors and more diverse organizational forms (Bevir, 2012). Hufty (2011, p. 405) similarly defines governance as 'the processes of interaction and decision-making among the actors involved in a collective problem that lead to the creation, reinforcement, or reproduction of social norms and institutions.' Both M. Bevir and M. Hufty see governance as decision-making linked to collective problem-solving. Similarly, D. Kong and K. Yoon define public governance as 'the overall process of decision-making and implementation in solving public problems in a country, where public agencies or institutions initiate the process or are at least partially involved in the process' (stressing the importance of a contextual approach) (Kong & Yoon, 2018). The governance perspective puts a premium on 'getting things done', which implies a sort of 'output legitimacy' (Haus, 2018, p. 53) for actors seen as capable of delivering such outcomes. On the other hand, Haus warns against an overly strong focus on problem solving within governance theories, as it may obscure issues of power, including democratic participation and accountability. In the governance model proposed in Part C, the process of problem solving is not only removed from external state agencies and returned to stakeholders, but power dispersal to correct power imbalances that are at the heart of many data use problems is treated as an essential pre-condition for enjoying mutualised governance benefits. In addition, conflict resolution opportunities proliferate in the model so that problem solving is constant, continual and co-produced as a key to accountable governance. We conjecture that a feature of *good data governance* is the ability of public and private stakeholders who hold significant data-related power to recognise the eventuality of conflict with data users, and to provide processes of accountability that can facilitate conflict resolution. But problem solving cannot be offered as some concession from the powerful to the powerless. Data subjects need opportunities to *own* conflict and participate in its equitable resolution.

Purpose of governance

Most writers define the purpose of (public) governance with reference to realizing public values and keeping the social order. According to Keping (2018), 'The purpose of governance is to guide, steer and regulate citizens' activities through the power of different systems and relations so as to maximize the public interest… As a political administration process like government by the State, governance also requires authority and power and ultimately aims to maintain a normal social order.' One work on global governance defines that purpose as the execution of collective functions in the common interest, adding that trust in institutions (facilitated by the existence of shared values and principles) and confidence that its functions will be performed justly and effectively are its key ingredients (Lopez-Claros, Dahl & Groff, 2020, pp. 391, 433). The proposed governance model in Part C sees the purpose of good governance as satisfying legitimate stakeholder interest, but not in a manner where the data-powerful dominate interest realisation.

[39] On digital self-determination, see fn 149.

Requirements for individual interests to be mutualised in order to achieve governance ordering are a feature of the model.

The concept of public value and order also emerges in smart city literature. In their book on city governance, J. Stenvall, I. Laitinen, R. Yeoman, M. Thompson and M. Mueller Santos state that the purpose of governance is 'to create order' and '[public] value creation' although, as they add, value is not only dependent on outcomes alone (Stenvall et al, 2022, pp. 76, 80). A similar view is expressed by M. Rodriguez Bolivar – public value is a new iteration of the common good discourse (Rodriguez Bolivar, 2019). But that of course could have different interpretations; in addition, there is little research in terms of the definition and assessment of public value in smart cities (Meijer, Gil-Garcia & Rodriguez Bolivar, 2016, p. 650).[40] The proposed model sees data access and use as disorderly if it is progressed in an environment of discrimination and exploitation. Data can be an agent of social order in urban space and its access and use can create challenges to social ordering if it occurs without respecting the interests of all stakeholders in the ecosystem.

By contrast, the purpose of governance in the private sector has been defined as creating 'owner-accountable organizations' (Carver, 2003, p. 4), achieving the aim of increasing their own value (Lubben, 2013, p. 894) or maximizing the shareholders' interest (Alban et al. 2019, pp. 22-23). Especially considering the disparity of motivations and values between the public and private sector, the involvement of the private sector as active agents of power dispersal and important influencers of good data practices is central to the proposed governance model and its co-production.

Data

A theoretical reflection of smart city governance must start from an understanding of data. What is data? Among the many competing theories, it is possible to see it quite simply as information or messages. In the EU, the content of a database can be protected by a *sui generis* right that prevents its extraction or reuse as per Directive 1996/9. This approach suggests that data can in fact be commodified, and is thus cohesive with a view of datafication as a new frontier of value accumulation, which has been criticized by writers such as Viljoen (2021), Nethercote (2023) and Srnicek (2016). The Report is informed by a theory of data transcending its market consideration as an economic resource (as it is capable of shaping people's destinies), on to data as a core element of smart city governance/ordering understood in terms of largely unchecked power.

Keeping in mind the importance of contextualizing data value in terms of stakeholder interest, we reject a formalist data/commodity/property paradigm as it is indeed currently not supported by developments in intellectual property and private law at large. The following analysis prefers to see data in a more direct relationship with data subjects and their communities – as messages across pathways of communication that gain their integrity from the original intent of that communication.[41] Seen this way, data does not need to have a property designation or to be valued in market terms of scarcity and alienation. Instead, data has a richer social value as do the message pathways across which it travels and is communicated. This novel theory on data is crucial for the understanding of data integrity, citizen-centricity, participatory governance and data integrity as cornerstones of this research work. The idea of data revaluing in terms of stakeholder interest is equally a feature of the modelling in Part C; the governance model recognises the value of data, its access and use in more than profit terms. Along with market valuing, and indeed of more significance, is the potential for data management to benefit citizen data subjects and their communities as a priority. This approach is consistent with seeing data as socially embedded – as a social fact that influences individuals and their community in how it is used and reused.

Some voices, however, still cling to quasi-property notions of data. Julie Cohen writes that a revolution in the legal status of data (and algorithms), which become '(de facto if not de jure) proprietary information property' is an important by-product of the access-for-data arrangement which characterizes platforms (Cohen, 2019, p. 44.) This presupposes that data could become something to be owned (e.g. Chik, 2021, p. 77 et seq.), and equally that the owner may be more than the data subject/producer of data (Lessig, 2002), or a wholly different entity – an approach that has been subject to criticism (Findlay and Remolina, 2021, pp. 22-23). We contend that if the notion of data as an object of private property rights and obligations is not settled, then ownership is moot and open to confusion. From a governance perspective it is more helpful to explore possession and use of data and the obligations that ensue beyond ownership paradigms. As demonstrated by Snapshot B, which engages with the Malaysian contact-tracing app MySejahtera, data ownership does not necessarily imply access and full control over it.[42] More pointedly, Fischer and Streinz (2022) suggest that it is not the fact of ownership of data that matters, but rather the question of who may extract value from data, i.e. who has access to the infrastructure that permits such an extraction. This dovetails with those criticisms of the slogan 'data is the new oil' which underscore that the data itself is not as profitable until it has been the subject of value extraction (Van Zeeland, 2019).

[40] Other scholars have defined the purpose of public governance in a broader way, e.g. 'to enable institutions, society and other stakeholders to work together and achieve policy objectives in a dynamic and changing environment' (Wirtz, Weyerer and Sturm, 2020, p. 822). Similar views can be found in the work of Asaduzzaman & Virtanen (2016) and Piso, Goralnik, Libarkin & Lopez (2019).
[41] Integrity here primarily refers to whether the data has retained its original purposes determined by data subjects in making data available for sharing. In this limited sense, this approach can be linked to H. Nissenbaum's theories of contextual integrity (Nissenbaum, 2009). See also fn 3.
[42] See Snapshot B

Regarding data valuing, feminist approaches are instructive. Observing that governments and corporations have long exercised control of people, objects and spaces through data, and indeed that the most powerful companies today are the ones that are able to collect, aggregate, and extract value from data, such approaches address data as a conduit of power (Klein & D'Ignazio, 2020). Since feminism is interested in mapping the way inequality is perpetuated through systems, data feminism inquires how data is used to maintain inequality through sociotechnical assemblages such as platforms. In this sense, feminist approaches argue that for the debate on data to address issues such as human autonomy, dignity and equality, data should be redefined as something akin to human bodies (Kovacs, 2020; Van der Ploeg, 2012). That is in view of the fact that so much of it (e.g. health data) is produced by directly measuring bodies, but also because of the headway that feminism has made through its debates on consent, allowing to view privacy policies and terms of use of digital products as originating in and perpetuating social power differentials and inequality (Kovacs & Jain, 2021).

Ordering

Ordering can be understood as an infinite process of organization of social life (Law 1993).[43] In this usage *organizing* is a social rather than an institutional activity and as such is directed at social bonding, through governance frames like trust. The field of STS (Science, Technology and Society), in which Law is a representative, claims that it is not just humans that may do the organizing – it can be facilitated through rules, architectures, but also through technology (Law, 2003). Different modes of ordering can co-exist, and are not necessarily coherent, in the sense of being within the language of a single discipline. Organizing for ordering can be either mechanical in the scientific sense, or organic as a social enterprise.

An example of ordering is how the modern city orders urban life (and in addition, these forces of ordering need governance as well so that they too are orderly).[44] Smart cities use technology and data to order life and to govern the essential living spaces and relationships. In particular, human lives are ordered through surveillance technology by means of tracking physical bodies in city spaces (e.g. via CCTV cameras), as well as human activity on the internet – for example, the Meta Pixel tracking code embedded in many websites has enabled the collection of information including private data and records of hospital patients accessed without their consent (Feathers, Fondrie-Teitler, Waller & Mattu, 2023; Alder, 2022).

Traditionally, public law and administration has tended to order society through state institutions, while governance has been controlled within frames of principle and process. More recently, technology and data have facilitated public governance, particularly in urban spaces. One such new kind of social ordering (distinct from bureaucratic hierarchies and market-driven ordering) is algorithmic regulation (Eyert, Irgmaier & Ulbricht (2018)), which K. Yeung (2018) defines with relation to 'decision-making systems that regulate a domain of activity in order to manage risk or alter behaviour through continual computational generation of knowledge by systematically collecting data (in real time on a continuous basis) emitted directly from numerous dynamic components pertaining to the regulated environment in order to identify and, if necessary, automatically refine (or prompt refinement of) the system's operations to attain a pre-specified goal.'

In smart cities there are many examples, through the sharing of public and private data and the reliance on information technology, that public and private social ordering are merging. As a result, governance in smart cities is not easily explained in a separation of powers or a market/state paradigm. What is this new 'ordering' (particularly through digitizing and surveillance) and how has the need for and achievement of governance in terms of principles and process changed, can only be answered after a more fundamental consideration of data access and use as languages and processes of ordering, as offered in the empirical part of the Report (Part B). The model in Part C views data as the object of governance, and ordering as its purpose – to produce order through transacting ordering via governing data access and use arrangements.

Ordering is a dynamic decision-making process and as such chimes in well with data management as an ordering frame. Our governance model stresses the creation of safe data spaces[45] which are trustworthy and orderly, and as such have the potential to order data management. In this model, ordering features as a dimension of data use. In addition,

[43] As emphasized by John Law, 'ordering' is necessarily expressed with a verb to reflect the infinite character of the task (Law, 1993).
[44] The importance of ordering needs to be considered as two-pronged. Mass data sharing is revealed in ways that data sharing and surveillance combine to 'order' urban life in smart cities. Additionally, the devices and applications of data sharing and surveillance are pervasive and as such require 'ordering' over and above their information management functions, and towards a deeper embedding in the social utility of the city.
[45] Our concept of data spaces is informed by the sociology of spaces and communities of Emile Durkheim and Roger Cotterrell. According to Durkheim, a social space is a space (whether physical or emotional) where relationships are created based on shared ideas, which connects to Durkheim's *conscience collective*. The nature of those spaces is thus determined by those relationships or, specifically, bonds of trust, according to Cotterrell. Following Cotterrell, by data 'spaces' we mean contexts in which social relations are constructed using data exchange rather than physical or temporal spaces. Following both Durkheim and Cotterrell, in social spaces, individuals become influenced to behave in certain ways through social facts external to those individuals (for a definition of a social fact, see fn 37). Thus, safe and trusted data spaces are those in which relationships between stakeholders will be evident, and in which the behaviour of individuals is determined by data as social facts, and then becomes part of the social fact of a data space (through contributing to creating more data). In this interactive enterprise data helps to determine social spaces in which citizen/residents are impacted in their daily lives. These spaces are 'safe' if the data subject can feel informed and enabled to trust other stakeholders using data and influencing digital personalities. See Durkheim (1982 [1895]) and Cotterrell (1999).

the model explores regulatory strategies for governing data's ordering functions. In particular, the model requires data-powerful stakeholders to divest some of their data dominance to data subjects through information sharing, which will make more apparent the reasons behind data capture and use. In this way the data access and use priorities of data-powerful stakeholders are called into account before a more informed data-subject interest base.

Method for theorizing, interrogating challenges and modelling

The methodology employed to reveal how and in what ways mass data sharing and pervasive surveillance is creating governance challenges (both unique and common) in smart cities in the SE Asian region is essentially comparative. In addition to empirical dimensions, the comparative method is equally a theoretical concern for this research. In coming to our preferred theory of governance and data we have analysed and drawn from other contending theoretical approaches. Further, as previously mentioned, Southeast Asia was chosen as a focus in recognition of:

- The commitment to smart city planning across the region and the development of some premier examples of smart cities worldwide;
- The extensive and deep operation of pervasive surveillance through public/private data sharing in countries of the region as part of Covid-19 control;
- The different levels of development in terms of urban planning and public/private data sharing in different locations in the region, particularly in administrative service delivery;
- The general ascription to Rule of Law governance, while at the same time
- The existence of governance continuums from paternal/authoritarian to neoliberal free-market styles.

These reasons provide a wealth of levels and perspectives for comparative research. So that this research did not get distracted away from its central governance concerns, the Report draws on the following disciplines:

- Law and society and anthropology of law to emphasise the context-specific examination of Rule of Law influences and the manner in which prevailing private and public law traditions operate and are influenced by the data production and management concerns of the Project;
- Regulation and governance[46] to widen the governance net and to include discussions of non-legal regulatory forms of social ordering;
- Social anthropology to ground the case-studies in real-life city experiences;
- Cultural studies in recognition of the traditions and histories that precede smart city evolution and to highlight the importance of social as well as physical infrastructure.
- Because of capacity and data constraints (and assuming that this is the first discussion and issues-formulation stage in a longer research period), the comparative method has featured the following:
- Desk research to uncover and consolidate available information on pervasive surveillance and mass data sharing in the context of smart cities, in the identified jurisdictions as well as more generally;
- Interviews or email exchanges with stakeholders involved in the planning and progressing of urban surveillance and mass data sharing from the public and private sectors;
- Formulating common challenges in terms of power relations, governance challenges and social ordering priorities;
- Organizing and facilitating conversations in the form of virtual roundtables involving interested scholars and policymakers, where hypotheticals and case-study snapshots provided a focus. While network-based conversations across smart city settings involving different stages of digital transformation and starting points for governing mass data sharing prove challenging, it is also a rich framework for testing relevance and policy applicability;
- Curating all this information down from the identified research orbits to a grouping of empirical snapshots interested particularly in the privatizing of administrative service delivery and the digitizing of major areas of public administration in cities (facilitated by the private sector), and finally
- Using these understandings to suggest a model frame for the inclusive and representative regulation of data sharing in smart cities.

To indicate and test the challenges identified in Part B, the analysis has adopted a limited comparative case study approach; a comparative contextual analysis (Findlay & Henham, 2007) growing out of research orbits in collaborations involving health, educational and governmental service platforms across SE Asia, and the sharing of surveillance data they produce. This has involved, firstly, an in-depth analysis of the comparative referents, and, secondly, comparing them against universal themes such as good governance, ethical and Rule of Law principles. The purpose of this exercise has been to develop a better understanding of the relationships between stakeholders in the new data relations across the three orbits, identify the principal governance challenges, and to speculate on the governance mechanisms necessary to ensure the interests of data subjects and their communities are protected. The case studies (snapshots) cannot be comprehensive, representational or comparative in all respects because all smart cities and their public/private

[46] As will be seen throughout the Report, our distinction between regulation and governance is that the former is an intentional enterprise designed to influence and alter specific behaviours, while the latter is a wider understanding of social ordering that may incorporate regulation, but also includes other forms of ordering technologies that may not influence behaviour.

confluences have their own unique characteristics. Rather, it is intended to use the case studies as a snapshot into more general governance challenges that mass data sharing in modern urban environments presents.

While the Project has not involved extensive original ethnographic fieldwork, the research team has devised useful analytical frames to inform future empirical testing. As an alternative to representative empirical study, an achievable and creative methodological approach has involved drawing upon a regional network of scholars and policymakers engaged in a *conversation methodology*[47] through which experiential understandings were developed and case studies fleshed out under the wider research orbits. From this foundation it became possible to suggest novel relations of accountability, trust, respect and responsibility, and analyse how these are undermined in examples from mass data sharing discussed among the group. The next step involved identifying any new kinds of social ordering the confluence of data may generate and require, and speculating on emerging necessities for governing (such as a re-ordering of platforms) that may arise in the development of future 'citizen-centric' data governance policy and regulation. It would be overstating the capacity of this methodology to expect or assert that it produces findings on which specific recommendations could be based. Realistically, due to limitations in representativeness and coverage, the value of the method is in the suggestions it generates that would be the focus of more definitive and specific empirical study as subsequent opportunities for this emerge across the region. The origin and some of the main takeaways from applying this methodology is discussed in the following subsection.

Conversation methodology

In line with the project's citizen-centric core and in the spirit of participatory research,[48] engaging with experts on data and AI governance, and, at the same time, 'smart citizens' from across Asian region has been the highlight of the methodological agenda of this work. The conversation methodology has drawn upon the workshop on the *Accountability of AI in Smart Cities* (24-25 April 2023),[49] which explored Asian attitudes and approaches to accountability of AI and data use in smart cities based on hypothetical case studies that sparked a lively and engaging debate around universal and specific challenges to smart city governance in the region, and served as a springboard to the research roundtables on governing data sharing in smart cities.

The workshop attracted scholars and policymakers active in the fields of regulation of data-driven technologies, representing most of Southeast Asia and some South Asian jurisdictions. The workshop involved a creative use of conversation methodology based on two hypothetical case scenarios involving unregulated data-driven technologies, algorithmic governance, and data sharing. Through the general discussion of both general and context-specific challenges and approaches to accountability (such as lack of legal frameworks, infrastructure gaps, digital divides, low AI literacy, limited funds for research, few experts, untrained personnel, and issues with trust) as well as conversations held on its margins, the workshop allowed us to broaden our understandings of accountability and brainstorm ideas of how accountability can be boosted and operationalized, as well as appreciate the challenges involved in finding a model that may work for most of the region. As almost every jurisdiction in the region had a representative able to highlight context-specific challenges and difficulties, the discussion was cross-sectional and nuanced.

Following the workshop, some of the participants sharing our concern about how pervasive data sharing and merging in smart cities puts a strain on traditional styles and models of governance agreed to continue the research conversations through a series of research roundtables devoted to comparative governance challenges across different Asian contexts and settings. Through these conversations, stimulated via mini-case studies and facilitated by some research questions, we have been able to deepen the research understanding of the respective governance and cultural contexts. This allowed us to put together a base of regional examples concerning public and private data sharing from across the region, gain useful feedback on our thinking, and test our ideas on governance models (frames) for mass sharing of data between the public and private sectors in surveillance and public service delivery that could ensure protection of citizen (data-subject) interests.

The first of the research roundtables (13 July 2023) discussed government-sponsored or developed apps involving crossovers with the private sector (in particular, super apps); a source for concern in different settings due to their nature and the challenges they pose to governance. The national digital identity project in Singapore and the linked applications (Singpass and LifeSG) were used as a case study and a starting point for the discussion regarding the different priorities and challenges to data integrity across the region, which progressed to cover insights into topics such as:

[47] This methodology is explained in more detail in the following subsection.
[48] See, for example, the rich body of work on Participatory Action Research (PAR), such as the work by Cornish et al. (2023) which defines this research approach as 'bring[ing] together community members, activists and scholars to co-create knowledge and social change in tandem', in order to 'interpret and address complex systemic problems' through 'collaborative, impactful, contextually situated and inclusive efforts' and in a manner 'prioritiz[ing] the expertise of those experiencing a social issue, and us[ing] systematic research methodologies to generate new insights'.
[49] This workshop was co-organized by the CAIDG SMU and Tilburg Institute for Law, Technology, and Society, and hosted at the Singapore Management University.

- The challenge of lack of knowledge and awareness of digital technologies among the public (Indonesia, Thailand);
- The challenge of data sharing not being regulated in one single document (Singapore);[50]
- A variety of social attitudes towards private technology in the public sphere ranging from convenience to dependency, and from trust to 'suspension of disbelief' (Indonesia, India);
- Attitudes to loss of data and in some cases privacy vis-à-vis feeling violated or manipulated (Korea);[51]
- Different 'starting points' across the region in terms of regulatory approaches to trust and regulation of the public/private interface (Korea);[52]

The second research roundtable (17 August 2023) involved a presentation of super apps developed from nation-wide contact-tracing apps in Indonesia and Malaysia. The discussion covered topics, including

- Omnipresence of the 'data-for-service' model and narrative (which implies loss of control over one's data and is thus a major departure from the original vision of its main architects for the internet as a free and open platform (Berners-Lee, 2017));[53]
- The additional local challenge of another narrative which is 'data for national development', or 'data for the community', now supported by exemptions in the data protection legislation (India);[54]
- The viability of to implement a different level of information obligations based on the level of sensitivity of data;[55]
- The challenge of tech colonialism in the Global South countries;[56]
- The difficulty in specifying the legal basis for transparency/information openness on data sharing even in countries such as India which have a data protection legislation, a constitutionally recognized right to privacy, and a right to information legislation.

One of the most fundamental outcomes from the conversation methodology with project network partners is the realisation that across the region, smart cities and their citizen/residents approach data sharing and integrity from different starting points and face different contextual challenges. These variations require understanding if a workable governance model (frame) on data sharing is to have its maximum purchase in these particular settings.

Method for modelling in Part C

To reiterate, a central analytical assumption (underpinning) of this research is that mass data sharing between public and private service providers in smart cities (and the associated surveillance possibilities) present challenges to the integrity of personal data. Working to address these challenges is a governance enterprise. Data governance is a very broad field and to refine its consideration here, we are interested in how data access, management and use can be regulated so as to better ensure that appropriate interests of citizen/residents are respected, and that these stakeholders are meaningfully included in any governance policy. Such an interpretation of governance requires that mass data sharing should be an orderly, accountable and responsible process in which the data subject is an active participant. The achievement of such governance objectives builds upon the following reflections which been informed by this research:

---

[50] Singapore Participant (research roundtables): 'data sharing is caught up between different laws and policies, and some confidential documents on public service. On the other hand, the PDPA [of 2012] exempts the government from its scope'.
[51] Korean Participant (research roundtables): 'People [in Asia] may be less concerned about the data itself, than whether they will feel violated in a certain way. Although Asians don't feel violated in the same way that Europeans or Americans do, they will if they feel manipulated or if they realize that that data has been weaponized against them; used to their detriment. Privacy in Asia is different from privacy in the US and EU and maybe in a sense it's more about autonomy even – but the idea that data sharing involves a privacy infringement is a foreign concept.'
[52] Korean Participant (research roundtables): 'the government had been collecting large quantities of data even before they had been put to pervasive uses, e.g. at airports. People have learned to expect that their data will be collected ([which is] not the same as cynicism), and they don't feel violated where the feeling is 'the government had so much data on me anyway'. At the same time, with the privacy law being very strict, people have learned that there are firm guarantees; there is also a minimization amendment on the way. So even if a common person in the street does not know what the privacy law says, there is a lot of built-in trust in the government, and a lot of reliance on government services, and people are willing to accept a higher threshold.'
[53] In the words of Tim Berners-Lee (inventor of the world wide web), 'the current business model for many websites offers free content in exchange for personal data. Many of us agree to this – albeit often by accepting long and confusing terms and conditions documents – but fundamentally we do not mind some information being collected in exchange for free services. But, we're missing a trick. As our data is then held in proprietary silos, out of sight to us, we lose out on the benefits we could realise if we had direct control over this data, and chose when and with whom to share it. What's more, we often do not have any way of feeding back to companies what data we'd rather not share – especially with third parties – the T&Cs are all or nothing' (Berners-Lee, 2017).
[54] Indian Participant (research roundtables): 'In addition to the narrative "your data is something that you must trade for your benefit," it's now become a big narrative of "the country's data is something it must sell for the country's benefit", and there's no way to really argue with that. The authorities say that it's fine if all of your data is gone, because thanks to that data, the domestic developers will come up with apps, and even if those apps could cause breaches of privacy, that's all fine as long as they are people from within the country. And so people know that there is sharing, but they don't actually know how it exists, what the sharing will be for, [and] what app they're planning to build with it.'
[55] Singapore Participant (research roundtables): 'Some personal data, e.g. health or financial data, may be more valuable than other kinds of data. (…) Proposing such a hierarchy of data might be more practical than trying to secure and govern everything with a similar approach, which may be very difficult to do in practice. If in a small city like Singapore it's already quite resource-intensive, a big country with more than a billion people will be infinitely more complicated and resource-intensive, also in terms of obligations such as making a citizen aware of how their data is being used.'
[56] Indian Participant (research roundtables): 'in global majority countries, the systems in use there are usually designed, trained and developed in Global North countries. So when you run them in global majority countries such as India, they tend to glitch and create a lot of bias for socioeconomic conditions that only exist in India and nowhere else. So in India, there has been a lot of caste discrimination based off of people's names, because the tech wasn't designed for it.'

- Understandings of the challenges posed, both contextual and universal;
- Identification of responsible stakeholders in the data ecosystem, and the duties and expectations they have in orderly data use;
- Options for key elements in any governance strategy which operates respecting the interests of citizen/residents;
- Proposals for ensuring that any such governance policy is compatible with Rule of Law principles and operates with data subject inclusion as a key feature of its sustainability.

Modelling a governance strategy/policy in this way is both a process of explaining the need for governance intervention, and offering potential solutions to challenges, recognising that policy and implementation will be dependent on pre-existing and prevailing regulatory preferences and capacity. Modelling is not the same as framing policy or determining implementation of regulatory policy. Rather, the purpose of our modelling was to construct a representation of what mechanisms of ordering data sharing could work for the benefit of legitimate citizen/resident interest.

Connecting thoughts: towards Part B

The first part of the Report, Part A, has generally established the purpose and significance of the project. The following part, Part B, will place in context the specific governance challenges generated by mass data sharing in smart city administration through merging public/private accountability and responsibility practices. These 'snapshots' discuss risks posed to the data integrity of citizen/residents by routine sharing of massive amounts of data in the everyday delivery of services, the mounting of surveillance and the incursions into the autonomy of data subjects through ever-expanding data dependencies. While each context differs in nature, extent, form, and consequences of data sharing, and as such the impacts on citizen/residents in each situation are specific, what nevertheless emerges are some common themes such as information deficit, absence of transparency, confused lines of accountability and power asymmetries.

It is hoped that through Part A's description of the explanatory theory which is applied in the analysis developed in Part B, the methodology used in analysing the use cases in the snapshots, and the elaboration of key terms and concepts, the reader will better appreciate the significance of the challenges revealed in Part B, and take ownership of the governance model offered in Part C.

Part A should not be left without certain key clarifications:

- This research project is limited in considering governance themes as they arise through public/private data sharing in smart cities, which are heavily dependent on the respective stages of digital transformation.
- The Report is neither a critique of mass data sharing as such, nor an attack on digitizing the delivery of city services. It is not an indictment of public administration channelled through private providers which has been a feature of urban development even before the 'smart city'.
- Where there is cautionary discussion of challenges to data integrity, the lens is directed towards public and private data sources from the perspective of data subject interests (citizen-centricity).
- The Report accepts that other data protection policy operates in those cities referred to herein, and that private sector best practice and ethical ascription may be present. That said, there is no specific governance modelling focused on mass public/private data sharing, and Part B identifies this as requiring remedy.
- The research was restricted by either limited information on the public record of the practices covered in Part B, or by contradiction and confusion surrounding reporting responsibilities, in particular. In addition, the snapshots, while indicative of governance challenges, are not representative of all data sharing occasions, and are not identified as the only problematic data sharing instances.

Part B: Challenges in governing mass data sharing. Snapshots and main themes in the governance of mass data in SE Asia

Introduction: Orbits and snapshots

Against the backdrop of the global Covid-19 pandemic, this research was inspired by concerns about the data governance challenges involved in the unprecedented transfer of many essential public services to the online sphere, where they would be delivered within a network of public and private operators, and the concomitant mass data sharing. This initial overarching interest in large-scale data sharing within the three orbits (large areas of service delivery which benefit the citizen), i.e. edu tech, health tech and gov tech, was eventually narrowed down into a number of empirical snapshots to sketch out the challenges involved in mass data sharing in the three orbits, based on particular empirical examples of public-private partnerships. The snapshots, some of which are anchored in more than one orbit, are meant to provide a window into the data sharing within the three orbits, and at the same time illustrate many of the problems experienced within platform governance globally. This research was primarily interested in mass data sharing through the amalgam of public sector service delivery agendas and mostly private sector information platforms. The description and analysis is enriched with some limited discussion of particular challenges that may be encountered throughout the region, or are snapshot-specific. Indeed, within those snapshots, we tend to focus on concerns and problems posed to good governance of data sharing, both established and likely to occur.

In other words, the primary purpose of Part B of the Report consists in uncovering the obstacles that could stand in the way of *good governance of citizen data*, understood as active processes that make data users accountable and protect the interests of the data subjects so that their data is accessed and used responsibly and respectfully. The discussion of challenges is not designed to address any policy applications beyond specific issues of data use, and should especially not be read as a critique of using technology in service delivery, or in government. This also means that we are not primarily focusing on the undoubtedly significant public benefit in having some of those partnerships implemented, but rather, focus on identifying areas where good governance of citizen data might encounter difficulties in the context of such partnerships. We are also not primarily interested in unauthorized sharing of data (i.e., sharing resulting from cybercrimes, for example), but instead, in its intentional sharing with partners, e.g. for the purposes of developing new products, research, and marketing purposes.

The following, empirical part of the Report is based on desktop research and enriched by the use of a conversation methodology (two research roundtables), supplemented by individual conversations with stakeholders and researchers. The snapshots that follow vary in detail and in emphasis due to the availability of data, its nature and detail. For that reason, in describing some of the snapshots it was only possible to focus on one or two major themes such as the challenge of sharing user data with advertisers, or on how potential conflicts of interest that may manifest in such public-private partnerships might impact the good governance of data. We have decided to include different levels of detail, considering that all snapshots raise challenges that range from obvious or definitive to suggestive and conjectural. The spectrum from well-developed to suggestive, is opened with Snapshot C (LumiHealth), followed by Snapshot A (SatuSehat), Snapshot E (Ruangguru), Snapshot D (Pair), Snapshot F (DELIMa), and closed with Snapshot B (MySejahtera).

Snapshot A: SatuSehat (Indonesia)

We ensure this platform [SatuSehat] will become a kind of 'toll road' where there are many entry gates from various parties…
Deputy Chief, Digital Transformation Office, MOH

Introduction

SatuSehat (meaning 'OneHealth', or 'OneHealthy') is a health 'super application' (super app) – an integrated health data exchange platform developed by Indonesia's Ministry of Health's Digital Transformation Office with the help of ad hoc tech consultants recruited primarily from start-ups. The platform adopts PaaS (Platform as a Service) infrastructure model connecting the entire health industry ecosystem and all facilities across the country (from pharmacies and clinics to laboratories and hospitals) for more efficient national health data exchanges (Mulyanto, 2022). SatuSehat has been directly converted from the Covid-tracing app PeduliLindungi, and incorporates some identifying features typical of government (ID function) as well as health apps.

Referred as a 'health data ecosystem', 'health data bank', 'citizen health app' by government officials and the media, it is intended to help transform Indonesia's health sector. According to the Minister of Health Budi Gunadi Sadikin, besides strengthening the integration of data and simplifying apps previously in use, the platform is also meant to build an innovation ecosystem (Mulyanto, 2022), in addition to increasing transparency and fight corruption.

Data sharing

Primary sharing

A super app like this is the front end of a platform where internal developers and third-party providers can publish mini apps (or mini programs) for users to activate as needed. Data sharing is thus key for a super app. SatuSehat will involve many partners and the sharing of large troves of confidential and potentially highly sensitive data including the ID card number, as well as medical and travel histories. The platform is to bring all services and complete health data (not only medical, physical condition and activity data, but also demographic and potentially genomic) together with analytics and trend predictions at the fingertips of Indonesians, for them to easily and conveniently access and share it with medical providers.[57] Users will equally be able to access services such as telemedicine, or book doctor's appointments through the app.

The key objective of the platform is the promise of a more efficient national health system and improved individual diagnostics through implementing and integrating EMRs (Electronic Health Records). Once the currently ongoing process of integration with the platform is completed, medical professionals and health facilities staff will be able to input patients' diagnoses and results into the digital system, to be shared on the platform and made accessible to the patients (via SatuSehat Mobile – the smartphone app), as well as for further sharing on the basis of consent (e.g. with referral clinics). Other information (including Personally Identifiable Information, or PII) and some health data is input manually by users in connection with setting up their SatuSehat user account (Kementerian Kesehatan, 2023a, para 4). The platform also collects other categories of data such as location, app usage information, data from camera- and Bluetooth-based activities 'to optimize the functions, features and use of SatuSehat according to SatuSehat's objectives' (Kementerian Kesehatan, 2023a, para 4). Location is used for contact tracing.

Data can be retained and used beyond the deletion of a user account, for a minimum of 25 years (Kementerian Kesehatan, 2023a, para 5). As per the information provided by the developer, i.e. the Ministry of Health, it appears that it is not possible to request the deletion of one's data from the platform (Google Play, 2023a), although the Personal Data Protection Law (PDPL) admits the user deletion requests. By contrast, this option is seemingly available in the Malaysian MySejahtera Covid-tracing app (GooglePlay, 2023b).

Secondary sharing (data reuse)

According to the Privacy Policy (PP), users' personal data (which includes health data) may, among others, also be used for purposes of statistical analysis for policy making; for health surveillance in the context of early detection and/or control of infectious diseases, non-communicable diseases, epidemics and extraordinary events by the Government, and for health efforts that are promotive, preventive, curative and rehabilitative as well as other purposes as long as they are permitted based on statutory provisions (Kementerian Kesehatan, 2023a). Though that is not clear from the Privacy Policy itself, statements given by government officials, who have repeatedly expressed interest in attracting start-ups and investors to build the Indonesian healthcare sector together[58] may suggest that health data of users may also be used for research and innovation. While in light of the regulation on medical records, accessing electronic medical records for purposes of education and research does not require patient's consent,[59] as confirmed by that regulation, that data ought to still belong to patients,[60] who might want to know how it is being used.

Governance concerns

It is certainly positive that the main documents about data collection and sharing, the Terms and Conditions and the Privacy Policy, are freely available on the MOH website. The T&C (of which the PP is part) are said to constitute a binding agreement between the user and MOH.[61] On the other hand, the document is available in Bahasa Indonesia (with an automatic translation into English provided). In spite of being the official language of this second most linguistically country in the world (Ethnologue & the World Bank, 2021), only a minority of Indonesians speak Bahasa as a first language.[62] However, the list of uses of data in the Privacy Policy is seemingly non-exhaustive; while it mentions health service implementation, health information provision, surveillance in the context of infectious diseases and analysis of

[57] The data will include patients' contact-tracing data linked to the Covid-19 pandemic, their medical histories, past procedures, lab records, exam results, scan images, immunization records and pharmacy purchases from their smartphones, as well as diet, allergy, and activity data combined with individual insights (patients will be able to share the data from third-party tech such as Samsung Health and Apple Watch with the platform to enable a more holistic view of their health).

[58] 'My focus is to ensure we can achieve a better quality, more accessible, and more affordable healthcare for all. Recognizing the urgency of healthcare investment, we openly collaborate with the private sector, and the government will support through the regulations to facilitate investors, therefore we must communicate with each other. [The organizer] has brought together notable investors and entrepreneurs to this event so that we can explore recent innovations and potential investments to work on for the greater good of Indonesia's healthcare,' said the Indonesian Minister of Health (Foundry, 2023).

[59] Regulation of the Minister of Health of the Republic of Indonesia No. 24 of 2022 on Medical Records, Article 35 para 1 letter e.

[60] In saying 'belongs' we are not entering the complex debate about whether data can be owned. Rather it is meant to refer to the control of personal data by data subjects. See also fn 82 and corresponding text in the discussion of Snapshot B.

[61] The MOH authority asserts that the agreement is binding, but it is not clear what legally actionable status this infers.

[62] Indonesian Participant, research roundtables.

personal data to obtain statistical information for policy making, it also refers to 'other promotive, preventive, curative and rehabilitative health efforts', and 'other lawful purposes' (Kementerian Kesehatan, 2023a, para 6.12). The manner of patient data sharing and resharing, especially who the data may be shared with (for example, who will analyse the data, where and in what form) is not entirely clear from neither of the documents (Kementerian Kesehatan, 2023b, para 5).[63] It is also not apparent what happens with that data once the consent for sharing is granted, for example, whether and to which extent a health facility can share data further, and in what circumstances. For example, will every employee of a health facility be able to access patients' medical records? Will physicians be allowed to save and store patient records on their private devices?

Considering the Privacy Policy's broad formula potentially permitting ulterior uses, one may wonder if the data will be shared with third parties such as academics or health start-ups or other private partners, and on what conditions (for example, it may be observed that in the LumiHealth app as described in Snapshot C, such uses of data are subject to a separate consent). The information on sharing appears possibly too limited and vague to allow accountability for data use,[64] as well as genuine and actual participation opportunities such as via truly informed consent. More broadly, the Indonesian public is not too well informed about digital transformation initiatives in Indonesia,[65] which makes good governance of data sharing through user oversight more difficult. With regard to SatuSehat, a digital transformation officer at the Ministry of Health said the government would need to promote outreach programmes to educate the community about electronic medical records (Arlinta, 2023), which may again encounter problems considering that Bahasa is not spoken equally throughout the country.

As a super app, SatuSehat will ultimately integrate multiple third-party apps and stakeholders which will access user data. In light of the Privacy Policy, 'users are advised to study and read carefully the personal information handling policies applicable to Third Party Sites/Applications and/or Third Party content' (Kementerian Kesehatan, 2023a, para 10). However, it may not be a reasonable expectation to require the data subjects to consult every privacy policy in a constellation of data sharers linked to the platform (health facilities and potentially many others).

With concerns about the security of data taking centre stage in the wake of high-profile citizen data breaches (Loasana, 2023), the issue of how and why the data will be shared on the platform as a matter of course seems to be given somewhat less attention. Raising questions of accountability for data use, the data that might be reshared with third parties includes large amounts of potentially highly sensitive data including sexual orientation, and data on stigmatizing illnesses such as HIV infections. It is not clear how sensitive data such as this will be protected from uses with which a patient might not be comfortable. Misuse of such data may generate particularly problematic consequences in a morally traditional society such as Indonesia, with its strong views of homosexuality as prohibited by religion (SMRC, 2018)[66] and a mixed record of upholding LGBTQ+ rights.

There may be particular concern in terms of what accountability can be expected of for-profit entities which may be party to data sharing. Specifically, will it only be directed to the patients, or will there perhaps also be a conflicting responsibility towards the owners of the business, who might profit from it amassing more data (expanding accountability locations from personal to market interests)? More broadly, an undertaking of such scale (storing data of potentially well over 200 million people, with many stakeholders, uses, and multiple entry and exit points) may encounter practical difficulties in ensuring that the data will be used solely for the declared purposes, and accountability sufficiently required of stakeholders given access to confidential, detailed and potentially very coveted population-level datasets which could not only improve health governance, but also advance research and innovation, offering significant commercial profits. Lessons learned from well-known public-private partnerships in health tech such as the Streams app developed by Google DeepMind at the behest of the UK NHS Royal Trust reveal that the scope of sharing of the health data of citizens even in smaller-scale partnerships may be broader than originally communicated (Powles & Hodson, 2017; Dickens, 2023). In the case of SatuSehat's direct Covid-tracing predecessor, the PeduliLindungi app, it proved problematic for the public sector to ensure the good governance of health data where private tech companies

[63] In light of the T&C, related to SatuSehat's purpose, the user may give consent to access data (such as medical record data, laboratory results, vaccination history, immunization, among others, and/or features such as the activation of location, camera, storage, Bluetooth (Kementerian Kesehatan, 2023b, para 5).
[64] Accountability to the patient is at issue, but what about accountability to any relevant external rule frame and/or regulatory agency? It is not sufficient to see accountability as having a single referent when the issues involved are relevant to the responsibility or concern of other important players.
[65] Indonesian Participant, research roundtables.
[66] In a survey of 1,220 respondents with voting rights (17 years old and above), 47.5% answered 'Agree' and 34% answered 'Strongly Agree' (SMRC, 2018).

were involved.[67] It is fair to speculate whether in this instance the government will be able to ensure accountability of private partners linked to SatuSehat, when it is presumably dependent on them for technical expertise, know-how and infrastructure.[68] SatuSehat's cooperation with private tech providers has given rise to voices of concern about possible arbitrary access and misuse of patient data (which unlike the PeduliLindungi data is now more comprehensive and shared across many health facilities), and its reuse in targeted advertising, for example from insurers or the pharma industry (Prabowo, 2023).

In terms of the legality of processing of data, the Indonesian Law No. 27 of 2022 on Personal Data Protection (the 'PDP Law') states that everyone who processes personal data must first obtain permission (consent) from the owner of the personal data. However, in light of the SatuSehat Privacy Policy, 'the legal basis for data processing is carried out based on statutory regulations in the context of implementing the duties and functions of the Ministry of Health, and for some processing [it] is carried out based on the patient's consent' (Kementerian Kesehatan, 2023a, para 3.)

To an extent, user consent within the SatuSehat platform appears to be designed in a problematic way. Firstly, as a generalized health platform, SatuSehat has repurposed the data initially collected by the controversial PeduliLindungi app on a surveillance basis[69] apparently without asking for the patients' separate consent; previous users of PeduliLindungi have only needed to update the app for it to transform into SatuSehat, and agree to the new Terms & Conditions (Rokom, 2023). Yet, transformation into a generalized citizen health data platform involved a significant modification of the original purpose behind the data collection, which was to contain the Covid-19 pandemic within the community. This could have easily warranted seeking a separate consent from the data subjects.

Secondly, it appears that consent for treatment in a public health facility (and therefore under public insurance) is inextricably linked with consent for the sharing of data with the government and SatuSehat (in private clinics, consent for sharing is to be sought separately).[70] On the other hand, the MOH Regulation no. 24 unequivocally states that health facilities must open access to the entire contents of the patient's electronic medical record to the Ministry of Health.[71] Indeed, voices from the civil society have already expressed doubts whether data subjects' consent for this will be sought (Loasana, 2023).

Thus, in each of these instances, the request for consent to share data is bundled with another request that patients may feel strongly incentivized to grant in order to access medical care and improved, more holistic diagnostics, as well as for reasons of convenience. According to the Office of the Australian Information Commissioner, bundled consent has the potential to undermine the voluntary nature of the consent (OAIC, 2022, paras B48-49).[72] The absence of alternative options offering these services or affordable competitive service providers further complicates any consideration of free choice by the data subject/patient. While the option of remaining off the platform remains in principle open, aggregator apps leave little feasible alternative to those wishing to remain off the platform (perhaps especially in the field of health), which may be thought to further undermine voluntariness in use and consent.

It should further be observed that SatuSehat mobile still tracks its users to handle the spread of Covid-19 and other diseases in the territory of Indonesia (Kementerian Kesehatan, 2023b, para 5), although the coronavirus disease is no longer considered a pandemic.[73] This feature of the app may raise Rule of Law questions, as justifications no longer prevail over universal rights and liberties beyond a clear emergency context. While the Constitution of Indonesia (1945) does not explicitly enshrine a right to privacy, it mentions related concepts such as a right to the protection of dignity, and to feel secure.[74] On the other hand, Law no 39/1999 on Human Rights states that Indonesia has a responsibility

[67] To increase the uptake of the app, the PeduliLindungi QR code was integrated with 15 large apps such as Gojek, Shopee, Tokopedia etc, allowing these platforms to access citizen data before it was handed over to the PeduliLindungi app (as a matter of fact, it is not even clear if it ever was). As a consequence, a human rights association brought an administrative complaint against the Minister of Health (case no. 102/G/2022/PTUN.JKT) saying that his decree which originally made these public-private collaborations possible violated the right to privacy of personal data that should be protected by the state. The complaint was ultimately rejected – which may or may not have been due to the emergency character of the app. It should be noted that in the case of PeduliLindungi, the additional motivation for the tech unicorns to integrate the QR codes (apart from the additional data) was the reopening of the economy after the pandemic – a true confluence in more than one sense.

[68] Like many other countries, Indonesia has been forced to rely on the tech sector during and immediately after the Covid-19 pandemic. Separately, it has recently opted for business-friendly approaches in regulating platform and digital service providers (Drazewska, 2023b).

[69] The design of PeduliLindungi was criticized for requesting personal data that was not essential for Covid-19 tracking and tracing (Idris & Pitaloka, 2022). Separately, the app's reliance on private tech companies was accused of illegality in connection to risks to user data privacy (see fn 67). The app has also suffered numerous large-scale data breaches resulting in leaking of the health data of President Joko Widodo, among others.

[70] As per the original information received from two Indonesian contacts. Considering the lack of publicly available information on this issue, confirmation has been sought but has not yet been received.

[71] Minister of Health Regulation No. 24 of 2022 concerning Medical Records, Article 28.1.

[72] The OAIC defines bundled consent as the practice of '"bundling" together multiple requests for an individual's consent to a wide range of collections, uses and disclosures of personal information, without giving the individual the opportunity to choose which collections, uses and disclosures they agree to and which they do not.'

[73] According to the Terms and Conditions of SatuSehat, 'When the User downloads the SATUSEHAT [app], the system will ask for User's consent to activate location, camera, storage, and Bluetooth. With the location active, SATUSEHAT will periodically perform tracing and tracking of the User's location to provide information related to the crowd, and the user's health status. The results of this tracing and tracking will be used by SATUSEHAT and the Government of the Republic of Indonesia to identify anyone who needs to receive further treatment so that the handling of the spread of COVID-19 and other diseases can be carried out' (Kementerian Kesehatan, 2023b, para 5).

[74] 'Every person shall have the right to protection of his/herself, family, honour, dignity, and property, and shall have the right to feel secure against and receive protection from the threat of fear to do or not do something that is a human right.' Constitution of Indonesia (1945), Article 28G, para 1.

to honour and implement the Universal Declaration of Human Rights, which includes the right to privacy.[75] Equally, the platform has also integrated the surveillance data collected by PeduliLindungi (the Covid-tracing app). Though rebranded as a new, generalized health app, the platform thus still carries its predecessor's data, including user profiles, Covid-19 vaccine certificates and tickets, antigen and PCR test results and QR codes (Arlinta, 2023) obtained primarily during the public health crisis that was the pandemic.

Public health is a field in which informational norms limit access to medical record information to respect patient confidentiality. Article 47 para 2 of Indonesia's Law no 29 of 2004 on Medical Practice states that a patient's health record can only be accessed by doctors and heads of medical facilities, who must nevertheless keep it confidential (Regulation 24 of 2022 on EMRs further empowers the latter to approve access of the medical facility's personnel). But the SatuSehat platform will enable a much wider circle of people to access patient data. On the one hand, the EMRs and other patient data will now be accessible to many more health facilities and potentially other stakeholders – an Indonesian MP has recently expressed concerns about the risk of third parties arbitrarily accessing the medical records of patients, which she saw as a human rights issue, as highlighted by Prabowo (2023). On the other hand, where many Indonesian families and communities of friends or neighbours share the use of one smartphone and therefore may need to share one SatuSehat Mobile account (especially in areas like Papua, where only approximately 35% of the population has a mobile phone – not necessarily a smartphone – as per Statistics Indonesia (2023)), the security and privacy of personal health data could even unintentionally be put at risk. Linking multiple family or community members to one account – a functionality of the SatuSehat app – can certainly be useful for families when it comes to children and the elderly, but it may equally involve multiple occasions for breaches of confidentiality which might otherwise not occur, and responsibility for which will be placed on the user (Kementerian Kesehatan, 2023a, para 10). In that sense, SatuSehat can significantly expand access to otherwise restricted information on both ends of the spectrum between private commercial access and community/family exposure.

The reality of stark digital divides across a populated and diverse country like Indonesia presents challenges of equitable access as well as data protection. It is difficult to ensure universal participation in this nation-wide initiative and enjoyment of its benefits by communities which lack access to the necessary technology, for example, if their customs prohibit it. For example, the Baduy community in Java has recently made headlines in connection with their petition to the government to cut off their internet access in an attempt to maintain their traditional way of life (Rayda, 2023). While equitable access is not a data integrity issue, it represents another variable in an uneven landscape of data generation, use and risk.

Challenges[76]

- As a super app, SatuSehat will most likely involve sharing of data on a large scale. The main concern is that the users may not be able to access sufficient information about how their data may be shared. On the one hand, the Indonesian public tends not to be too well informed about digital transformation initiatives such as SatuSehat, and remedying this issue may encounter difficulties because of language barriers and levels of digital literacy. On the other hand, the platform's privacy policy (available in Bahasa and in English) does not clearly and exhaustively state whether, and if so, how, the SatuSehat data may be reused, with whom it may be shared, and subject to what limitations (statutory or contractual). Yet even allowing users to better understand 'how exposed they are', and a list of entities their data might be shared would already involve a simple but meaningful change in the policies of many platform-based services (Litman-Navarro, 2019). *The challenge is transparency and the response should not be simply more information, but along with that a concerted effort from the data holders/sharers to ensure that data subjects will understand the ramifications of entering their data, involving a reflection on natural barriers to such understanding.* The model presented in Part C recognises this need and provides opportunities for data subjects to ask about the meaning of information on data access and use.

- If information on the data practices of those third parties can be presented in a concise and approachable way (even if multiple parties are involved in the sharing), more users ought to be able to understand it without significant hardship. By contrast, simply referring users to the usually complex privacy policies of potentially multiple third parties (assuming they can be identified) may place an unreasonable burden on users, particularly when most are either not digitally well skilled or their knowledge of data and protective expectations is limited. *The challenge is to make clear that the responsibility to inform and explain data access and use rests with the data-powerful, and as such they need to enter into meaningful negotiations with data subjects (as referred to in Part C), based on effective information dissemination.*

- The known limits of the notice-and-consent approach to citizen data may well suggest exploring alternative ways

[75] UN General Assembly, Universal Declaration of Human Rights, 10 December 1948, 217 A (III), Article 12.
[76] At the risk of repetition each snapshot will contain a section summarizing challenges (both from contextual and then more lateral/universal analysis and link these to governance imperatives and options. Of necessity, and particularly where common challenges emerge across contexts this might appear as repetitive or unbalanced in favour of which snapshot is positioned first and out of which many of these challenges are initially apparent.

of stakeholder engagement. The model in Part C is based on dynamic engagement and not a once-and-for-all consent regime. However, within the notice-and-consent model itself, there may be ways to improve the design of consent so that it maximizes the autonomy of data subjects rather than being a gateway to boundless data sharing. For instance, broad consent can often be linked to generic, vague language without sufficient details or examples (for example, where users consent once to their data being shared with unknown 'third parties', for a number of broadly described purposes, rather than consent being sought each time information is used or shared.) Observed in the majority of snapshots analysed, these information deficits have an impact on accountability (understood as providing a sufficient explanation for the use of user data) and on the extent to which consent may qualify as informed (as partial disclosure can only result in partially informed, broad consent). *The challenge, therefore, in each context as well as more generally is to move beyond transparency, and to back it up with accountability attribution.* The model presented in Part C recognises that at present, data access/use arrangements typically represent significant power asymmetries. One important way to level these is for the data-powerful to recognise their accountability to data subjects and participate in meaningful negotiations as part of the accountability process. In so doing, accountability works towards enhancing trust relationships on which good governance depends.

- As illustrated by the SatuSehat example, the design of consent for data sharing is typically a top-down decision of the more powerful parties who have an interest in the sharing taking place. In most cases they will be the ones deciding when consent will be asked for, and when it will be presumed (this problem is of course not limited to the health tech industry, but routinely encountered in users' interactions with powerful and dominant tech providers such as Apple or Facebook (Doctorow, 2023)).[77] In order not to presume consent, it might be worth reflecting whether it might be split into separate consent requests, for every instance of sharing that users may not be comfortable with, such as e.g. linking consent for medical treatment with consent on sharing data on the platform, or consent to use a platform interwoven with consent to repurpose data previously collected under unusual circumstances and for another reason. In the absence of other opportunities for more extensive stakeholder engagement and participation such as co-creation (which is anticipated and explained in the model in Part C), asking clearly and separately for active and free consent for individual aspects of data sharing, especially those potentially contentious, may offer a true means of correcting power imbalances and informed participation (empowerment through information). *The challenge has two parts. Will consent be a sufficient governance boundary to protect the legitimate interests of all stakeholders, and if so, what basic requirements should be in place to see that consent is actual and incremental?* It should be observed that model in Part C is not reliant on consent, preferring to work from meaningful engagement to form relationships of trust.

- Unlike some other parts of SE Asia where the data protection legislation is only addressed at the private sector, the Indonesian government is bound by the recent data protection legislation (PDPL). While there is no standalone constitutional right to privacy in Indonesia, it is possible to link it to legislation enforcing international human rights obligations. Confidentiality of patient information is also mandated by special provisions. Thus, it becomes crucial for the government to ensure accountability for the way data is shared and how that sharing is governed in connection with SatuSehat. One way to increase accountability would be to supplement and concretize the missing and vague information in the privacy policy. Accountability for data use connected with SatuSehat could equally be imposed through data sharing agreements. While such agreements are typically not public, the public interest in the sharing of sensitive health information could justify making some of this information available publicly. *The challenge here is the secrecy associated with private contracting and the limitations it poses on open governance (transparency and accountability referred to above).* The model in Part C requires that more powerful players mutualise their interests through negotiations with data subjects. This can entail introducing those players which formally remain outside of the contract to some of the stipulations which may affect them.

- As a good governance condition concerning data access and use, it is posited that sharing of highly sensitive information could be subject to additional constraints, such as the possibility to limit the sharing of information on sexual orientation through an additional opt-in or at least a clear opt-out option,[78] or obscuring that information through encryption (while at the same time regulating who has possession of the decryption keys). *The challenge here is whether some data might be too connected to the personal integrity of the data subject to be shared outside of narrowly circumscribed contexts.* The model focuses on data integrity for the data subject; therefore, concerns about what should not be shared freely would be important in negotiating respectful and responsible data relationships.

[77] As an example, C. Doctorow describes how since Apple had allowed iOS device owners to block Facebook surveillance (96% of them choosing to avail themselves of this opportunity), 'Apple continued to spy on those users, in precisely the same way that Facebook had, without telling them, and when they were caught doing it, they lied about it' (Doctorow, 2023).
[78] For example, see California Civil Code Sec. 1798.135.

Snapshot B: MySejahtera (Malaysia)

Introduction

MySejahtera is an application launched in April 2020 to help the Malaysian government to monitor and control the spread of the Covid-19 virus on the Malaysian territory through the use of digital contact tracing in tracking the movement of persons.[79] The app collects personal data of users such as name, IC (ID) number, contact number, email address, age, gender, ethnicity, and home address, in addition to data from health assessments, location data and QR code 'check-ins' at public places (Martinus, 2022). In addition to containing sensitive (health) data, such a data pool may reveal a lot of information about the app's users (whose number at one point was 38m) such as where they live, work, eat and who they meet with, and could be used to predict spending habits, personal preferences or behavioural patterns. While the Minister of Health has always assured that the government owns the data and that it will only be used for digital contact tracing (Salim, 2022), it would certainly be risky to let that data fall into unintended hands. According to the President of the Malaysian Bar, those risks include 'unregulated management and abuse of personal data collected, and at worst, possible breaches of privacy, social engineering, and data abuses affecting national security' (Cheah, 2023).

While initially developed for the purposes of Covid-19 control, MySejahtera (as was the case with PeduliLindungi, which later changed into SatuSehat) has expanded the contact-tracing function to other infectious diseases which currently do not have pandemic status (MySejahtera, 2022).[80] Thus, the app has integrated additional functions through which it continues to be relevant in the endemic phase of Covid-19. At present, the government is exploring various options for the future development of the app, which include converting it into an all-purpose health app for the entire Malaysian population, or a government super app used to distribute targeted subsidies, among others (Bernama, 2023; Lee, 2023). This last objective is linked to plans to enact the Omnibus Act in order to facilitate intra-government data sharing, and to build a centralized nation-wide database and platform called Malaysian Main Database (Pengkalan Data Utama, or PADU) meant to provide a 'granular view of households' and enabling the socio-economic profiling of Malaysians (Sharon, 2023). The future 'governance by data' tool has raised questions about data security, third party access and even potential Big-Brother-like surveillance (Gilbert, 2023).

Data sharing

Regarding the infectious disease contact tracing function, the MySejahtera app is meant to detect all contacts of persons identified as infected (i.e., whose data has been uploaded to the MySejahtera servers), although there is no horizontal exchange of data, i.e., information on infections is not shared between users themselves. According to the Editor-in-Chief of Code Blue (a non-profit human rights activist and health information portal), during the pandemic data was only extracted from MySejahtera for the purposes of contact tracing in case of identifying a positive Covid-19 case, and accessible to a handful of government agencies (Boo, 2020), although access could conceivably be broadened when the new legislation on data sharing is adopted, and with the rollout of PADU. It should be noted that although the data is technically provided with user consent, during the pandemic the sharing of personal and other information via public check-ins was made obligatory in light of the Standard Operating Procedures (SOPs) linked to the Prevention and Control of Infectious Diseases Act 1988, which treats refusal to provide information under the Act or related regulations as an offence (IDEAS 2022; Tong & Tay, 2023, p. 114). And although in light of the Privacy Policy, user interaction data is deleted after a period of time, that most likely means that just a fraction of the data collected by the app will be erased regularly.

Other functions performed by the app at present, which include clinic and vaccine appointment booking, telemedicine, emotional support hotline, psychosocial support information, and self-screening assessment (Tong and Tay, 2023), will also involve data sharing. Following the potential future conversion of the app, any additional data sharing patterns and data practices will depend on the purpose the new app will ultimately serve, and on the contractual arrangements in place.

Governance concerns

MySejahtera's uses of data (including when that data may be shared with third parties) have never been very clear (Martinus, 2022, p. 43; MySejahtera, 2022).[81] But more broadly, the app has been steeped in controversy and faced with waning levels of public trust almost from the outset (Martinus, 2022). Concerns over data privacy, security and third party involvement intensified after the government announced its plans to sell the app to a private company with alleged political and business connections to parties in the government (Zahiid, 2022). It was then that a report by the

[79] Relying on technology such as Bluetooth, GPS and other smartphone sensors, the digital tracing method allows to identify persons that were in close contact with infected individuals, who are then notified and usually requested to self-monitor or self-isolate to stop the spread of the virus within the community.
[80] These diseases include Dengue, Rabies, Measles, Hand-Foot-Mouth Disease (HFMD) and Tuberculosis (MySejahtera, 2022).
[81] 'MySejahtera may share your personal information with our service providers where we need to do so and only to the extent that your personal information is necessary for those service providers to perform their services or obligations' [Translation from Malay] (MySejahtera, 2022).

Parliamentary Public Accounts Committee (PAC) revealed that the app had been developed by a private company called KPISoft (soon renamed as Entomo Malaysia Sdn Bhd) as a CSR initiative (philanthropic project), whereby the Malaysian government was to start paying for the app after a free year-long service on 31 March 2021, and that there was no formal tender or contract between the parties (PAC, 2022). According to PAC, this initially 'free' period followed by the government paying a negotiated price for the app in particular appeared to have been meant to bypass the official public procurement route to a government contract. The price finally requested from the government to continue using MySejahtera was quite high, which may be viewed as evidence of a situation of dependency on a particular tool to provide essential public services by the Malaysian government.

Had a formal contract been in place, it might have shed light on who has control over data and the platform itself. Nevertheless, the Malaysian government has consistently maintained that it owns the app, the IP and the user data (Boo & Batumalai, 2022), and that the matter of ownership was settled in a non-disclosure agreement (NDA) signed with the Singaporean-owned Entomo on 1 April 2020. While the NDA is not public and therefore its exact contents are not known, it did not stop Entomo from transferring the intellectual property rights of the app to another private company, MySJ Sdn. Bhd., in 2021, granting it a license to use the KPISoft software and a perpetual and exclusive license to the MySejahtera app (including its further development and support), thus making MySJ Sdn. Bhd the government's partner in the initiative without asking for its consent (Tong & Tay, 2022). It therefore appears that, contrary to the government's claims, it does not own the app or the IP.

The situation is even less clear with regard to the data generated by the app which can potentially include personally identifiable data. As recently observed by the Institute for Democracy and Economic Affairs, 'until now, it remains unclear as to who actually owns the data as the various sources seem to contradict one another' (IDEAS, 2022). However, the crucial question here is not necessarily that of ownership of the data in the legal sense, but rather – to apply a broader outlook – the real-world *access* and *control* over the data.[82] Unless a thorough digital audit of the app is carried out, it seems possible that these rest in the hands of the private companies which control the software and the platform itself. That might also explain why the Malaysian government has continued the arrangement with MySJ to use the app for another two years in spite of the exorbitant 196m RM in taxpayer money it was asked to pay at the end of the initial 'free' CSR period (Hussin & Kamal, 2023). Not surprisingly, the PAC report recommended that the government must ensure that both the MySejahtera app and the data are fully owned by the government so that no third party may claim ownership, copyright, payments linked to the development of the app in the future, and that the data is secure and not misused by any party (PAC, 2022). Transparency International Malaysia equally expressed its concern over the lack of transparency in the ownership, as well as the lack of parliamentary supervision of the MySejahtera app, and possible dependency in providing essential public services (Basyir, 2022; Mohad, 2022).

If the data is under control of the government, it is worth noting that together with the state governments, the federal government is excluded from the scope of application of the Malaysian PDPA,[83] and hence formally not accountable for any failure to conform to its provisions. If the data is, however, controlled by a private company (or companies), the question is how the government may ensure its privacy and security, as well as accountability for any uses and misuses of the data (e.g. sharing it with data brokers abroad), especially without a contract in place. With any potential further conversion of MySJ into a (super)app in the future, the above mentioned challenges could relate to an even wider dataset.

Similar to the SatuSehat snapshot,[84] such potential conversion might raise the issue of legitimacy of repurposing data originally collected for digital contact tracing in response to a public emergency, and possibly impact public trust in the government's digital transformation initiatives. Additionally, it also raises the issue of retention – section 10(2) of the Malaysian PDPA would require that data no longer required for the purpose for which it was to be processed is 'destroyed or permanently deleted' (assuming the PDPA is applicable at all). One of the options for MySejahtera's future could see it reimagined as a super app with commercial features such as e-wallet, e-commerce, and multiple applications (Liew, 2021); in such a case leveraging that database, initially constructed through behaviours made compulsory by the government, to build a for-profit business could prove particularly controversial. But in terms of other plans for the app, those linked to the Ministry of Economy's ambition to launch a huge, national socio-economic database (PADU) meant to allow to 'identify households by their income, location, and commitments, to get a view of their net disposable incomes in order to calculate how much to give in subsidies' (Ramli, 2023), the question is who will control this inevitably much larger platform, the relevant software and data this time, and whether the government will be able to ensure accountability for the data sharing in what it hopes will be a new era of 'governance by data'. Separately, these far-reaching plans necessitate ensuring the high quality and reliability of the data as a matter of data justice (Taylor, 2017). This may be linked to ensuring sufficient data security, which has been a notorious concern with the MySejahtera app. But perhaps more importantly, with a view to future potential uses of that large database, it would be crucial to provide avenues of access to justice in the event of algorithmic discrimination, e.g. with regard to automated decision-making on socio-economic scoring (Yeung, 2022), for which access and control over the system will be essential.

[82] On the difference between 'data ownership' understood as an exclusive right to data, and 'data access', see Drexl et al (2016), and Thouvenin & Tamò-Larrieux (2021).
[83] Personal Data Protection Act 2010 (No. 709 of 2010), Sec 3 para 1, https://www.pdp.gov.my/jpdpv2/assets/2019/09/Personal-Data-Protection-Act-2010.pdf
[84] See Snapshot A.

Challenges

- The uses of data in MySejahtera have not been fully transparent during the pandemic. Since then, the uses of the app have expanded to tracking other infectious diseases, and may be expected to expand further. But the MySejahtera snapshot emphasizes the need for transparency and accountability in the use of health data more broadly. While the government guaranteed this data would be used for the specific purposes of contact tracing only, ensuring accountability for uses of this large dataset, which includes sensitive data, may be difficult if it is in fact in the hands of private providers, whose 'move fast and break things' credo may be particularly jarring in the governance of healthcare (Khalaf, 2018). The concern is that private, for-profit entities are driven by the maximization of shareholder value and might therefore be inclined to regard the data as its own business data potentially useful for developing new products and apps. *The challenge here builds on the transparency and accountability concerns highlighted in the preceding snapshot. In this case, because of the nature of the dependency relationship between the public and private providers, attribution of responsibility for transparency and accountability is blurred.* The model presented in Part C requires that the public and private service deliverers/data sharers both be actively engaged in building safe data spaces and trusted relationships. Their role in how these are negotiated will better reveal the relationships which exist between the data-powerful players, and how this might act for or against accountability and transparency. In this way, these negotiations may expose and address any unhealthy dependency relationships.

- More generally, where the public sector is excluded from the data protection laws (or like in Malaysia and Singapore where other data control laws apply besides a PDP law), there may be less formal legal obligation on part of governments to reveal the extent of data sharing and merging. On the other hand, if the private partners have a limited incentive to reveal what they consider to be private or business data (for example for reasons linked to the willingness to protect proprietary technologies), there is concern about whether users will be able to access sufficient information to enable informed consent and decision-making on interacting with the app. *The challenge is to reach a common and acceptable (to data subjects) standard of data openness across private and public data holders.*

- But transparency and accountability are equally necessary as far as the terms of the underlying public-private partnership itself are concerned. Accepting free arrangements such as that involved in the development and initial operation of MySejahtera, without a formal contract, may lead to lack of clarity in terms of power and accountability, as private companies may want to use philanthropic projects for whitewashing, or even to bypass the procurement laws to secure government contracts. *Building on the last challenge, unhealthy dependency relationships between public and private service deliverers cannot hide behind private law market governance shields.* While the model described in Part C does not specifically argue for the accountability facility of public contracts, it does recognise the utility of other supportive governance mechanisms like public contracts known in French law (Bell & Lichère, 2022, p. 139)[85] that could give greater capacity to the negotiations between data subjects and data users.

- There seems to be an accountability gap, as it is not clear who may be holding and using the data on the platform, or even who is accountable, i.e., required to satisfactorily explain these uses to the data subjects. On the one hand, the role of private companies in the MySejahtera debacle has not been formally acknowledged by the government, making the question of possible access and any uses of citizen data on their side uncertain. On the other hand, not being bound by the data protection legislation, the government may not feel required to provide comprehensive information about the extent of data sharing in the app. Moreover, even democratic accountability could be additionally limited if the government is put in a position of a conflict of interest. *Whether a conflict of interest, or an unbalanced data dependency is the concern here, the challenge is that the conventional separation between private and public accountability possibilities is blurred in these market-driven arrangements.* This blurring is addressed in the model discussed in Part C, as it requires a new and mutualised governance project where no stakeholder can deny their responsibility to contribute to trusted data relations.

- Dependency in public service delivery may equally contribute to lack of transparency and accountability, and it may additionally limit the government's capacity to protect citizen data. If there is a clandestine dependency initially based on goodwill and subsequently turned into a commercial obligation, the app users (most of whom will at the same time be the taxpayers) will need specific parliamentary or executive protection – including where their data and its sharing is concerned. T*he challenge here stems from the development of a commercial relationship whose fundamental terms may change without reference to the users as data subjects (or even as taxpayers).* In the context of representative democratic governance, which the model presented in Part C takes as a given, for the state to expose itself to any exploitative dependency with tech providers reduces the reach of democracy as a governance mode. The model in Part C emphasizes the type of representative accountability which is compromised in such commercial arrangements.

- The lack of transparency around procurement, the conflicting, partial statements from government officials on the

[85] French law differentiates between public contract law governed by public law rules and adjudicated by the administrative courts, and private contract law governed by private law and adjudicated by the ordinary courts. Public administration can enter into both kinds of contracts. Unlike private law contracts, public contracts (contrats publics) are the ones which relate to providing a public service, and they reserve exceptional powers for the administration to protect the public interest. See Bell & Lichère (2022).

role of third parties and control over app and data, bypassing of laws, as well as potential risk of dependency in public services might have harmed public trust, which would need to be regained for citizens to have confidence in the good governance of public-private data sharing. It may be important to observe that trust is not an automatic and immediate consequence of greater transparency, particularly where it unveils bad practices which may further endanger trust. However, transparency may provide the springboard for correcting bad executive action and regaining public trust in the medium term. *The challenge here is to the short-term legitimacy of data-powerful stakeholders as they open up through transparency and accountability, revealing questionable access and use practices and undermining trust from data subjects.* The model presented in Part C does not expect that trust will come immediately through stakeholder engagement, but rather, that data openness will lead to the negotiation of mutual interests, and thereby to a revision or correction of untrustworthy data practices.

- Aside from raising concerns of post-pandemic health surveillance not unlike those discernible in SatuSehat (Snapshot A), expanding the functionality of the MySejahtera app to tracking non-pandemic infectious diseases (and possibly to other future purposes as well) involves a change of the original purpose and scope of the public-private partnership beyond the original, limited user consent to contribute data for the purposes of Covid-19 control. *This challenge to the good governance of data would have been less glaring if the app had been developed and rolled out in co-creation with civil society interests.* The model presented in Part C moves on from emergency health exigencies and requires that all parties share in the governance enterprise through the meaningful negotiation of data interests.

- Plans to build a central socio-economic database of all Malaysians are likely to present new accountability challenges, such as providing avenues of access to justice in the event of algorithmic discrimination. But they equally highlight the importance of ensuring the high quality and reliability of the data, which warrant allowing for data subject confirmation of the accuracy of the data as part of a regulatory strategy. The model presented in Part C assumes such a confirmation would be possible and beneficial to all stakeholders, as data subjects would be informed of the nature of data use as well as its direction and frequency, and better positioned to challenge what they know is inaccurate. But more generally, based on the experience with developing and rolling out MySejahtera, ensuring accountability and transparency might be easier this time, if the new initiative invites the participation and input of civil society rather repeating the pattern of a top-down, secret alliance between the state and private tech.

Snapshot C: LumiHealth (Singapore)[86]

This partnership between Singapore and Apple will enable Singaporeans to lead healthier lives, but equally important, it will contribute valuable insights to improving the health of people all over the world.
Heng Swee Keat, Deputy Prime Minister, Singapore[87]

Introduction

On 15 September 2020, the government of Singapore and Apple Inc. announced their initially 2-year long partnership to launch LumiHealth – a personalized health and lifestyle program to encourage healthy activity and behaviours among Singaporeans. The program was designed by Singapore's Health Promotion Board (HPB) in conjunction with Apple as part of the country's Smart Nation initiative, and created with help from physicians and public health experts. LumiHealth uses gamification and behavioural insights to foster the adoption of health and wellness habits among users through completing tasks such as walking or deep breathing, tracked by an Apple Watch synchronized with an iPhone. In light of the LumiHealth Privacy Policy, 'the Programme combines sensor and health and usage data to allow participants to complete certain healthy living activities and to receive smart message notifications about healthy living engagement activities' (LumiHealth, 2023a), rewarding positive behaviours with the ultimate objective to help Singaporean residents lead healthier lives. The app equally reminds users to undergo health screenings and vaccinations, motivates them to take part in challenges that aim to improve sleep/wakefulness cycles, and promotes better food choices.

Data sharing

The data flowing through the portal comes from tracking the sleep and activity of participants via the Apple Watch, as well as from activities manually recorded by the users (e.g. walks) and responses to questions and surveys meant to personalize user experience and improve the app, for which the users receive points that can later be monetized. The data collected (referred to as Programme Data) includes Personal Information (PII, such as personally identifiable fitness and wellness information), as well as Non-Personally Identifiable Information (Non-PII), i.e. data that has been de-identified or aggregate. The user is assured (through the LumiHealth website and the Program Participation Agreement, or PPA) that PII and non-PII will never be combined, and that Apple will not combine non-PII with identifiers such as the AppleID. To ensure the security of data, the Programme Data, as well as Engagement and Usage data[88] is encrypted on the user device, in transit, and on Apple servers. According to a press statement published on Apple's website, 'The collection of any personal[ly] identifiable information will be limited to information that will enable the app to provide an experience that is personalized and relevant to the user's needs,' though no more details are provided (Apple Newsroom, 2020). Programme Data can be used for a range of purposes in support of the Programme, to adjudicate rewards, to identify and deliver messages that may be helpful to the user's health and wellness or otherwise of interest to the user or other communication, as well as 'to develop and refine the Programme, [and] to understand overall Programme effectiveness and/or its impact on the healthcare system in Singapore' (LumiHealth, 2023), among others. Both Apple and the HPB assert that data from LumiHealth will not be sold or shared for marketing purposes,[89] although this does not mean it will not be shared for the purposes of building other products that a user may disagree with, and used for non-commercial purposes (Loo, 2022, p. 21). On the other hand, the user is prohibited from attempting to collect data from the app.[90]

HPB

Programme Data shared with the HPB may include: contact data, Singpass data (a Singpass account is required to use LumiHealth), demographic data, health app data (upon joining the programme, the user is invited to share their Apple Health data with LumiHealth from up to 24 past months), body weight and measurements, vitals, blood glucose levels etc, even the number of falls, and self-reported health information, app engagement and usage data (for example, frequency of using the app, responses to surveys, rewards collected). Programme Data can be 'collected, used and disclosed' for a number of purposes listed in the Privacy Policy and the PPA, both primary (linked to the main purpose of the programme) and secondary, such as 'e. to develop and refine the Programme, [and] f. to understand overall

[86] This snapshot is presented in more detail than any or the others. This should not be read as indicating that this context of governance challenge is more pressing than any of the others. It simply shows that more explanatory data was available to the research team.
[87] Apple Newsroom (2020).
[88] As explained in the PPA, this category of data shared with Apple 'may include information such as when and how frequently you use the LumiHealth App, the goals you set, whether you participated in quests and activities, your quest progress, your achievements, your responses to surveys, rewards earned and redeemed under the Programme, information about notifications you receive such as content, frequency, your responses to notifications, your user profile, your device model, software and LumiHealth App version, and App related logs' (LumiHealth, 2023c).
[89] On its website, Apple states that 'no data from LumiHealth will ever be sold or shared for marketing purposes' (Apple Newsroom, 2020), while the LumiHealth Privacy Policy states that 'Your Programme Data will never be sold or used to drive sale of any products' (LumiHealth, 2023a).
[90] 'You may not, and shall not permit any other party to: (…) (8) create a database by systematically downloading and storing LumiHealth; (9) use any robot, spider, site search/retrieval application or other manual or automatic device to retrieve, index, "scrape", "data mine" or in any way gather LumiHealth or reproduce or circumvent the navigational structure or presentation of LumiHealth without HPB's express prior written consent' (LumiHealth, 2023b, para 6 letter b).
[91] '[Singpass data collected for the LumiHealth Programme] may include information such as your Singpass ID. Your NRIC Number/FIN will not be collected' (LumiHealth, 2023c).

Programme effectiveness and/or its impact on the healthcare system in Singapore' (LumiHealth, 2023). It is not clear what categories of data are requested from SingPass via MyInfo (in order to communicate with users and send them personalized experiences).[91] The PPA makes it clear that Programme Data (though it is unclear in what form) may be transferred and stored in the US.

Apple

By default, Apple is only able to access to Engagement and Usage data, although users may opt in to sharing (de-identified) Programme Data with Apple, 'so that Apple can help to design more customised experiences and actions' such as more tailored recommendations, and improve the app. The opt-in can be withdrawn at any time, although any data already shared will continue to be stored and used as described in the applicable PPA.

Third Parties

The Programme Data can be disclosed to third parties for certain reasons, such as where the recipient has been authorized to provide services on behalf of LumiHealth. The PPA only nominates one such extraneous private company – Evidation Health (which however is not mentioned in the PP), though it does seem that data will be shared with other, unnamed third partners too. As the programme operations partner, Evidation Health is tasked with helping to manage the Programme and Programme Data – specific purposes as listed in the PPA include storage of the PPAs, coding the Programme Data, and storing a copy of the Programme Data before passing it to the HPB. The Programme Data may also be disclosed where that is required by law 'such as pursuant to a subpoena, regulatory oversight or other legal process' (LumiHealth, 2023a). Considering that the app tracks its users and entails sharing their location, users could speculate if activating the app might potentially get them into trouble with state authorities.[92] According to the PPA, LumiHealth may also use cookies and other user activity tracking technologies such as pixel tags and web beacons (LumiHealth, 2023c). While the presence of these data collecting tools does not per se imply that user data will be shared with third parties, some of those technologies are commonly used by marketers, among others.

Governance concerns

Information on data collection and sharing practices through LumiHealth is enshrined in three main sources: the Programme Participation Agreement (PPA), the Privacy Policy (PP), and on the LumiHealth website (FAQ section). The LumiHealth Privacy Policy describes the practices of the developer, i.e. the Health Promotion Board (HPB) in connection with the data collected, used and disclosed through LumiHealth, though not those of Apple (LumiHealth, 2023a). It also does not specify what is encompassed in the 'Engagement and Usage Data' category (which is shared with Apple)– this is only explained in the PPA. It should be noted that the Privacy Policy (PP), while widely available, is apparently 'not intended to and does not create any contractual or other legal rights in or on behalf of any party', and additionally, like most privacy policies, it can also be easily changed without notice. While it has been suggested that privacy policies lack the binding appeal of contracts (Norton, 2016), and in any case it is not clear whether they may produce actionable rights, this privacy policy is the main publicly available source of information on data practices and sharing by the HPB; it may raise concerns in light of the public sector's exemption from the Singapore PDPA if that information is not entirely reliable.

The PPA is, therefore, the most reliable and comprehensive source of information to list the categories of data collected and used in connection with the Programme, and on the categories of data collected and shared by the HPB as well as Apple. Still, the PPA does not tackle any potential data reuse (for example for research purposes) in any detail, and is somewhat vague on some declared purposes of sharing, such as 'to develop and refine the Programme', as well as on who are the third parties with whom the data will be shared and for what purposes. The PPA is also not widely and freely available, but only upon downloading the app on the iPhone and logging in with one's Apple ID (i.e., is conditional upon having access to specific devices and the initial sharing of some data), when the user is being asked for consent to participate in the programme.[93] Therefore, prospective participants cannot access this sharing information in advance of making the decision to join the LumiHealth programme, while the publicly available information on the sharing (and re-sharing) of data is apparently incomplete.

Likewise, it is after they join the programme that LumiHealth users receive a text message informing them that their data will be shared with Evidation Health. While (unlike the Privacy Policy) the PPA does in fact mention the Evidation Health company (once), feelings of panic have been reported upon receiving the text message by one LumiHealth user, who felt that this information should have been transmitted more clearly earlier, during the signing-up for a LumiHealth account (Kong, 2022). The availability of comprehensive and clear sharing information only after downloading the app

[91] '[Singpass data collected for the LumiHealth Programme] may include information such as your Singpass ID. Your NRIC Number/FIN will not be collected' (LumiHealth, 2023c).
[92] The experience of TraceTogether in Singapore (Elangovan & Tan, 2021) suggests that the use of citizen data outside of the declared purpose of the programme (which for LumiHealth is limited to public health considerations) may cause community disquiet.
[93] Following that, the PPA is available for reference within the app.
[94] 'HPB and/or Apple may engage other third parties to support and/or to provide services in relation to the Programme and accordingly, your Programme Data may be disclosed to such third parties for the purposes set out in this PPA' (LumiHealth, 2023c).

instead of a public and *ex ante* accessibility might hinder informed and empowered decision-making as far as joining the programme is concerned. But more broadly, without the LumiHealth users knowing who the data is shared with, for what specific purpose and based on what arrangement there can be no accountability of the data sharers for their treatment of user data.

Adding to the concern, the role of the third parties and the distribution of tasks relevant to processing citizen data is somewhat unclear in the PPA (for example, which data uses will be handled by the main stakeholders (HPB and Apple), and which will be left to third parties).[94] Specifically, users may not appreciate to which extent the programme partner, Evidation Health, might be leveraging its capabilities as a health-tech start-up active in the field of health tech research, analytics and machine learning, for the LumiHealth programme or otherwise. While the PPA makes it clear that each LumiHealth user will be separately asked to consent to any health and wellness-related research studies involving Programme Data, the engagement of Evidation to apparently handle all user data (not only in cases of such separate consent) might suggest that other forms of data could conceivably be shared for research purposes.

Separately, ensuring accountability along a chain of potentially profit-oriented, third-party private partners can involve particular challenges insofar as they are not bound by the 'no sale of data' commitment in the PPA, unless this has been separately stipulated in agreements to which they are party (and which are not public). In particular, it is not clear what is the specific arrangement with Evidation Health (i.e., what is its role and the purposes it will use the data for), especially considering that it is very active in the relevant field of health research[95] and, as Cao (2022) submits, this appears to be its main source of revenue. Specifically, it is not clear who analyses the LumiHealth Programme Data, what is the relevant sharing arrangement and whether the data is de-identified at that point (Evidation Health, for that matter, does appear to have access to PII insofar as coding, or de-identification of data is among their tasks listed in the PPA). Some LumiHealth users may be expected to feel less trust toward the initiative if they come to suspect that the role of Evidation Health might conceivably extend beyond coding and storing data as stated in the PPA. For instance, in light of Apple and Evidation Health's past collaborative research study on how personal digital devices may help in the identification of mild cognitive impairment and Alzheimer's disease dementia (Lilly & Co, 2019) users may fear that unbeknown to them they are being tracked for serious illnesses, or otherwise worry about how their information might potentially be used (e.g. to build health scoring).[96] Conversely, other users may be very glad to know whether and how they might be contributing to important studies that improve the health of people around the world, as intimated by the Deputy PM. In either case, having more clarity on the role of Evidation and other third parties in the LumiHealth programme certainly would enhance the overall transparency and accountability of the programme.

According to the Privacy Policy, 'We *acquire* Programme Data when you use the LumiHealth App and through your participation in the Programme' [emphasis added], which stands in contrast to the otherwise used verb 'collect' otherwise used by the PP (Lumihealth, 2023a). Likewise, a number of international news outlets have stated that Singapore pays its citizens to get healthy (Savov, 2020), and that citizens are offered a 'deal' whereby they get paid for wearing an Apple Watch (Sharwood 2020). Nevertheless, LumiHealth is advertised as a 'free, personalized health programme' (LumiHealth, 2023d). While at least some of the users appear to perceive the Health Promotion Board eVouchers rewards (Sharwood, 2020)[97] earned for certain amounts of points as free (EDMW, 2023), their presence may further muddy the waters in terms of how the data is shared (i.e., involve challenges to transparency and accountability). The presence of monetary rewards can arguably be seen as legitimizing the 'sale' of user data. By the same token, reward incentives ostensibly validate the view of data produced by the users' own bodies as a commercial asset rather than a key issue in personal integrity. On the other hand, the attractive 'rewards' for data in LumiHealth appear to have been a major incentive for some Singaporeans to join the programme compared to other similar programmes such as Healthy 365 or the National Steps Challenge (running on typically less expensive Android devices).[98] This preference might conceivably be explained by the prevailing cultural norms and attitudes in Singapore; in particular the so called *kiasu* or *kiasuism* (from a Hokkien word for 'fear of losing out') has been described as an 'indigenous psychological construct [which is] synonymous with the Singapore identity' (Cheng & Wee, 2023).[99] Monetary rewards encourage users to complete as many challenges as possible and therefore extend their interaction with the platform, ensuring a steady flow of data. 'Rewards for data' may equally have implications in terms of accountability for the further sharing of this data, but crucially, they shift power and control over data sharing away from the users as data subjects, making the quest of its good governance all the more difficult.

[95] Evidation Health has actually developed its own health tracking platform offering health and lifestyle insights back in 2012, integrating health-related data from third-party platforms and trackers. Through its platform (which follows a somewhat similar 'health data for rewards' concept as LumiHealth), Evidation 'connects platform users interested in tracking their health data and research organizations wanting to conduct studies on those same users' (…) and 'leverages its monetized research partnerships to incentivize user engagement and retention' (Cao, 2022). Evidation equally holds international patents in health research, which include a prediction system for subjective recovery from acute events using consumer wearables (WIPO, 2023).

[96] For example, the Swiss insurtech company Dacadoo uses data from fitness trackers to calculate individual health indexes which allow insurance companies to calculate morbidity and mortality risks, and apply differentiated pricing toward different customers (Dacadoo, 2023).

[97] HPB eVouchers are non-transferrable electronic vouchers issued in Singapore currency, which however cannot be exchanged for cash. They can be used at participating merchants in Singapore.

[98] As observed by one Reddit commentator, participation in LumiHealth allows to buy the iPhone back (and more) very quickly. See r/askSingapore (2023). See also Random (2022) and Random (2023).

[99] Kiasu has been viewed as synonymous with greed, selfishness and risk aversion, but equally with competitiveness, diligence, and a hard-working attitude (Bedford & Chua, 2018). Kiasu has been reported as relevant to the Malaysian society as well, though to a lesser degree.

Like in all other snapshots analysed in this chapter, the decision-making about data use by the LumiHealth platform and associated consent requirements follows a top-down model, exacerbated by information asymmetries which hamper informed participation. This is true not only of decisions directly associated with data governance proper (such as the collection or uses of data), but also those that involve a perhaps less direct link with data, such as the point/reward incentive system (which involves placing a value on user data). When the programme was first launched in 2020 for a period of two years, the participants were able to earn rewards worth up to S$380 over its initially two-year duration. Since 2022, the level of activity required to gain rewards has increased exponentially, while rewards have decreased at the same time; as a result, user data is now apparently 'worth' much less. This was apparently decided without taking the participants' views into account. Users have expressed disappointment online over the decreased rewards, complaining that the app tends to 'suck in' and then abandon its users.[100] Some have announced their intent to leave the programme over this change.

LumiHealth incentivizes users to share their Apple Health data with LumiHealth on the promise of help in progressing towards their health goals and earning rewards. Equally, the opportunity to customize and improve the user's LumiHealth experience by opting into sharing data with Apple, albeit vague, may be viewed as a measure of inducement to consent to broader sharing of data. Yet, the terms of this 'transaction' (personalized experiences in return for more data) may cause uncertainty among the public – a recent qualitative study reported security, privacy and lack of trust as the main factors determining the adoption of digital healthy lifestyle technologies among the Singaporean public, with concerns voiced around the sharing of personal information with various health-promoting digital tools and platforms in order to receive personalized, authentic, and meaningful experiences (Roystonn et al, 2023).[101] The underlying governance concern here is whether consent is still 'free' and 'informed'. The autonomy of data subjects may be diminished when faced with the powerful incentive of customization (personalization) to allow the maximum degree of sharing of their data. Customization is equally a strategy for the powerful stakeholders to enhance user engagement and retention, which means that the user might indirectly be agreeing to a lot more data sharing in the long run; of various kinds and with different stakeholders. But in that context it might be important to ask what customization really means in the given context (does it offer meaningfully tailored service that helps the user attain his or her objectives, or is it principally geared toward increasing the users' interaction with the platform?), and is it explainable, meaning that users are offered information on how their data is used to make it happen.

If a user decides to quit LumiHealth, their data will not be deleted, and will still be held and used for the purposes specified in the PPA and the Privacy Policy. On the other hand, a user cannot access that data with the purpose of modifying, erasing or confirming it, either. Under Singapore law, there is no legal basis to request its deletion, or that it no longer be used after a user has withdrawn from the programme. But this arrangement contributes to a diminished level of the user's 'data-power' compared to HPB and Apple, and their partners. Another issue which may impact power asymmetries is data encryption (the Programme Data is encrypted end-to-end). Encryption is certainly useful for protecting data against unauthorized access in that it obscures PII and thus protects privacy. However, the important governance question when determining duties and obligations seems to be: who holds the encryption keys, and is therefore uniquely able to decrypt the data? Although it can be assumed that the tech companies are involved, there is no information about this. One thing seems clear: the encryption keys are certainly not handed to the users.

Although the Privacy Policy HPB states that 'Your Programme Data will never be sold or used to drive sale of any products', participation in the programme may in turn incentivize users to purchase Apple products, typically more expensive than other brand equivalents. An Apple Watch paired with an iPhone (to enable data sharing) is necessary to participate in the programme, as is downloading the LumiHealth app from the Apple Store. The LumiHealth app does not work on other trackers, or on Android-using phones, which limits the general public's access to this particular pathway into a healthier lifestyle enabled by technology. Moreover, the programme may be indirectly encouraging Singaporeans to purchase Apple products, potentially leading to dependencies on the brand if they learn to rely on the app in the pursuit of that lifestyle, and choose to remain on the platform to monitor their level of activity. LumiHealth programme participants may subsequently find themselves incentivized to purchase more premium-priced products from the Apple family (Apple devices are notorious for their lack of interoperability with non-Apple products, which is a way for Apple to ensure that it alone monetizes the data it collects from users (Hern, 2021)). This further assists Apple in the quest for a continuous supply of more kinds of data, as well as further revenues. Separately, through the Apple/HPB collaboration (which could be seen as restricting competition in that the particular service is not available outside of the hardware manufactured and sold by the preferred provider), Apple and its products become entwined in a leading public policy programme (Engels, 2020), with commercial apps such as LumiHealth app featuring in the Ministry of

[100] See r/Lumihealth (2022) and EDMW (2023).
[101] As an example, in the words of a male interviewee in his early 20s, 'I'm not very comfortable with the idea of sharing such personal data with technology companies….I think that's really very scary to me. Yeah, taking over our lives. And what we can or cannot do. So yeah….because they really do steal data from their own (users)' (Roystonn et al, 2023).

Health's strategy to roll out personalized health plans for Singaporeans (Teo, 2023).[102] To an extent, the same criticism could be raised about any of these snapshots where only one access pathway is offered through private technology (hardware or software) to benefit data subjects in the provision of public services.

Additionally, endorsement of commercial trackers for the quest of improving public health and services may effectively legitimize the use of technology and approaches whose reliability is not scientifically proven (Zhang, Godin & Owens, 2018; Kuosmanen, Visuri, Risto & Hosio, 2022; Cox, 2018) and the tech-solutionist narrative more broadly, as well as normalizing constant tracking, self-quantification and comparison associated with a 'cult of numbers' (Mau, 2019). All of this raises questions of accountability for shaping lifestyles and identities in a certain key, in addition to stakeholder dominance and power dispersal cutting across public and private spheres.

Challenges

- Consistently with all the other snapshots analysed, the enumerations of categories of data accessed by the app described in the Programme Participation Agreement (PPA) and the Privacy Policy appear to be partial and exemplary rather than specific and exhaustive. This may raise concerns among users of the app, considering that they will have no formal expectation of minimization in data collection (this principle is not enshrined in the PDPA), and in any case, the PDPA would only be binding on the private stakeholders involved in the LumiHealth programme (if at all).[103] If the private provider is acting for the government or with some delegated authority, confusion could arise as to whether the PDPA would apply, considering its specifically private sector focus. This kind of data sharing situation generates governance uncertainty from the perspective of conditional personal data protection regulation. More specific information provided to users on the data collected and shared could help engender more trust among users and ensure informed participation of users. *The challenge here stems from inadequate transparency, but is exacerbated by ambiguous language and limitations on the conditions offered.* The model in Part C would require that the private provider be more candid and less conditional at the negotiation table. As the model presented in Part C will be activated in specific contexts of data use, stakeholders might additionally require that the language used in these accountability exercises be open to criticism by data subjects along with their requirements for more context-specific clarity.

- Consistent with most snapshots analysed in the Report, the information relating to the sharing of data on the platform is somewhat limited. While admittedly containing much information about the LumiHealth programme, neither the LumiHealth the Programme Participation Agreement (PPA) nor the Privacy Policy make it entirely clear who are all the third parties with whom the data is shared, what is their overall role in the programme, as well as what are the conditions and purposes for the sharing of user data. Users interested in understanding how their data may be shared might be interested in clarifications of the role of third parties such as Evidation Health (only mentioned once in the PPA), and of some of the generally described purposes of sharing, such as 'to develop and refine the programme'. More in particular, users may wish to know who is the stakeholder analysing Programme Data for insights, whether the data is de-identified at that point, how data might or might not be reused, e.g. for research, and, relevant to encryption as a mechanism for ensuring the security of data, who holds the encryption keys. *The challenge of information deficit is one of the most universal across these snapshots. What makes this instance particularly challenging is the extent of data immersion and the sensitive nature of health data.* Part C describes a contextually specific, bottom-up governance model in which the particularity of data and addressing information deficits through more clarity are significant considerations in the process of creating safe data spaces and trusted data relationships.

- The PPA – the most comprehensive and reliable[104] source of information on data collection and sharing – is relatively inaccessible in that it requires prospective users to possess specific devices, download applications[105] and share authentication data. Advance public availability of complete information on data collection and sharing could help data subjects in making informed decisions about joining the platform and facilitate forming trusting relationships from the get-go. *This challenge highlights that information deficit is not cured by more information alone.* The Part C model will not be satisfied if the user of the platform is flooded with information which is not accessible because of where it is located and how it is expressed.

[102] The initial versions of the personalised health plans under Singapore's preventive care strategy are likely to be more generic, Mr Ong [the Minister of Health] said. But over time, this will change as individuals use apps such as Healthy 365 or LumiHealth to track their health goals and nudge them to stay on top of their plans. "When you see the doctor again, the doctor will look at what the health app says and further personalise your health plan. So there will be an iterative process"' (Teo, 2023).
[103] Singapore Participant, research roundtables.
[104] In light of the Privacy Policy of LumiHealth, it is 'not intended to and does not create any contractual or other legal rights in or on behalf of any party' (LumiHealth, 2023a). While privacy policies such as this one might not necessarily produce legally enforceable rights anyway, this clause might have repercussions in terms of the perceived reliability of the PP, and user trust more generally.
[105] As the Federal Court of Australia stated in the case against Meta, '[t]he ACCC submitted that it is significant that users of Onavo Protect were only asked to accept the Terms of Service having already been induced to download the app. I accept those submissions.' Australian Competition and Consumer Commission v Meta Platforms Inc [2023] FCA 842 (26 July 2023), para 35.

- Observing the significance of customising data in this snapshot and the impact it has on consent, data production, use and sharing, an additional governance challenge is presented. It would not be sufficient or appropriate simply to prohibit or deter customisation because many users choose it for perceived commercial value, perhaps especially so in societies with a culturally embedded preference for 'getting the best deal', or the 'fear of losing out' such as *kiasu* in Singapore. Instead, it may be enough to require that users are fully informed about the why and how of customization, and what are the competing benefits and potential downsides in this practice (for example, that they are likely to spend more time on the platform, which also implies sharing more data) so that their choice participate is informed. The abovementioned information deficits might impact accountability understood as providing a sufficient explanation for the use of data, and on whether consent can be thought of as 'informed', considering that partial disclosure can only lead to a partially informed consent).

- *Recognising the complexity of information deficit as a challenge that appears in various degrees and forms throughout all the snapshots, a common challenge concern relates to the willingness of powerful data holders/users to tell the whole story, and tell it in a way that genuinely empowers data subjects.* The model presented in Part C invites more powerful stakeholders to foster user trust and data empowerment through specific use concessions such as enabling users to opt out from user data retention upon leaving the platform. Particularly if there is no requirement to delete user data in personal data protection legislation, data holders/users going further in data concessions than other governance strategies require will enhance trust and overall transparency about the data practices, particularly when the data subject no longer wants to participate in the purposes for the original data collection.

Snapshot D: Pair Chat (Singapore)

Introduction

The collaboration between Microsoft and the Singapore Government for the use of Pair Chat, an AI-powered writing assistant for government officers, was first announced in early 2023 in the Straits Times newspaper (Chia, 2023a). The technology, which is essentially customized ChatGPT technology embedded in Microsoft Office products, was launched at the annual Open Government Products hackathon (Hack for Public Good, 2023) with the promise of helping to free up civil servants to pursue more complex working tasks by identifying information, summarising long chunks of text and preparing complete drafts for government staff within seconds. The generative AI tool has already been rolled out in several departments including the Smart Nation and Digital Government Office. The plan is to eventually have around 150,000 government staff use the chatbot across all agencies (Min, 2023). This comes at a time when governments around the world are increasingly exploring the potential of generative AI technology for a variety of tasks; the UK government might soon follow in Singapore's footsteps by using similar technology with the aim of increasing productivity in the public sector (Strauss & Foster, 2023).

Although Pair Chat was made possible thanks to a special agreement between the Government and Microsoft's Azure OpenAI Service, ChatGPT is based on a Large Language Model (LLM) originally developed by the tech company OpenAI, in which Microsoft holds a large stake. The GPT technology, which uses natural language processing algorithms to respond to text-based human inputs broken down into smaller pieces (tokens), raises concerns about privacy and the proper use of personal information, among others. The model comes pre-trained on huge amounts of information, scraped indiscriminately (and not always legally) off the Internet, including names and surnames, email addresses, phone numbers, addresses, medical records etc. The data may contain information which allows to identify an individual, and is therefore governed by data protection laws to protect the autonomy of the data subject. As an immediate concern, if no privacy protections are applied, that data could appear in the outputs produced by the LLM either intentionally, through so-called inversion attacks or divergence attacks leading to the system 'regurgitating' chunks of memorized training data (Nasr et al, 2023), or unintentionally, for example when the model is used as a search engine (Borji 2023, p. 28; Van Daalen, 2023). More broadly, there is a general uncertainty about ChatGPT's third-party data practices, which are not exhaustively addressed in the OpenAI privacy policy (Gal, 2023). Some recent LLMs have adopted the practice of filtering out personally identifiable information (Zhuo, Huang, Chen & Xing, 2023). However, regarding the PII-removing technology used by the commercially available ChatGPT, Emily Bender, professor of computational linguistics at the University of Washington, was quoted as saying that it would be almost impossible for OpenAI to identify and remove all the personal information from the data provided to ChatGPT (Kim, 2023).

Data sharing

Unlike the other snapshots analysed in the Report which typically involve three kinds of relationships (public sector-private sector; public sector-the citizens; and private sector-the citizens), this partnership does not involve a relationship between the private sector provider (Microsoft) and the public.[106] The main relationship here is the contractual relationship between Microsoft and the Singapore government. While it will not involve the provision of services directly to citizens, the rollout of Pair nevertheless raises governance questions inasmuch as

1. The technology requires data contributed by the public (including Singaporean citizens) to learn and to function;
2. it may therefore involve data sharing of citizens (including the data of civil servants) and in so doing, it may have an impact on their lives, and finally
3. its use may create dependencies which shift the balance of power in ways that are new and not too well understood for now, but may be relevant to data-driven technologies which will be used to directly provide public services to citizens.

GPT-3.5, which is the digital backbone of Pair, has been pre-trained on a large dataset scraped from the internet. Therefore, the training in many instances might have been unlawful for using mass quantities of personal data without consent (Garante per la Protezione dei Dati Personali, 2023),[107] and of copyrighted material without a license (Associated Press, 2023). While OpenAI claims that the training data for GPT-4 came from 'a variety of licensed, created, and publicly available data sources, which may include publicly available personal information' (OpenAI, 2023a, p. 53), the director and co-founder of the company did not elaborate when asked for more details (Wiggers, 2023). It should be noted that some of that data would have been placed on the internet with no valid legal basis whatsoever such as through doxing, which is illegal in many jurisdictions, including Singapore. The legality of ingesting and training on much of that data is therefore highly controversial, and is presently the subject of a number of investigations and lawsuits worldwide.

[106] In saying this it is intended to see the public as an active third-party stakeholder. It could be said, on the other hand, that the state and executive government are for the people, and thus accountability to the public in the exercise of executive services is implicit.
[107] Following an inquest, Italian DPA has found no valid legal basis (no consent) for the collection and storing of mass amounts of personal data by ChatGPT for the purpose of 'training' the algorithms underlying the operation of the platform, and i.a. ascertained that the users whose data is processed in connection with their use of ChatGPT do not receive any notification of that processing, which is essential for the consent to be informed. As a result, access to ChatGPT was temporarily suspended from the territory of Italy.

ChatGPT has been found to be prone to training data extraction 'attacks' to recover individual training examples (Carlini et al, 2021). This training data may thus on rare occasions be 'leaked', or accidentally shared with any number of users upon entering the right query to the chatbot, and exposed to further non-consensual uses. In such cases, it would be impossible to monitor who has had access to that data. In this setting the issue of risk is not measured only by the probability of an occurrence, but by the harmful consequences if it does occur.

Civil servants in departments working with Pair Chat may use the technology as a productivity work tool for a wide spectrum of purposes such as research, preparation of policy papers, contracts, regulations, leadership speeches, parliamentary queries, briefs and internal updates on news, as well as emails to citizens (OGP, 2023, at 9:20-12:40). Considering that public administrations maintain and manage large databases of sensitive information, Pair has been cleared for Restricted/Sensitive Normal information (Developer Portal, 2023). As stated by the Minister-in-charge of GovTech in Parliament, work that contains highly confidential or sensitive information will still be written exclusively by civil servants, without the use of Pair (Smart Nation, 2023). Features yet to be released in the future will allow Pair to access and analyse information from official databases, and for civil servants to upload documents directly to the writing assistant.

In principle, the above-mentioned tasks could still involve the sharing of sensitive data of the user or of third parties in civil servants' interactions with the chatbot, causing the risk of leaks or of further data sharing for unknown purposes (such as training new models), as is by now well-documented with regard to the ChatGPT technology (ART GSC, 2023; Eliot, 2023). It is for this reason that organizations around the world (including Apple, Amazon, and Samsung, and even Microsoft itself) have been strictly forbidding their employees to input any sensitive information into the chatbot, or even banning the technology completely for internal company use for fear of sharing sensitive data such as client information (Kim, 2023; Mok, 2023; Gurman, 2023). Importantly however, Pair Chat possesses improved data confidentiality features compared to the commercially available ChatGPT offered by OpenAI. Pursuant to a special agreement, Microsoft has agreed to turn off logging in Pair, which is taken to mean that all prompts (which may include personal data) will be kept out of sight of Microsoft and OpenAI.

Yet, even if conversation history is not collected, other types of data linked to interactions of the civil servants with Pair might be gleaned. In particular, metadata for every interaction is usually collected and can likely be used by OpenAI or Microsoft for purposes which include improving the capabilities of the model, troubleshooting and tracking performance. Although metadata does not include personal data as such, it could still provide information about how Pair is used in government, such as the number of tokens used (which may imply the length and complexity of the inputs and outputs); reason for the conversation finishing (e.g. because the reply was completed, or because content was flagged by a content filter), response times or timestamps (OpenAI, 2023b). Such data might still conceivably be used for statistical analysis and predictions, linked e.g. to the required token length to help OpenAI gauge how much traffic to handle for a specific group of users. On the other hand, it might provide a glimpse into the volume and frequency of use of Pair in government, the length and complexity of conversations with public servants, task variety, or even the work patterns and rhythms in that group of users. Somewhat similarly, government agencies will likely be able to have oversight of staff's interactions with the chatbot through logging the information locally, allowing for employee evaluations. The issue of who should be accounted to here, along with the independence and autonomy of the user are important, and yet somewhat clouded. The civil service is a hierarchical structure in which one's status will determine their opportunity and capacity for independent choice and decision-making. Such factors not only have an impact on the nature and occurrence of the choice to share data, but more simply on who is the data subject, and who should control the data that she generates. In this context, individual staff are encouraged to interact with the platform, and work with its outputs. The data that comes out of such exchanges is in large part personal, and as such has consequences for the data subject if it can be attributed and reviewed by others.

Governance concerns

Publicly available information about Pair and the extent to which it involves data sharing is somewhat limited, with most of the information coming from several articles published in the Straits Times newspaper, and from the content of the Pair website. But more broadly, the limited availability of information about the way data is used in Pair has to do with the fundamental opacity of the proprietary technology involved (Walsh, 2022). This is a classic black box problem, and the newly released GPT-4 (to which is where Pair is likely to be upgraded in the future) is even less transparent than GPT-3 in this regard (Barr, 2023). Another reason for the information deficit is the secrecy of contracts; specific privacy arrangements such as a privacy policy will only be known to parties to the contract.[108] Additionally, there is a risk that like so many others, these policies might not even be very clear on the use of data. In particular, Microsoft's European subsidiary is currently under investigation from the Hungarian Competition Authority over the allegedly contradictory and unclear character of its privacy policy in terms of providing information on what happens to personal data of users during the use of another generative AI product similar to ChatGPT (GVH, 2023). Separately, the Mozilla foundation has recently launched a public petition to Microsoft to clarify whether or not it is repurposing the data of users of its 130

[108] Standing in private law has always been an impediment to the involvement and interests of the wider public when public/private partnerships are governed by private contract arrangements.

products to train generative AI models – something that nine privacy experts have not been able to get clarity on from the newest Microsoft Services Agreement (in effect from 30 September 2023) (Mozilla, 2023).

The use of Pair in government may involve challenges linked with accountability, control, and the power asymmetry between the stakeholders involved in the partnership. For now, accountability of the ChatGPT technology remains unregulated. From a technical point of view, it seems impossible to carry out a thorough audit of the technology due to lack of access to its proprietary algorithms and training data. Additionally, as there is virtually no way of determining how a particular output was generated, and how and why particular data was used (ART GSC, 2023), it might be exceptionally difficult to attach responsibility in the event of the chatbot producing an output containing personal data. Any misuse of data using the technology is therefore impossible to prove in court, potentially leaving data subjects without a private law remedy. On the other hand, while it has been stated that Microsoft does not log the data, there does not seem to be any way of empirically verifying its data practices on the cloud, and therefore of keeping it accountable through data openness practices. It might appear as though there is little else to do than have trust in the unilateral assertions of contractual compliance.

As Pair offers an extra layer of security meant to ensure the confidentiality of sensitive data handled by the Government, public officers will most likely be inputting such data into Pair Chat, trusting that these will not be logged. At the same time, as observed by the computer science professor and AI expert Vincent Conitzer, Microsoft may benefit from collection of data which could then be used to improve the ChatGPT technology – and though the company is unlikely to behave irresponsibly, 'it's always good to be aware of incentives' (Kim, 2023). While there is indeed no reason to assume that contrary to contractual commitments, Microsoft will log data from interactions with Pair, the company might be said to be facing a conflict of interest here that is similar to the one which led it to prohibit its employees from inputting sensitive data into the commercially available ChatGPT offered by OpenAI. The perils of entrusting the confidentiality of valuable data to a company which profits from the accumulation of data might be illustrated by Meta's 'Onavo Protect' app which promised to protect the data of users when browsing online content through its VPN service, but was later revealed to be pulling in user data for its own purposes (Constine, 2019).

In terms of power asymmetry risks, technological dependency (IEEE, 2016) in the public sector is a potential concern, both in terms of reliance on foreign private infrastructure, and of staff consistently outsourcing writing skills to generative AI. In light of a presentation by a member of the team developing Pair at the OGP hackathon, the use of Pair Chat is expected to accelerate the knowledge work done in the government by taking over some of the 'writing' tasks to free the officers up to do 'thinking' tasks (OGP, 2023, at 9:20-12:40). This may well be so; however, research suggests that working on writing actually helps cognitive skills, and in that sense, doing one's own writing might be very useful for intellectual work and analysis (see e.g. Menary, 2007). These concerns are taken a step further by Ang, Ho & Jayakumar (2023), who warn that 'the tool, if not judiciously used, may well end up shaping the public service', arguing that public servants must understand all the 'building blocks' of higher-level tasks, and equally develop research and analysis skills thanks to knowledge passed down from more senior officers. There is also concern that too strong a reliance on Pair for research might involve its own risks – insofar as it is trained to predict the most likely order of words in a sequence but without any capacity to engage with their meaning, ChatGPT has been known to hallucinate, or fabricate answers (Narayanan & Kapoor, 2022), going so far as to inventing inexistent academic references to back up its own statements. To address such risks, the guidelines by the Ministry of Communications and Information on AI usage in the civil service on the safe and effective use of Large Language Models (LLMs) in civil service remind officers that they are responsible for checking any content generated by these models (Chia, 2023b). But even more caution might be required.[109] Scholars have raised concerns that such technologies may tie users to themselves either through what is known as automation bias, i.e. when users place too much faith and trust in automated systems, overestimating their performance and accuracy (Jones-Jang & Park 2023) or, alternatively, by becoming increasingly indispensable in their lives, so much so that relationships of trust are replaced by ones of necessity or convenience (Findlay 2023, p. 6).

Challenges[110]

- The use of generative AI such as ChatGPT (on which Pair is built) involves challenges to transparency and accountability; both key principles in public administration. This is owed to its unregulated character as well as the proprietary, highly complex and changeable code, which makes it impossible to understand how these technologies work and how they learn. Generally speaking, transparency is not helped by the lack of information openness from the developer of the ChatGPT technology, OpenAI, whose technical report on GPT-4 released

[109] In his book, *Machines of Loving Grace*, John Markoff, former tech reporter for the NYTimes, sets out to answer the question of whether AI will help us or replace us. In his view, outsourcing tasks to AI may actually cause one's autonomy and control to diminish. According to Markoff, 'For today's younger generation, the world has been turned upside down…Rather than using computers to free them up to think big thoughts, develop close relationships, and exercise their individuality and creativity and freedom, young people were suddenly so starved for direction that they were willing to give up that responsibility to an artificial intelligence in the cloud. What started out as Internet technologies that made it possible for individuals to share preferences efficiently has rapidly transformed into a growing array of algorithms that increasingly dictate those preferences' (Markoff, 2015, pp. 341-342).

[110] It will be apparent by now that many of the challenges that are identified for governance attention are interrelated and even recurrent across snapshot contexts. This is to be expected when central concerns like information deficit are at the heart of much data sharing practices linked to the use of data-driven technology, including in all smart city administrative arrangements.

in early 2023 justifies its lack of information openness with the need to preserve competitive advantage and by reasons of safety (OpenAI, 2023a).[111] But more broadly, it is nearly impossible to fully understand the inner workings of the technology, and by extension, how it may have used specific data points in a given instance. On the other hand, the information asymmetry with regard to Pair may be exacerbated by the collection of metadata which might offer a glimpse into the work rhythms and patterns linked to policy formulation in the government to a private company. *The challenges to transparency and accountability are more complicated by the source of information that is not accessible or easily appreciable.* No matter whether the information at the heart of the deficit is produced by human sources or comes from AI-assisted technology and applications, the model described in Part C requires its availability to data subjects as part of the empowerment objective. Even if the information comes via complex algorithms, it is also attendant on these data users to make clear how the data is processed and what conclusions are drawn from it. Even though this might be a technological question, the requirement for open and accessible data information remains.

- The use of Pair requires broad-based trust in the largely intransparent, and by now controversial technology currently facing multiple lawsuits and investigations. One area which calls for trust is the data confidentiality arrangement. This trust might not be automatically given considering, on the one hand, the opacity of the technology and the sheer impossibility of steering it, and on the other hand, potential discrepancy between the public and private sector stakeholders in terms of values and incentives. For example, private tech providers may have little incentive to comprehensively share information about their technologies (especially up to a democratic standard that rests on the requirements for inclusion and representativeness), and yet all other smart city stakeholders must rely on their assurances concerning realities they themselves cannot access or experience. *The challenge here is for trust to be generated and maintained in the long run through a shared commitment to address the fundamental information asymmetry as far as possible given the arguably inscrutable nature of the technology*. The model presented in Part C considers that the quest for accountability (understood as capacity to explain how data is used and how reliable are the systems and processes used for its protection) depends on data openness, whether it is tech-processed or otherwise.

- *A challenge that results from dependencies on generative AI as well as the complexities of understanding where answers come from and how they were derived is essentially another example of power asymmetries in data arrangements.* The model presented in Part C holds that power will strengthen trust in a situation of a potential conflict of interest, such as this snapshot seems to reveal. While data-subject participation could conceivably lend more legitimacy to the use of the controversial technology in public service delivery, it would need to do so from a base of informed engagement rather than dependency. The key governance objective of power dispersal does not call for a major restructure of data market relations. More simply, providing more information to data subjects about the nature and extent of data sharing in which they have an interest, could constitute the foundation for negotiating opportunities where data subjects can participate in the control of such sharing.

- *Another challenge is how to incorporate technology into the transparency mission*. To maximize transparency and accountability, technology could curate and document the training datasets and involve stakeholders early on (Bender, McMillan-Major & Mitchell, 2021, p. 615) in a sense of co-production so that it no longer is a limitless pool of lawfully and unlawfully procured data mixed together. The recent recommendation by French, Italian and German governance proposals around mandatory self-regulation suggest the inclusion of model cards in generative AI applications to ensure better transparency throughout the operation of the application (Italy, France and Germany, 2023). Greater transparency around the technology would better fit the non-profit, representative and inclusive, public welfare-oriented public sphere, and its commitment to accountability toward the public.[112] At least where these technologies are used in the public sector (and thus may impact citizens), and to support the openness mission of the model in Part C, it should be possible and indeed necessary to strive for a greater degree of transparency through agreements between stakeholders that are designed to benefit their individual and shared interests. Such agreements could take the form of either 'public' contracts (McLean, 2019) or even contracts which admit the interested public as a third party with contractual standing (Perez, 2002).

[111] 'Given both the competitive landscape and the safety implications of large-scale models like GPT-4, this report contains no further details about the architecture (including model size), hardware, training compute, dataset construction, training method, or similar' (OpenAI, 2023a, p. 2).
[112] The concern is that endorsement of such technology by powerful social actors without sufficient transparency and data subject participation avenues will only further consolidate the power of tech companies to the disadvantage of all other players, who will be locked into cycles of dependency by the promise of a 'smarter' work, education and improved life experiences.

Snapshot E: Ruangguru (Indonesia)

Introduction

Ruangguru is an online tutoring platform built by one of the largest edu tech start-ups in SE Asia. According to its website, 'Ruangguru develops various technology-based learning services, including virtual class services, online exam platforms, subscription learning videos, private tutoring marketplaces, and other educational content that can be accessed via the Ruangguru web and application' (Ruangguru, 2023a). With a broad scope of online courses ranging from the K-12 education to employee training, the technology has played a role in expanding access to education in Indonesia – according to one of the platform founders, only 30% of its users come from big cities (East Ventures, 2023). The platform played an especially significant role during the Covid-19 pandemic through its Free Online School which offered online distance learning classes (taught live) every Monday to Friday (Kemendikbud, 2020), which also led to significant growth of the user base. As of November 2023, the app has had over 10m app downloads, and has 22m users as reported by Ruangguru itself (Ruangguru, 2023a). The platform follows a 'freemium' (two-tier) model: the basic (free) tier is mainly meant to attract users, while the paid (premium) brings profits.

Data sharing

The app appears to collect large quantities of data including personal, academic, technical and app usage information – some of it input by the users themselves, and some of it coming from tracking users to create a record of everything they have done in the system, down to every mouse movement – and it shares this data with third parties which are unnamed in the privacy policy (Ruangguru, 2023b). Pursuant to the investigation of the app by the Human Rights Watch as part of its 'Students Not Products' initiative, these 'third parties' very likely include Facebook (currently Meta) and in some instances, may include Google (HRW, 2021a), among others. The app is largely addressed to children – while users under 18 must express consent to join the programme through a parent or guardian, data on them will equally be collected and shared as per the policy as they use the app.

The privacy policy makes it clear that the platform shares user data for advertising purposes (Ruangguru, 2023b, paras 3 c.v and vi; para 4.b.2). This means that it likely makes it accessible to a potentially wide network of advertisers which can then target users with ads, like other similar apps have been revealed to do (the report by the Norwegian Consumer Council refers to 'thousands of interconnected entities', typically without any direct relationship with users (Forbrukerrådet, 2020, p. 14). Additionally, the app has the capability to collect and share with third parties data such as user location, among others (HRW, 2021a). Tracking location data may lead to discovering a lot more information about the user – not only where they live, study and work, but also potentially personal characteristics of the user such as who they spend time with, as well as what is their religious and political affiliation and sexual orientation, among others (Forbrukerrådet, 2020). The Ruangguru app can also share data with government for purposes linked to law or policy.

Governance concerns

The Ruangguru privacy policy is part of the T&C, and both are available online (in the Bahasa language). The privacy policy is quite long (the equivalent of 10 pages in a single-spaced Word document), and somewhat fragmented.[113] This fragmentation exacerbates the relative lack of clarity of the policy – the document does not explain some key issues in a language that is accessible to all,[114] such as whether all user data shared or disclosed to third parties will be anonymized or not, and whether Ruangguru users will be asked to separately consent should the policy change, or if any changes are thought to be covered by the initial consent (Ruangguru, 2023b, cf. paras 12 and 3(f)). The purposes of sharing are also quite broad and open-ended: for example, Ruangguru can disclose users' personal data to 'members of [its] business group, which includes branches and subsidiaries, as well as the main holding company and its subsidiaries ("Affiliates"), in order to, including but not limited to (i) providing support for the provision of Applications by Affiliates, (ii) implementation of [its] business activities by Affiliates, and (iii) data processing by Affiliates' (Ruangguru, 2023b, para 4.a).

Further, it is not too clear either what kinds of personal data are actually collected by the app. The privacy policy provides a very broad, open-ended definition of data (which includes e.g. genetic data, criminal record, phone numbers of all contacts as well as, very generally, 'data and other information that we legally obtain either from you directly or from other third parties'), without saying plainly which categories of data is actually collected by the platform in question. It is also not clear what third parties that data may come from, let alone be shared with by Ruangguru especially when it comes to data collected directly by platform, i.e. user tracking data. Confusingly, developer information (which is

[113] In a case involving a consumer law complaint, the Federal Court of Australia agreed with the Australian Competition and Consumer Commission that 10 pages of a privacy policy of an app used by consumers might be too long. Australian Competition and Consumer Commission v Meta Platforms Inc [2023] FCA 842 (26 July 2023), para 35.

[114] A reminder of different levels and capacity for comprehension among users who access these apps in very diverse social settings was offered in our conversation sessions with regional experts. When looking at clarity and accessibility in privacy policies it is first necessary to ensure a simplicity in common languages and experiences, supported by logical and approachable structuring of the important issues that users must first understand to be deemed to consent to.

altogether easier to understand than the privacy policy given its format, and easily available to access on the Google Play platform wherefrom Ruangguru can be downloaded), contradicts the privacy policy by saying, for instance, that no data is shared with third parties (Google Play, 2023c). Assuming that good data use practice recognizes the need to provide informed user participation, mixed messaging like this is a fundamental challenge for governance.

The Ruangguru privacy policy makes a reference to accessing the data of third parties (contacts of the user) through the 'Contacts' app on user devices (i.e., their phone book), ostensibly to identify connections between existing users for them to interact on the app's social features (Ruangguru, 2023b, para 3.e). It is not clear whether the app will ask for additional user consent before searching the contact lists on user devices, or whether this will be covered by the broad initial consent sought when joining the app. In any case, that initial consent is not for that user to give – it seems to be a wholly unreasonable expectation that the user will ask every one of their contacts (including people he or she might no longer be in contact with) for consent for such sharing. While the Privacy policy says this information will only be accessed once and 'not stored' by the platform (ibid.),[115] Ruangguru would certainly have a commercial incentive to put such data to use. Address book data are valuable sources of information about users, other people that they are associated with and connections they have with them, and usually in high demand from both public and private players, as revealed in high-profile data abuse scandals (Gellman & Soltani, 2013). The many platforms that try to access that data on their users' devices tend not to be too transparent about the purposes they collect that data for, and who they share it with (Kelly, 2021a), and some of them were revealed to be doing it in spite of users' refusal to consent (FTC, 2013). Facebook has infamously claimed back in 2019 that collection of data of millions of people through a similar feature had apparently happened 'accidentally' and promised to delete the database (Hern, 2019), although both these claims are notoriously hard to verify. But even if the data of third persons is not monetized and only used to maximize user engagement with the app, the potential scale of such sharing (necessarily involving not only the sharing of data of existing Ruangguru users, but of all contacts stored on the given user device) and its likely unlawful character are a source of concern.

By its own admission, Ruangguru shares the data of users for marketing purposes (Ruangguru, 2023b, para 3.c.v and vi; para 4.b.ii), which means that is it selling it. While in light of the privacy policy, this is supposedly only meant to include aggregate information rather than complete personal data (Ruangguru, 2023b, para 4.b.ii), the Human Rights Watch investigation has found this information to be misleading. It discovered that the app was actually sharing users' unique Android Advertising IDs with third parties (HRW, 2022, p. 61). In any case, aggregate information can still include some personal details and be sufficient for users to be segmented and targeted by advertisers based on data collected by Ruangguru. Information collected by marketers and put together at the back end can lead to user profiling and building multiple and ever-refined digital twins of individual users which may ultimately allow for predicting their behaviour as customers (to be sold on to the highest bidder for a potentially infinite number of times). Data reuse in this setting will also be driving tactics such as differentiated pricing (e.g. of health insurance) for clients who are subjectively seen as wealthier or more motivated to make a purchase, on top of increased exposure to hacking and fraud, ID theft and other digital crimes. Of course, at that point initial users or even Ruangguru have no control whatsoever over who has the users' data and what uses they are putting it to. In fact, according to the Human Rights Watch, Ruangguru itself is likely to be profiling its users, as it admits to supplementing its data on users with data on the same users coming from other (unknown) sources (HRW, 2022, p. 62). All of this is said to be covered by the same one-time, bundled consent, although this is clearly an additional use of data vis-à-vis the main purpose of the app and thus could easily warrant a separate consent request.

As per the Ruangguru policy, it is possible to request deletion of one's data or withdrawal of consent. However, the policy seems to suggest that at that point data will be anonymized and only deleted after 5 (sic) years, provided that Ruangguru no longer needs it at this point (Ruangguru, 2023b, para 11),[116] even though Article 16(2)(g) of PDPL appears not to foresee any retention period in case of user data deletion requests.[117] Apart from the fact that anonymization does not always guarantee data privacy (AEPD & EDPS, 2021), it seems that users might still be targeted with advertising based on the data held by Ruangguru during that 5-year period. In any case, there is no way of tracing that user data which has already been shared with other entities by Ruangguru, and by extension, not possible to request its deletion.

Ruangguru has a record in collaborating with the public sector in providing educational services. According to its website, Ruangguru has partnered with most provincial governments as well as many city and regional governments in Indonesia. During the pandemic, Ruangguru collaborated with the central government on its pre-employment card programme.[118] The platform had equally benefitted from 'zero-rating' by the government-owned Telkomsel

[115] 'If you use social features on our application, you can share our users' telephone numbers stored in the contact list application or address book on your device so that we can help make it easier for you to connect you with your friends or acquaintances who use these social features. We only use the data contained in your contact list application or address book for 1 (one) contact synchronization process, and we do not store it on our server and/or database' (Ruangguru, 2023b, para 3.e)
[116] '[I]f the Personal Data is no longer needed for the services or purposes stated in this Privacy Policy, and if we no longer have legal or business purposes to store your Personal Data' (Ruangguru, 2023b, para 11).
[117] As per the automatic translation from Bahasa Indonesia. To the best of the authors' knowledge, an English translation of the law is not available yet.
[118] A social protection and competency development program for job seekers and laid-off workers in Indonesia, involving job skills training offered through digital platforms. The program was introduced in 2020 as part of a pandemic-time economic stimulus package.

wireless network provider company in an arrangement that offered Indonesians 30 days of quota-free access only to the Ruangguru app (up to 30 GB) to support distance learning policies during the Covid-19 pandemic (News Desk, 2020). Judging from the kind of incentives for telcos involved in Facebook's Free Basics (a similar free connectivity programme), this arrangement likely was beneficial to Telkomsel by giving it access to users which might someday buy full access to the internet (Pisa & Polcari, 2019). On the other hand, Ruangguru no doubt stood to gain by accessing prospective clients which might continue with the platform later. But perhaps more importantly from the perspective of governance challenges, through this partnership to deliver apparently 'free' educational services to the public both Ruangguru and Telkomsel got access to large amounts of user data, which were likely monetized by both, as Telkomsel has equally been optimizing big data services to carry out advance profiling of the target market to sell to advertisers (Silalahi, 2017).

As an additional challenge for good data governance in public-private digital service delivery in the Indonesian context, collaborations between the public sector and start-ups may be facing some public distrust linked with 'revolving doors' between the two sectors. In the recent years, young people strongly linked with homegrown tech companies have been taking up positions in the public sector, often without relinquishing their private sector jobs. Ruangguru co-founder, co-owner and CEO, Adamas Belva, was appointed as President Joko Widodo's special staff in 2019. Founder of ride-hailing decacorn Gojek became the Minister for Education and Culture in 2019, and co-founder of the e-commerce unicorn Bukalapak was given a job in data analytics management at Telkomsel in 2020 (Hicks, 2021, p. 15). Such arrangements might lead to conflicts of interest, real or perceived, as such individuals may be expected to lobby for their companies' involvement in government projects. During the pandemic, the Indonesian government relied on Gojek's help in the distribution of Covid-19 vaccines and providing loans for SMEs (Darmawan, 2021). Belva's appointment as special staffer to the President coincided with Ruangguru's selection as one of the government's partners for its pre-employment card programme. Although the Cabinet Secretary insisted there was no conflict of interest, the public controversy suggested otherwise which eventually led to Belva's resignation (Arbi, 2020). Another startup-founder-turned-special staff, Andi Taufan, the CEO of fintech Amartha, was caught sending letters encouraging all sub-district heads across the country to involve Amartha's services in their Covid-19 mitigation efforts, using the official letterhead of the Indonesian Cabinet Secretariat (Hakim & Galif, 2020). Apart from potentially compromising the reliability of such private/public partners, conflicts of interest could represent a gateway into data sharing without transparency. There is particular concern that individuals crossing over from the largely not transparent private sector could carry with themselves these approaches and attitudes to transparency into the public sector, where they might impact the governance of the sharing of citizen data in public administrative contexts, which ought to be more attuned to representative responsibility.

Challenges

- Consistent with most others, this snapshot presents issues of transparency in terms of basic issues of data collection, use and sharing. User data is most likely reshared widely, but some of the information on sharing presented in the privacy policy was found to be unclear and even potentially misleading. In particular, the purposes of sharing user data on the platform are described in a vague and open-ended manner. It is also not clear who are the third parties the data relevant to the user will be shared with, and by. These information deficits adversely impact accountability and participation, in that user consent cannot be truly informed. *Building on already-noted concerns about information deficit, transparency and accountability, this snapshot adds the concern about the misunderstandings and even misleading information on which consent is based.* If consent is meant to legitimize data access and use, it needs to be fully informed and grounded in the certainty that consenting parties understand all the consequences for their data in consenting to its incremental use (this is a requirement of the model presented in Part C).

- Somewhat like in the SatuSehat snapshot (Snapshot A), consent for data sharing on the Ruangguru platform is quite broad and appears to be designed in a somewhat problematic way. As it is not directly required to access educational material, sharing of user data with advertisers could well be the object of separate consent. Presuming this consent appears particularly problematic given that Ruangguru has collaborated with the government to offer educational programmes, and has been relied on at schools (which means that students had no choice but to accept all terms in order to access the mandatory educational content). Separately, Ruangguru relies on deemed consent (i.e. pushing the obligation to request it onto the user) to access the potentially large and profitable troves of personal data of third persons stored in the address book ('Contacts' app) on the phones of the app's many users. *Considering that the platform freely admits to sharing user data for purposes unrelated to the platform's educational services (i.e., for advertising), the challenge rests in a concern whether the data of user's contacts will not be shared further by the platform, for example together with the data of given user, and whether such approach to consent can validate or legitimate data access and use.*

- As per the Privacy policy, the execution of user data deletion requests is seemingly made conditional on the platform no longer needing that data, and in any case, it appears that the deletion of data may happen long (up to 5 years) after the request, during which time the data might conceivably still be shared for advertising purposes. In any case, there is no way to demand the deletion of any user data from the third parties Ruangguru will have shared it with already. *The challenge rests in unconditional data retention.* The model in Part C is intended to produce a mutualizing of interests as a basis for trusted data relations. Any absolute claim for data retention would not be consistent with the spirit of mutualizing interests.

- Public trust in good data governance on the Ruangguru platform, and possibly in public-private partnerships for the digital delivery of public services more generally might have been damaged by patterns of conflict of interest and even suspicions of possible favouritism in accessing government contracts. *The challenge here is that if conflicts of interest underpin the dynamics of data sharing, trust is at risk.* For the model in Part C, trust is crucial for safe data spaces. If the relationship between the main data sharers is suspected to undermine trust, it will not be consistent with the good governance outcomes that the model defends.

Snapshot F: DELIMa (Malaysia)

It is no longer okay to use chalk and board in a classroom… the DELIMa platform is a novel way to extend the learning culture.
Dr Habibah Abdul Rahim, former DG of Education, MOE, Malaysia

Introduction

Digital Educational Learning Initiative Malaysia (DELIMa) is a learning platform for schools offered by MOE in collaboration with US Big Tech companies: Google, Microsoft and Apple.[119] The platform focuses on digital teaching and learning involving teachers, students and parents, integrating free digital learning technology and resources such as Google Classroom, Microsoft O365 and Apple Teacher Learning Centre. Advertised as a one-stop platform for all of one's educational and teaching needs (KEM, 2020a), DELIMa integrates artificial intelligence (AI) and machine learning to provide personalized, interactive learning and teaching experiences (Chin, 2022). It provides teachers with the tools to deliver real-time daily online lessons, create assignments, track student scores, share educational content and communicate with students. DELIMa equally incorporates ready-made learning modules contributed by third parties – for example, the STEM and road safety modules for secondary school students in Malaysia have been developed by Shell Malaysia in collaboration with the Malaysian Institute of Road Safety Research, the Road and Transport Department and the police (Aziz, 2022).

DELIMa was introduced in 2020 as the platform of choice of the Ministry of Education to support online learning and teaching during the Covid-19 pandemic (Hii & Sofwan, 2022). By now the use of DELIMa is embedded in the educational system of Malaysia, linked with its vision to leverage ICT to transform the sector (KEM, 2013), and the promise of a democratized, accessible and affordable learning for all Malaysians (Kapong, 2023). According to UNICEF, nearly 5 million users (99% of teachers and 85% of students) accessed the platform in 2022 (UNICEF, 2022).

Data sharing

As an aggregator for third-party educational services and products of its strategic tech partners as well as other providers, the platform itself apparently does not collect much data (KEM, 2020b). At the same time, DELIMa's partners will most likely be collecting extensive data in the form of digital traces left by users in all instances of interaction with the platform-embedded educational tools and services. According to recent studies such as the international investigation of edu tech platforms by the Human Rights Watch, the collection of edu tech data, made possible through the tracking of students as well as teachers and parents, tends to be pervasive, and the opt-out opportunities are very limited (HRW, 2022). Notably, in a 2022 study conducted by the non-profit Internet Safety Labs, 96% of edu tech apps used in US schools were found to be sharing student data with third parties, mostly for advertising purposes and usually without the consent of users or schools (ISL, 2022). Indeed, the Human Rights Watch's investigation found 3 ad trackers on the DELIMa platform to be sending user data to Google alone (HRW, 2021b)

The data collected on users will be used for providing the services as well as the customization of user experience on the platform, but it could also potentially be used for other purposes, such as advertising or machine learning. Google and Microsoft, two of DELIMa's strategic partners and at the same time contestants in the global AI arms race, have an incentive to access large quantities of data to train AI models. While the 2023 update of Google's privacy policy states that the company only uses 'publicly available data' to train ChatGPT's rival Bard, data from products such as Gmail can nevertheless be used to train other AI systems which increase the attractiveness of Google products (Fowler, 2023).

DELIMa leverages Google, Microsoft and Apple technology as well as other applications, tools and services to provide online courses, online classrooms, and student materials. Notably, DELIMa's main website is directly hosted by Google. The platform relies heavily on Google products such as Gmail and Google Drive as well as Google Classroom (which used to be the primary feature of the MOE-supported platform before its pandemic-time expansion and rebranding as DELIMa). Google Classroom can be integrated with any number of third-party apps, which potentially gives Google access to a wealth of valuable customer usage data generated across such products and companies (Williamson & Hogan, 2020, p. 45).

To use DELIMa, each student and teacher is given access to a dedicated Gmail account and Google Drive storage. In fact, the only way to access DELIMa is by logging in first with a Google account, which involves sharing data with that company. Google's practices in relation to gathering analytical information from these platforms are widely commented on; in particular, Gmail is known to collect much information about users, some of which is shared with advertisers (O'Flaherty 2021). And while the company promises 'not to sell' the data of users, this promise is not to be taken too literally as it largely obscures a sprawling, complex system of automated ad placement through real-time bidding auctions, where sensitive data of users is shared with multiple AdTech companies and Google is involved at almost every step of the way (Cyphers, 2020). The promise sounds even less reassuring considering that an earlier wording of

[119] Apart from the three major technological partners, Ministry of Education Malaysia has partnered with UNICEF, DiGi Telecommunications and Malaysia Digital Economy Corporation for this initiative.

the same privacy policy clause used to read 'we do not transfer your data' up until the mid 2000s, which is when the tech giant began consolidating its new data-based advertising business model (Warzel & Ngu, 2019).

Equally, in light of Microsoft's Privacy Statement, the company uses the personal data of users for advertising and marketing to them 'which includes sending promotional communications, targeting advertising, and presenting [them] with relevant offers' (Microsoft, 2023). In the investigation of edu tech platforms by the Human Rights Watch, Microsoft Teams and Zoom (both learning tools accessible through DELIMa) were found to embed code capable of collecting contacts' information, saved photos and call logs, as well as precise location, and to use software development kits (SDKs) – third-party code, which may serve as a persistent channel sending user data directly back to the third-party developer of the SDK (HRW, 2021c and 2021d). SDKs tend to run somewhat 'behind the scenes' of the app they are embedded in. While some of the purposes of sharing may be quite innocuous, SDKs tend not to ask for the user's consent to access his or her data (LeVasseur et al., 2021). On the other hand, it is not always clear why these applications collect user location data, which alone can be quite destructive to privacy (Kelly, 2021b). As reported by the Human Rights Watch, Zoom did not disclose that it was collecting location data in its privacy policy, whereas Microsoft said that MS Teams collects that data 'for many purposes', which include compliance with US laws (HRW, 2022).

Governance concerns

Edu tech data is continuously generated and analysed, primarily using private technology in a public (educational) context, raising questions about who will exercise control over it throughout its lifecycle involving collection, use, storing and repurposing, or 'extraction, enclosure, aggregation, analysis, and transformation into intelligence' (Komljenovic, 2021). The pervasiveness of surveillance allows to profile the data subjects to understand them in minuscule detail, including insights they themselves may not be privy to, such as how and they learn best. These profiles allow for adaptive, personalized approaches to learning, but can equally be very much in demand for other reasons, including correlating with other datasets, leading to more valuable insights not limited to the educational sector. While a lot of academic commentary focuses on the data of students, data-driven educational technology can also allow for the monitoring, ordering and assessment of teachers, limiting their autonomy in the workplace (Kwet, 2017). Additionally, their data can equally be sold and used for profiling in obscure ways through edu tech platforms (Arantes, 2023). The first step to gain some control over one's data is to understand how it is being used, which would correspond to the right to information as enshrined in Section 7 of the Malaysian PDPA, if applicable to a private sector entity working through government mandate. But that may involve significant hardship for users of DELIMa, for reasons which will be explained in the next paragraphs.

The Terms and Conditions of the use of DELIMa are quite limited, and do not address data practices or sharing. The privacy policy of DELIMa, on the other hand, insists that no personal data will be collected 'unless submitted via electronic form or email'. It may at first appear as though this policy also applies to DELIMa's partners whose website or application links it hosts; however, any suggestion that no user data is collected automatically has to be rejected in light of DELIMa's capacity to provide personalized digital services (KEM, 2020b). Indeed, the privacy policy urges users to 'review and understand the privacy policies of each website they visit' in following any links from the DELIMa portal, but without naming any such partner websites (or linked products) upfront. Since DELIMa is essentially a constellation of apps and products by different providers, using multiple DELIMa features may trigger different privacy policies (and terms of use). These in turn may additionally refer the user to the privacy policies of third parties (partners), and then their partners etc., in what is known as a 'cascade of data sharing' (Forbrukerrådet, 2020, p. 123). As an example, authors of the 'Out of Control' report have established that to understand what the dating app Grindr and the only advertising partner named in the app's policy do with user data would require reading the policies of over 160 partners of Grindr, as well as the partners of partners, whose numbers run into the thousands (Forbrukerrådet, 2020, pp. 123-160). Likewise, it might be expected that understanding how the data of DELIMa users may be used and shared by its partners and their respective networks as a consequence of using the platform might involve a potentially overwhelming amount of reading for those authorized to grant consent.[120]

Another issue is the matter of understanding the terms of those policies, underpinning, as they do, consent and participation. An investigation by a children's digital rights charity in the UK found that while Google Classroom and ClassDojo (free edu tech products popular during the pandemic) exposed children's data to commercial exploitation, their terms of service (including privacy terms) were opaque and confusing even to teachers (Hooper, Livingstone and Pothong, 2022). Swedish researchers have argued that the privacy policies of Google edu tech products conceal the company's true advertising-based business model, allowing it to pose as an ethical, free public institution (Lindh & Nolin, 2016). When threatened with a ban on its edu tech products over privacy concerns in the Netherlands, Google was forced to amend its terms of service and renegotiate contracts with schools (Criddle, 2022). Malaysia is a very linguistically diverse country, some parts of which do not speak Malay. While English is the country's second language, that does not necessarily imply a general linguistic capacity up to the standard required to understand privacy policies, which are notorious for using inaccessible jargon and giving words idiosyncratic meanings (Ng, 2021), or not explaining

[120] In light of the Personal Data Protection Regulations (2013), for data subjects under the age of eighteen, the data user shall obtain consent to process their personal data from an adult, such as a parent or guardian.

key terms at all (Lindh & Nolin, 2016). Considering that privacy specialists have struggled to understand the policies of Microsoft,[121] it might be especially unreasonable to expect a user base comprised largely of non-native speakers of English (many of whom may in fact hope to learn it through the platform) to have a better understanding of them.

Aside from the potentially unclear message of the privacy policies, there is also the question of whether these will be sufficiently relevant to the Malaysian regulatory context. As reported by the New York Times, Google has been somewhat reluctant to commit to the compliance of edu tech products offered in a number of schools across the US with local laws and policies (Singer, 2017). On the other hand, a study of edu tech apps in an Australian educational context has found that most of them tended to refer to laws and frameworks of their countries of creation rather than Australian ones (Rennie et al., 2019), which may be suggestive of patterns of tech colonialism. All these information-related challenges could have an impact on the validity of consent, and on accountability for user data sharing. While discharging information duties in a sloppy and potentially even obstructive way might be nothing unusual in the platform economy (experts such as Shoshana Zuboff in an interview with Mance (2023), or the non-profit Norwegian Consumer Council (Forbrukerrådet, 2020) go so far as to claim that a lot of 'informing' by the private sector is merely meant to extort consent from users), transplanting these practices into educational (public) sector delivery is a matter of concern.

The processing of edu tech data into intelligence leads to insights which may, on the one hand, allow to constantly improve and tailor an existing product or service to the user (e.g. to enable and refine data-based automation of assessments of ability, and adjust the difficulty of material using algorithms in schools), or to offer brand new products and services, such as intelligence on students or tailored advertising (Komljenovic, 2021). Features such as adaptive learning or personalized learning approaches will be offered to DELIMa users based on algorithms which will almost certainly be proprietary. This means that there will be no way for users, schools or even government officials to understand how they gave rise to specific outcomes. This implies there will be no accountability for decisions which may have an impact not only on the users' experiences with the DELIMa platform, but potentially limiting their life opportunities more generally, e.g. through mistaken assessments of ability.

Public-private partnerships imply a kind of official public sector endorsement of the private partner. There will be an expectation that the government has vetted those partners to assure data user safety in educational service delivery, especially considering the vulnerability of the main target group (schoolchildren), and the effectively compulsory character of using DELIMa in schools. While the platform provides some options such as whether to use Apple or Microsoft products within the platform, there seem to be no alternative edu tech platforms available that offer a similar service, or enjoy similar government support in Malaysia. However, no official or unofficial statements or assurances relating to the data practices of the tech partners of DELIMa from the point of view of the privacy of users or compliance with the PDPA have been identified from generally available sources. What is more, DELIMa's T&Cs appear to exclude the liability of the MOE for losses originating from, among others, unauthorized access or alteration in users' transmissions or data, and statements or actions of third parties in DELIMa (KEM, 2021). On the other hand, the government itself is not bound by Malaysia's PDPA, and thus cannot be formally held accountable for its breaches in terms of data sharing. Moreover, even if there is some public accountability back to an elected government, it may be additionally limited in practice if the government is one side of a conflict of interest. Therefore, it may seem as though there is an accountability gap in terms of the data sharing happening in the context of this public-private partnership. But equally, users of DELIMa as an edu tech solution endorsed at the school and ministerial level may appear to have no real choice of whether to consent to the data practices, also considering the pre-existing power imbalances in an educational context. Data-based ordering in schools is generally used by hierarchies of power to discipline data subjects and determine their merit (Jung Han, 2020), with little opportunity for data-subject participation and accountability (Hillman 2019; Lindh & Nolin 2016, p. 657). But importantly, Malaysia has the highest power distance index in the world according to the framework developed by the Dutch social psychologist Geert Hofstede, which indicates the willingness of the less powerful members of institutions or organizations to accept that power is distributed unequally (The Culture Factor, 2023).[122] Indeed, the GDPR makes it clear that in situations of power asymmetry, consent given by a data subject is not to be considered 'free'.[123] It is also not 'free' if it is a condition for a provision of a service.

The DELIMa platform engages with foreign Big Tech companies which are among the most powerful international players in the field. This raises concerns whether this power imbalance may translate into domination of these partners in terms of decision-making, and an inability of the less powerful partners to assert themselves and realize their goals and obligations such as ensuring an adequate protection of users of the platform. Separately, from the point of view of local tech companies, engagement with foreign Big Tech may amount to data and tech colonialism. One such stakeholder criticized the MOE for 'arbitrarily giving data to foreign technology providers allowing them to achieve their ambitions, [and] leaving us becoming glorified consumers' (Al-Shahab, 2020). It is no coincidence that proponents

[121] A Concerning the lack of clarity on whether Microsoft may or may not use the data of users to train generative AI, see p. 43
[122] By comparison, Singapore and Indonesia score 74 and 78 respectively, which still qualifies them as 'high' power distance countries (The Culture Factor, 2023).
[123] In light of recital 43 of the GDPR, 'In order to ensure that consent is freely given, consent should not provide a valid legal ground for the processing of personal data in a specific case where there is a clear imbalance between the data subject and the controller, in particular where the controller is a public authority and it is therefore unlikely that consent was freely given in all the circumstances of that specific situation.'

of the data, or digital colonialism thesis have been sounding an alarm about the often exploitative practices of data extraction from the Global South to the North, made possible thanks to Big Tech's sprawling infrastructures which often tend to suffocate the domestic tech industry (Couldry & Mejias, 2019). In the realm of public-private partnerships, what might be of particular concern is the risk of a government's dependency on (foreign) Big Tech for the delivery of essential services such as education to the public, resulting in an inability to protect the citizens from de facto 'paying' the tech providers in the data they generate through interacting with their platforms. Infrastructural domination is typically cemented by the limits proprietary software imposes on understanding its inner workings on users' own computers, which ultimately thwarts the quest for accountability of the Big Tech (Kwet, 2019).

Largely against the backdrop of the global Covid-19 pandemic, Microsoft and Google especially have become 'globe-spanning infrastructures for remote education' (Williamson, 2021) through embedding its software and cloud infrastructure in public education. At the same time, Google in particular has been able to consolidate its position through acting on several levels: reaching out directly to educators and thus often bypassing district officials (Singer, 2017); working with international institutions such as UNICEF;[124] and offering their products and services for free to governments as CSR, which in turn has enabled it to avoid traditional procurement routes and processes (Williamson & Hogan, 2020, p. 44). While it should be observed that apparently 'free' technology tends to be paid for through advertising[125] and therefore paradoxically is often the most attractive option for schools that are chronically underfunded and unable to afford costly data safety and privacy solutions (Kwet, 2017;[126] Lindh & Nolin, 2016), it is not clear how the Malaysian government appointed its business partners to develop DELIMa. If the example of MySejahtera[127] is anything to go by, it is not inconceivable that procurement procedures were not followed in the case of DELIMa, either. Microsoft, on the other hand, has tended to secure key partner roles in post-pandemic public-private partnerships for education delivery as well as endeavouring to influence policy with its vision for the future of education (Williamson & Hogan, 2020, pp. 43-44). It may be observed that the Malaysian government has recently emphasized digital education, digital literacy and infrastructure for virtual learning environments as a priority in the 2024 budget (Malay Mail, 2023). Some of the key areas of focus of the Digital Education Policy (which is soon to be implemented) include embedding a culture of digital technology usage among education leaders and strengthening strategic partner networks (Daim & Radhi, 2023).

If J. Komljenovic is right that 'what becomes valuable in digital education is power over the direction of student and staff teaching, learning and work patterns' (Komljenovic, 2021, p. 325), the ensuing power asymmetries and dependencies on the private sector are a source of concern in an essentially public sphere, especially where education is compulsory. Part of the educational endeavour of edu tech seems to involve teaching teachers 'what works' in educational technology. While this may in itself be biased and based on reductionist methodologies as suggested by Williamson (2021), the more urgent concern is that it might effectively capture fundamental educational purposes and policies. This can be linked to an overarching ambition of the technology providers to maintain their position in education service delivery, but equally to the hope of retaining the users of these products and services beyond school, to ensure a steady, unrestrained flow of user data. While Google claims not to use any data of K-12 (primary and secondary school) students for advertising purposes (Google for Education, nd.), that does not mean that it does not monetize their data in other ways (Williamson & Hogan, 2020, p. 46; Lindh & Nolin, 2016, p. 658). In any case, any such age limits will no longer apply as the students grow older, all the while relying on the long-familiar products and services for work purposes and in their day-to-day lives. From that point of view, Google edu tech products have faced criticism of 'training Google consumers from infancy', and accused of being a brand loyalty scheme in the guise of a free education revolution (Krein, 2020).

Challenges

- The DELIMa platform provides a gateway to sharing the data of users (teachers and users) on a potentially large scale, in ways that are not very transparent, impacting accountability for sharing and user consent. *The challenge here is particularized by the educational context and the relationships of moral agency and student/parent vulnerability.* The model presented in Part C recognizes the dangers inherent in power asymmetries. If the pre-existing power hierarchies in a context like this can be compounded and expanded by tech dependencies, then the powerful stakeholders need to be especially mindful of the governance needs of data subjects, and how these can be mutualized.

- While the user is somewhat compelled to trust the app if they want to participate in its effectively monopolist services, there is a lack of easily accessible basic information on data practices and sharing linked to the use of the platform. Clear, approachable and detailed information on sharing could foster a more vigorous debate and would

[124] It may also be worth observing here that the DELIMa initiative has enjoyed the support of UNICEF.
[125] This paradox was addressed by the Italian Antitrust Authority (Autorità Garante della Concorrenza e del Mercato, or AGCM) in a decision in which Facebook was fined €5 million for having misled its users into believing that the social network was free rather than paid for by the commercial exploitation of the users' personal data. The decision was later upheld by the Regional Administrative Tribunal (TAR) of Lazio in 2020 and by the Italian Council of State (Consiglio di Stato) in 2021.
[126] See in particular the remarks by an official of the Gauteng Department of Education (South Africa) quoted by Kwet (2017) at fn. 41.
[127] See Snapshot B.

allow the users, if they so wished, to gain a degree of empowerment from at least knowing how their data might be used, and whether and if so, how they may be 'paying' for education in data. *The challenge is in the compulsory environment, and how this exacerbates data subject vulnerability.* The model in Part C requires data subject autonomy. In situations of data exchange where that autonomy is denied, mutualizing will be difficult and additional steps may be required preceding participating in the model to modify compulsion.

- In particular, an apparent failure by the public sector stakeholder to provide any 'outer layer' information on data sharing linked to the use of the platform, obscuring an invisible cascading of data sharing (likely triggering multiple privacy policies at once) could be an important impediment to knowledge and comprehension which are preconditions to participation. *The challenge consists in a dereliction of governance responsibility by one particular stakeholder*. In the model presented in Part C, there is an expectation that stakeholders will come to the negotiating table recognizing and representing some of the fundamental responsibilities they have in governance at large. That means, for instance, the public sector players continue to retain their 'political' obligation to citizen/residents.

- If DELIMa users (most of whom are not native speakers of English) are interested in learning about the use of their data, they are left to rely on multiple and usually long and complex third-party (partner) privacy policies. On the other hand, the often incomprehensible, potentially confusing and even manipulative statements found in privacy policies across a wide range of services offered from major tech and media platforms has been widely commented on (see e.g. Litman-Navarro, 2019). These shortcomings in the partner policies themselves could probably be fixed relatively easily, provided there are incentives to do so. A more difficult, but at the same time impactful option would involve, in line with the proposal by N. Couldry and U. Mejias, emphasizing public education – 'in the form of citizen research, literacy campaigns, decolonial thinking – to understand the dangers of data colonialism' (Couldry & Mejias, 2023, p. 798). *The challenge could consist in the failure of data-powerful stakeholders to motivate data subjects through a more informed and generous openness to participate in the governance of their data.* The model described in Part C requires from stakeholders a willingness to participate in respectful and responsible data relations and this may require more energy and concessions from powerful data holders to generate trust from the outset.

- The public-private partnership underlying the development and use of DELIMa and the architecture of the platform imply an endorsement of private tech providers, but at the same time it is not clear if those partners and their data practices have been vetted. Without contractual limits on DELIMa required by public sector partners, an external governance opportunity is squandered. *The challenge here is that the legitimacy of the public/private partnership is not tested by conventional public sector probity requirements.* If parties come to the negotiating table with the disposition to distrust each other, the initial effort to establish safe spaces and trusted relationships as foreseen by the model presented in Part C will be all the more difficult.

- *Flowing from the preceding challenge, the accountability for data sharing between the public and the private sector involved appears to be diluted.* Before the PDPA was enacted, the Malaysian Ministry of Education was reported to have been selling student data to private, for-profit colleges who then targeted the students with advertisements (Ismail & Yong Cieh, 2013, p. 18). Although the PDPA is now in force, it is not clear to which extent it would regulate the sharing of data on the DELIMa platform; in any case, the government is not accountable in light of the PDPA. Additionally, the PDPA is only applicable to commercial transactions defined as involving 'any matters relating to the supply or exchange of goods or services'. Graham Greenleaf supposes that non-commercial educational services would therefore be excluded from its scope (Greenleaf, 2010); but in any case, this limitation of its scope of application can result in confusion (Walters, Trakman & Zeller, 2019, p. 214). On the other hand, potentially making (foreign) Big Tech accountable for uses of data of DELIMa users is not straightforward, even less so considering that the PDPA is not applicable to the processing of data outside of Malaysia. Additionally, it might be difficult to understand particular choices and decisions made by digital products (e.g. to adapt the learning to the user) as these will likely be determined by proprietary algorithms. It is not sufficient to claim property protections over publicly endorsed programmes or applications in order to deny the possibility of information openness.

- There is also concern about lack of accountability for possible interference with public policies and programmes, and for embedding given products and services into the lives of users at an impressionable age as a result of the public-private partnerships underlying DELIMa. *Vulnerable data subject demographics are a significant context in which special challenges arise from data sharing.* The model described in Part C is context-specific and as such when vulnerable data subjects are present, there is an additional need for external support (perhaps through data stewards or data cooperatives) to assist in achieving data parity.

- There is a significant power consolidation/disparity aspect to this partnership, possibly made worse by an apparent existence of a dependency to provide digital education services to millions of Malaysian users. *The challenge is connected to pre-existing, customary top-down power arrangements (Malaysia's high distance culture) in schools and classrooms, further exacerbating the complex power asymmetries at play.* As a result of this, user consent for data sharing is unlikely to qualify as 'free'. The cultural aspect and how that plays out in appreciating autonomous consent of the user demographic underscores the importance of bottom-up participation and engagement, as well as the need for power dispersal, both essential to progressing the model described in Part C.

- While algorithms embedded in edu tech applications indirectly teach the students, so do students 'teach' (or train) the algorithms through their data (human-machine feedback loops). *The profound challenge of such data-sharing platforms is that as enriching as the educational experience may be, there is a need to ensure that data-based educational technologies do not leave students impoverished as data subjects, through creating opportunities to view, understand and interrogate data practices and uses relevant to them. What may be at stake is the same self-development that is the great promise of education* (Risse, 2023, pp. 160-172). The model in Part C has as its entry point an acceptance of the priority of data subjects, whoever they are, and if they are vulnerable, the model will require greater vigilance in the processes of power dispersal.

Connecting thoughts: towards Part C

The snapshots presented in Part B are evidence of multiple challenges, both context-specific and more universal, to the good governance of sharing of citizen data encountered in the context of public service delivery enabled by collaborations with private tech providers. While many of the challenges that emerge from the analysis are addressed in a normative sense by principles linked to the Rule of Law such as fairness, equality, and accountability, they remain largely out of sight for current regulatory approaches to privacy and data protection, oblivious as they are to the dynamics of power underlying contemporary data relationships through mass data sharing. Reference is repeatedly made to the need for data openness and for more meaningful information sharing with data subjects. Equally, we argue for clearer lines of accountability, which tends to be impacted where the distinct governance domains of the public and private sectors seem to become blurred. In particular, there is a need to move beyond reliance on inflated consent and unreasonable data subject diligence requirements to achieve accountable governance of data sharing.

The governance model which follows does not propose an overhaul of existing governance and regulatory approaches, or suggest that tech companies ought to entirely abandon their data-driven business models. Instead, it offers a supplementary governance strategy of mass data sharing emphasizing power dispersal in relations of data sharing in the spirit of empowerment of data subjects, which is capable of coexisting with the extant approaches and contexts. The conceptual work behind the model to some extent channels the thinking which sees systems of social collaboration such as peer review (Riles, 2020, p. 15) as well as digital self-determination and grassroots data co-operatives as capable of generating norms, and even as alternatives to legal enforcement in their own right.

Considering the main gaps consistently emerging across all the snapshots analysed in Part B, i.e. power asymmetries; information deficits and the concomitant challenge to accountability for data use and sharing, and lack of trust, the model emphasizes three important elements:

- Vulnerable individuals and communities (data subjects) as the specific targets and addressees of novel governance strategies;
- Transparency and accountability relative to data as key to power dispersal as well as informed and meaningful participation;
- Availability of a mechanism of contestability at every step of the governance process (in order to safeguard trust as the cornerstone of relationships between stakeholders).

Importantly, recent discussions of new directions for governing AI and data use from the UK, EU and the USA flag these considerations as crucial. The new discourse makes it clear that there is a need for more governance options premised on data subject awareness and inclusion. In the model presented in Part C of the Report these aspirations gain detailed development against grounded normative frames such as the Rule of Law. The model does not centre on sanctions, relying instead on longstanding commitments (within public administration in particular) to follow Rule of Law principles. Rule of Law enforcement might strengthen the collaborative endeavour in encouraging initial participation, and whenever good will and trust are strained as a matter of dispute resolution. The principles of justice, fairness and equality would guide the operation of dispute resolution conversations and processes as developed by the stakeholders in any contextual operation of the governance model.

Inspired by the work of social theorists such as Durkheim (in our view of social spaces and the germinating and nurturing of a collective conscience out of mutualised data interests) and Cotterrell (regarding communities as bonds of trust), and to some extent by the related thinking of law and anthropology professor Annelise Riles on a 'platform for collaboration' as 'part regulation, part technology and part social organization… [that] incorporates elements of both states and markets and welcomes different kinds of stakeholders' (Riles, 2020, p. 18), the proposed governance strategy focuses on social collaboration and forging relationships of trust as a cornerstone of a truly communal data sharing. More particularly, in the spirit of ensuring inclusion and meaningful participation of all stakeholders (data subjects, i.e. citizen/residents and their communities; public administrators, and private service providers) in the data-sharing governance enterprise, the model emphasizes co-creation and co-production by representatives of all stakeholder groups as a pathway to identifying governance agendas based on sustainable, responsible and respectful data relationships. The creation and maintenance of such relationships does not appear to have been considered in conceptualizing and operationalizing the public-private partnerships underlying the six snapshots; if it was, it requires a far stronger emphasis. In addition to a renewed legitimacy of the data sharing governance project in an era of data-

driven decision-making as well as tech-solutionism more broadly (Bayamlıoğlu and Leenes, 2018),[128] these governance underpinnings offer a way to chart novel, citizen-centric directions for governing data-sharing across the three orbits, as well as other areas of public-private service delivery. In that sense, the model argues against the dispossession of public sector's governance duties inherent in the traditional notion of 'public and private partnerships'. Equally, it endeavours to expand the typically narrow outlook, focus and composition of the classic private law distinction between contractual parties and non-parties, in order to extend a degree of contractual standing to those with genuine interests, as advanced by Perez (2002).

[128] As E. Bayamlıoğlu and R. Leenes write, 'Big data constrains the possibilities for political and moral choices by reducing governance to a technical process of adaptation, and law to a process of optimisation – rendering politics a mere question of "better-doing"' (Bayamlıoğlu & Leenes, 2018, p. 312).

Part C: A suggested governance model

Overview: towards an alternative strategy for governing data sharing

Part B of the Report has established the need for a purpose-designed governance strategy addressing the challenges posed to personal data integrity by mass data sharing in smart cities. The analysis has targeted some of the most vital areas of service provision where public and private data[129] is shared in the fields of education, health service delivery, infrastructure, consolidated welfare service delivery and the management of state-centred information. As also discussed in Part B, good governance of large-scale sharing of citizen data in the region may in some cases be hampered by challenges which include:

- Reliance on imported tech infrastructure and capacity (tech colonialism and dependency);
- Reduced capacity to understand the risks (lack of awareness and information deficit),
- Digital, cultural and linguistic divides (for example, some parts of Indonesia don't speak Bahasa; additionally, traditional communities may reject digital lifestyles)
- Cultural/social attitudes (trust/distrust, convenience, defeatism, high-distance culture);
- Narratives such as tech solutionism or 'service-for-data' narrative (implying that there can be no benefit without sharing data),
- Powerful stakeholder dominance, with added difficulties stemming from secrecy, exclusion, lack of competition.

Why should data subjects and their communities be concerned about these challenges and be interested in governance alternatives? As explained in more detail in Part A, it is because the prevailing governance modes are either not designed or not well-suited to manage and control mass data sharing between public and private governance styles. On the other hand, large-scale data sharing involves novel risks to conventional good governance indicators, such as accountability, transparency, information openness, legitimacy, Rule of Law, and citizen-centric participation, which need to be urgently addressed by an alternative governance frame.[130]

What follows is a proposed model for such an alternative, but complementary governance strategy[131] designed to address both local and universal challenges posed by mass data sharing between public and private providers in smart cities.[132] It is based on the intersection of governance and regulation; that being where a wider requirement to 'order' a specified behaviour (in this case the practices of data sharing) is actioned through particular regulatory instruments and processes. While the model is not presented or developed as a determined regulatory policy or an exclusive regulatory strategy, it provides an agenda for a much needed behavioural change in the realm of data relationships,[133] and suggests regulatory approaches that could be taken to operationalise this agenda. As a foundation for informed policy planning, the governance model is meant to specifically address the gap posed by the problem of mass data sharing on the one hand, and the outcome of responsible and respectful access on the other.[134]

Addressing the gaps in pre-existing regulatory regimes

The model is intended to complement rather than replace other governance and regulation regimes that influence data access and management. These include personal data protection regulations, industry best practices, ethical principles, standardisation, constitutional rights to privacy, and a variety of consensual model frameworks. Part B has revealed that while pre-existing governance/regulation regimes may have some positive effects on data access/management, they are largely not directed to mass and repeat data sharing, and more often than not they originate and operate from either public or private sector authority, values and responsibilities. Public-private collaborations involving large-scale sharing of population-derived datasets, such as the one to build and operate the Indonesian health super app SatuSehat (Snapshot A), the Singaporean app LumiHealth, or the Malaysian e-educational platform DELIMa, demonstrate the difficulties involved in ensuring that individual consent, i.e. the cornerstone of most data

[129] It might be argued that the distinction between public and private data is no longer clear, and in some cases even no longer defensible as a result of multiple and proxy sharing as well as deployment of both public and private technology tracking user behaviour for increasingly overlapping purposes (as Cory Doctorow suggests, 'distinguishing between state and private surveillance is a fool's errand' (Doctorow, 2023)). For the purposes of discussing governance stakeholder responsibility, the analysis retains the distinction based on which major agency initially stored and used the data for public or private purposes, or a cross-over of both.

[130] It moreover became apparent in our regional conversations that many of these indicators were too general and oblique, and/or understood very differently in different contexts, and hence that a more lateral and holistic approach to appropriate indictors of good governance was necessary.

[131] The model is alternative (supplementary) because it could coexist with the currently existing data governance modes (which do not adequately address public/private mass data sharing). At the same time, the model is complementary as it does not principally concern itself with privacy protection, data rights or ethical compliance, focusing instead on inclusion and empowerment of the data subjects.

[132] In proposing this duality, the analysis recognises that for centuries there has been no purely distinct and separate domains for the provision of urban services. However, what has made the confluence of public and private service delivery in urban environments presently unique is the practice and process of mass data sharing and the consequent strains this places on very different governance approaches (public administration or market forces).

[133] Behavioural change is the generic purpose of regulation in terms of Julia Black's celebrated definition, whereby 'regulation is the sustained and focused attempt to alter the behaviour of others according to defined standards or purposes with the intention of producing a broadly identified outcome or outcomes, which may involve mechanisms of standard-setting, information gathering and behaviour modification' (Black, 2002, p. 26).

[134] 'Responsible and respectful access' to data is the ultimate goal of governance approaches based on the theory of digital self-determination (Findlay, Seah, & Wong, 2023, chap. 3). See also fn 149.

protection regimes, is truly informed, comprehensive and meaningful. In parts of SE Asia, the added difficulty may consist in ensuring that consent is truly informed in jurisdictions where there is no freedom of information legislation, or where the public sector is not bound by the data protection legislation, as demonstrated in some of the snapshots. Mass data sharing in smart cities merges public/private data sources, and therefore governance approaches designed from either public administration institutions or from market relations will not comprehensively cover a data interface where the public and private converge. In light of this gap in the regulatory tapestry, there is a need for fresh governance thinking on mass data sharing in smart cities. Responding to that need, this chapter proposes a model that recognises the importance of co-creation – which emphasizes collaborative interaction between a plurality of actors in networks and partnerships (Torfing, Sørensen & Røiseland, 2019) – and co-production,[135] incorporating the interests of public and private data users and data subjects into the governance of mass data sharing in smart cities.

In this vein, this proposed model works on a commitment to inclusion and participation, wherein data subjects and their communities (residents and neighbourhoods, whether located in large conurbations or smaller community settings) are central to the governance exercise, and share duties to ensure that data management and access is responsible and respectful towards other stakeholders in the data ecosystem. Considering the concerns of citizen/residents when addressing the governance of mass data sharing, it is vital to regulate both the mundane, 'everyday' kind of data sharing which impacts the data subject, as well as that kind of sharing which is sometimes specifically considered by the powerful data users as worthy of restriction (for example, health data). The objective of the model is to create safe data spaces (involving data relationships of respect, responsibility and mutual interest), within which the nature and impact of sharing, particularly where it may involve manipulations of data integrity, will be better understood. It is the principal aspiration of the model to give all stakeholders the information about data access and use in an environment where these can be more equitably negotiated.[138]

Recent governance proposals

Recent governance options proposed in Europe, the UK and the US have recognised the importance of policy developments (particularly regulating AI) that address fundamental concerns similar to those informing our model. On the side-lines to the negotiations of the EU AI Act, France, Germany, and Italy have reached an agreement on how artificial intelligence (AI) should be regulated, as set forth in their joint Non-paper from November 2023 (Rinke, 2023). While conceding that the intrinsic risks lie in the application of AI systems rather than in the technology itself, the initiative supports 'mandatory self-regulation through codes of conduct' for advanced AI foundation models, including large language models (LLMs), which are designed to deliver a wide range of outcomes (Italy, France & Germany, 2023). The UK is also preparing an AI (regulation) Bill[139] which imposes duties linked to ensuring public engagement in governing AI on the special body created by the new legislation,[140] while in the US senators have introduced a bill for AI regulation that intends to balance accountability and innovation.[141] Equally, the UK government recently held an AI Safety Summit, and Germany has followed with a digital summit bringing together industry, policy-makers and research interests. What this means is that major economic powers are moving with a more concerted urgency to regulate AI, and, by extension, Big Data, in languages that go beyond consent and internal (voluntary) compliance.[142]

---

[135] Following Røiseland (2023), we understand co-creation as an umbrella concept, composed of co-design, co-implementation and co-production. In the application of co-creation and co-production in the model to follow, co-creation emphasises that the engagement with the data subject and her community to co-create governance solutions must take place from the outset of the model, and that co-production gives meaning to participation and inclusion by focusing on governance outcomes. In the model, data subjects come together with the more powerful data stakeholders to negotiate respective interests in the creation of safe data spaces, and the respectful and responsible data relationships that are nurtured in these spaces.

[136] For a definition of safe data spaces, see fn 45.

[137] For a definition of data integrity, see fn 3.

[138] Information deficits lie at the core of data disempowerment for data subjects. If these stakeholders have the opportunity to know more about their data and its uses, they will be able to choose if and how they will participate in its control.

[139] The AI (regulation) Bill will create an oversight body (AI Authority) to coordinate AI regulation across government agencies and businesses employing AI technologies. According to the Bill, 'The functions of the AI Authority are to ensure that relevant regulators take account of AI; ensure alignment of approach across relevant regulators in respect of AI; undertake a gap analysis of regulatory responsibilities in respect of AI.' The Authority will be tasked with coordinating reviews of existing legislation, including aspects like product safety and consumer protection, to gauge their fitness for addressing the challenges and opportunities presented by AI. The Bill delineates principles for the AI Authority to consider in regulating AI, including 'safety, security and robustness; appropriate transparency and explainability; fairness; accountability and governance; contestability and redress.' The notions of contestability and redress are where that proposal aligns with our model, which assumes the capacity for negotiating and mutualising data management concerns among the stakeholders. The UK Bill also requires businesses involved in AI to be transparent, to thoroughly test their AI systems, and comply with existing laws, including those related to data protection. It also emphasizes the need for AI applications to be inclusive and non-discriminatory, catering to diverse socio-economic groups, such as the elderly and disabled people. In a significant move to ensure ethical use of AI, the Bill requires businesses to appoint a designated AI officer, responsible for 'the safe, ethical, unbiased and non-discriminatory use of AI by the business.' Like our model, the Bill recognises the importance of identifying power imbalances, and approaches the dispersal of power in favour of vulnerable stakeholders by having AI (and data use) cater to the interests of less powerful stakeholders.

[140] In light of Sec. 6 of the AI (regulation) Bill, the AI Authority must:
(a) implement a programme for meaningful, long-term public engagement about the opportunities and risks presented by AI; and
(b) consult the general public and such persons as it considers appropriate as to the most effective frameworks for public engagement, having regard to international comparators.

[141] What is compatible with the spirit of our model is that bill's recognition of consumer (in other words, the data subject) interests and the need for this stakeholder group to be better informed through concerted education programmes.

[142] As the recent regulatory push to introduce elements of more rigorous oversight, and even compulsory notification of AI potentials and practices, in state governance language there is a move away from voluntary compliance.

Much of the thinking behind the minority EU proposal (the 'non-paper') resonates with governance modelling that focuses on data, and recognises the importance of stakeholder interests beyond business and corporate governance. Despite the emphasis on risk, both the AI Act and the minority proposal emphasize the need to focus on AI applications, and implicitly on data access and use to fuel and direct AI in contexts of human engagement, risky or otherwise. In relying on 'mandatory self-regulation', the proposal is echoing J. Braithwaite's theory of enforced self-regulation (Braithwaite, 1982), which has informed the model to follow. This approach accepts that stated principles and best practice suffice for the great majority of powerful data stakeholders to engage in respectful and responsible data management. For the few that do not respond to this, some degree of compulsion from an external authority might be required at least in the early stages of data negotiation. Our model recognises the importance of universal principles and best practice invocations, as well as contextual ownership of the process by data subjects. It also anticipates the utility of tools such as stewardship and data licensing to stimulate initial engagement and negotiation among stakeholders.

Another feature of the mandatory self-regulation model which aligns with our model is the basic requirement for openness in the way that data is used in LLMs and other mechanisms of decision-making with direct influence on the interests of data subjects and their communities. The suggestion about 'model cards' made in the Non-paper (Italy, France & Germany, 2023) is intended to concretise openness and transparency. Additionally, these cards would be a repository to which other stakeholders in a data ecosystem could always return for information about data applications. Also, there is the potential for this information frame to produce reflection and form the basis for ongoing assessment and evaluation which hopefully would have positive impacts on trusted data spaces and trust relationships that give definition and parameters to space.

The mandatory self-regulation strategy relies on external oversight and potential intervention, even sanctioning (as a last resort), which are not particularly emphasised in our model. However, the mutuality on which self-regulation is premised is the fundamental and overarching commitment in our approach. While the compulsory self-regulation proposal institutionalises openness processes and reporting requirements, we strongly encourage the data-powerful stakeholders to assume more formalised transparency and accountability obligations that are directed to the interests of data subjects in first instance, in the particular contextual operations of our proposed model. But above all, the core belief underlying both mandatory self-regulation and our model is that regulating technology alone is not sufficient if we want a more responsible digital environment. The model to follow goes further by offering all stakeholders, and data subjects in particular, active roles in co-producing trusted data spaces and respectful/responsible data engagements.

The importance of these proposed developments (and those already in action) is not only that they address the application of AI tech and the importance of data (not just tech) as foci for good governance. To various degrees, the proposals consider the interests of data subjects, discuss accountability to users, anticipate contestation from data subjects, and advocate wider education and awareness about AI application at all stages. The recently revised AI strategy in Singapore also identifies data, trust and empowerment as significant factors in encouraging community knowledge and confident participation in AI expansion (MICI & Smart Nation Singapore, 2023). These themes are evidenced, as well, in our proposed model.

Central principle of the model: citizen-centricity and power dispersal

The model places the data subject at its centre; in a smart city context, this is sometimes referred to as citizen-centric policy.[143] We are aware that operationalizing citizen-centricity may encounter two kinds of hurdles. First, it would be operationalized in data ecosystems typically already constructed around strong power asymmetries, which may not recognise respect for and empowerment of the data subject (citizen) as a priority in data control, management, access or transaction. Second, even where the data subject/citizen can be empowered through data governance, this may not prevent their capture by other powerful interests, their negotiating away of their primary preference not to share their data in order to readily access a digital service, or even the apathy and lack of interest in empowerment among disinterested or disengaged data subjects/citizens. Recognising these impediments to the repositioning of data subjects as truly central in the data ecosystem, the model seeks to overcome these hurdles with, on the one hand, inducements to participation, and, on the other, by encouraging powerful data holders to realise what are the benefits of negotiating mutual interests for themselves (Findlay, Seah & Wong, 2023, chap. 3; Alfredson & Cungu, 2008); Verhulst, 2023).[144]

The role of trust and respectful data relationships

Any governance approach that relies on inclusion and participation, starting out in data-use environments where data-related power between stakeholders is deeply asymmetrical, will benefit from external assistance for trust relationships to grow and to enable more organic and mutually beneficial interests to be recognised and valued. In smart cities it is

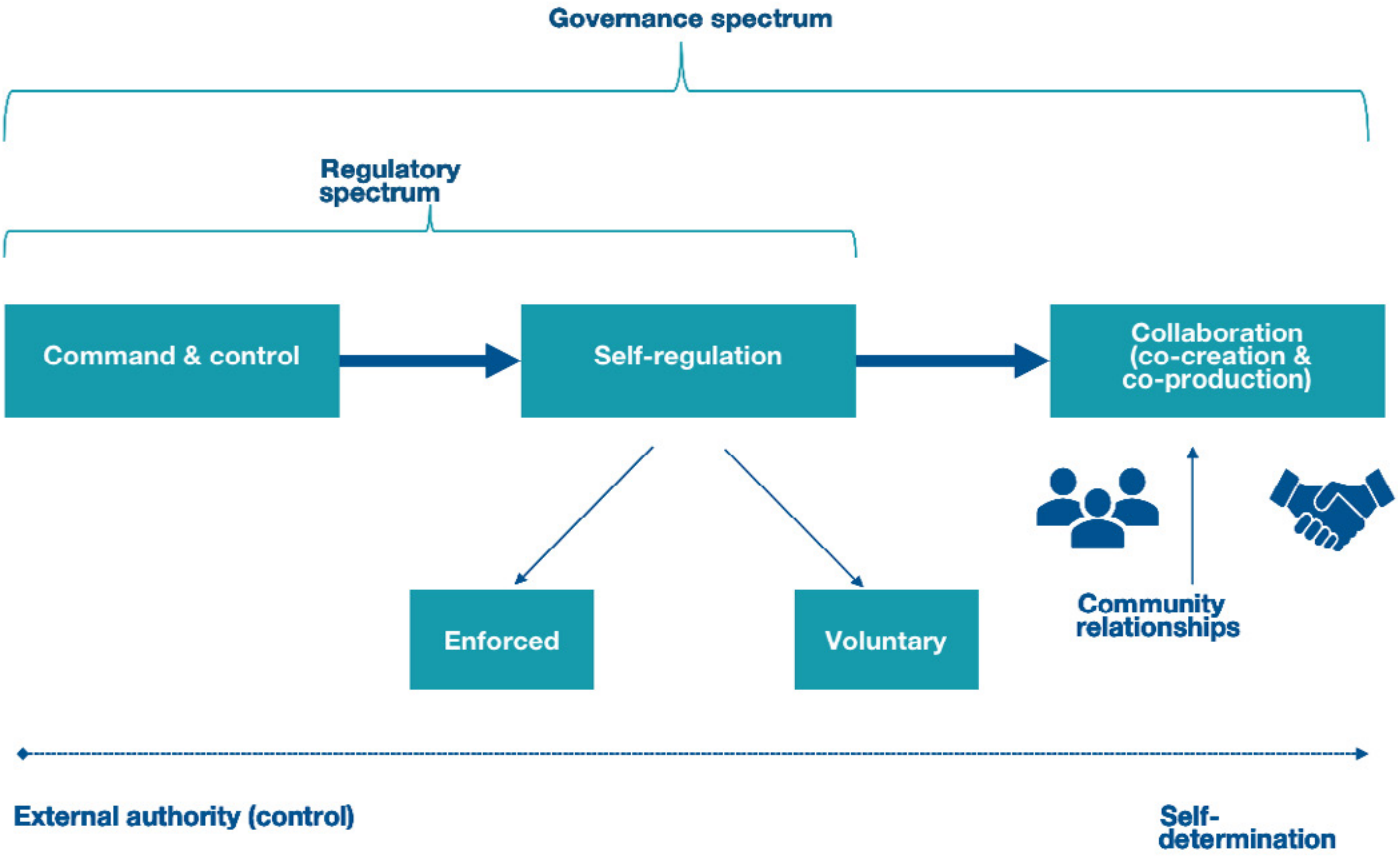
[143] For an explanation of citizen-centricity, see p. 17 ff.
[144] Mutuality is a key aspiration for the theories of digital self-determination. Grounded in principled negotiation theory, it operates based on the recognition that stakeholders will initially approach negotiations over control and access from different perspectives, but will eventually move to recognising the subjective and collective benefits of compromise towards mutuality. This happens once they realize that they will be able to satisfy their respective interests better through a process of finding a mutually agreeable solution than unilaterally.

the state and the private corporations that hold the data-related power, and therefore it is necessary for these players to divest some of that power to data subjects in order to achieve sustainable mutuality of stakeholder interests. While it is often assumed that power is inevitably an exponential enterprise, lessons learned from open finance challenge that assumption (De Pascalis, 2022). By the same token, it ought to be possible to motivate data-rich stakeholders to assist in repositioning data subjects when the consequences of so doing benefit their own interests as well as those of the data subject, which is the essence of mutuality. Expanding the view of data beyond a market commodity on the one hand, and the province of state authority on the other (i.e., moving the narrative in directions proposed by scholars rooted in postcolonial and feminist theory, such as Klein & D'Ignazio, Kovacs & Jain, and Couldry & Mejias),[145] is helpful in this endeavour. Data subjects regularly create and communicate data with no commercial intention, and its value consists in the message it conveys to its intended recipient communities rather than its market profitability.

Stakeholder motivations

It is expected that many public and private actors will see the benefit in better data subject inclusion, whether it be through improved legitimacy, opportunity to lead in corporate social responsibility, or greater access to higher quality data when access pathways are opened and strengthened through data subject confidence and confirmation.[146] That said, a minority of public and private players will need more stimulus to divest power and engage in an inclusive governance project, as anticipated in John Braithwaite's work on enforced self-regulation (Braithwaite, 1982). Braithwaite works on the understanding that the majority of private sector players are willing to engage in responsible corporate practice, provided the terms of market participation are clear and the responsibilities of different stakeholders are certain (only a small minority of private players require regulatory strategies that have a carrot and stick dimension). Braithwaite's enforced self-regulation requirements could provide the initial external validation for wider inclusion and participation, and therefore facilitate the power dispersal pursued by the model. In such a reading, external requirements to endorse initial compliance by the data-powerful may be the 'command and control' component referred to in the governance (regulatory) spectrum from enforced self-regulation to command and control.[147] The overall projected progress of the model is from command and control onto self-regulation (from enforced to voluntary), and then hopefully towards the emergence of full collaboration (via co-creation and co-production), assuming that all stakeholders come to the realisation that there is mutual benefit in integrating and working together without requiring external enforcement.

Projected progression of the model



[145] See Part A.
[146] This could offer an improved mechanism of ensuring data quality compared to, for example, prohibiting users of an application which relies on tracking their activities from sharing their individual accounts with others (cf. LumiHealth, 2023b, para 7.c.i and Snapshot C more generally).
[147] Enforced self-regulation anticipates that there will be no need for external intervention unless that integration and common purpose cannot be achieved. Conflict resolution is designed to get things back on the rails if tensions emerge, and before it is necessary to revert to external regulation.

The four unique features of the model

As a supplementary governance regime, the model has 4 unique features:

1. *Co-creation/co-production.* The model is unlike other governance styles or strategies in that it vests ownership over the regulatory project (and each of its active components, also referred to as mechanisms) in those who are the most vulnerable in the data ecosystem. Even so, data-powerful players are invited to participate in the governance project and shape its mutual benefits.

   If stakeholders with different levels of data-related power are to come together in a shared regulatory project, it is necessary that they should jointly negotiate at the outset to determine what is to be governed, when, why and how. Genuinely shared ownership of the regulatory project from the get-go will ensure better 'buy in' at the implementation and operational stages. Participation in achieving agreed regulatory outcomes is also difficult to stimulate and maintain, if trust is not created through a divesting of data-related power. Co-production will offer various opportunities for participation to achieve individualised interests for mutual benefit. Mutuality can mean that all three key sectors of smart city life (civil society,[148] public administration, private corporations) partake in both the duties to better govern data sharing, and the advantages that can flow from respectful/responsible data management and access.

2. *Self*-determination. Self-determination relies on understanding the 'self' as meriting empowerment through collaboration in responsible data use.[149] To be sure, data subjects and their communities need more than recognition. In view of their fundamental role in the data economy, they need an independent and central place in determining the use of their data.

   Considering the difficulties in recalibrating policy priorities toward citizen-centricity referred to in Parts A and B, the model works from positioning the data subject and their communities first in achieving data autonomy. Autonomy in this sense is not to be understood as egoist individualism which could seemingly run counter to the preference for collective good in many Asian cultures (in fact, see Chang, 2022),[150] and that will eventually lead to contestation with stakeholders holding greater data-related power. Instead, it refers to the genuine capacity to make decisions about how data is controlled, managed and accessed, and can be exercised by any and each of the participant interests (civil, public or private). In this regard, the model advances concerns for the data subject's autonomy as a way of addressing 'data disability', understood as a lack of agency and control over decisions about access and use.

3. *Safe (trusted) data spaces.* It is not enough to ensure data subject autonomy, however. Data subjects and their communities need respectful and responsible spaces to engage with negotiating and transacting their data. Trust relationships between stakeholders are essential if data management and access is to be negotiated across pre-existing data-related power asymmetries, and if vulnerable players are to feel encouraged that their inclusion will be taken seriously. There is much contemporary discussion about the need to create safe data spaces wherein these trust bonds can flourish (see generally Otto, ten Hompel & Wrobel, 2022, esp. the chapter by Huber, Wessel, Brost & Menz; see also Findlay, Seah & Wong, 2023; p. 91). This model takes the notion of safe data spaces, understood as trusted social relationships[151] rather than temporal or spatial 'places', and gives it specific location through the orbits and snapshots that have been discussed in Part B.

   Part B of the Report has identified some common challenges posed by mass data sharing, and context-specific issues that vary across demographics and localities. A safe data space may be unique in the social space it covers, but also in the social experiences and needs it addresses, particularly of vulnerable participants living in each space. Safety could be measured in different ways, as might also be the case with whether a space is distinct or porous, real-life or virtual (Li & Lalani, 2022; Li & White, 2023; Muhammad Ashraf, 2023). The model and its applications will need to accommodate both universal and relative understandings of what makes a data space, and what makes it safe. Additionally, the form and formulation of 'data' can vary across time and space. Thus, different appreciations of 'data' may need negotiation among stakeholders if data governance is to be achieved in ways that recognise the particularities of data as it is created, used and shared.

[148] Of course, we acknowledge that the civil society landscape varies across the region, and so do their roles and levels of societal influence. Like any governance project, the model operates with both universal principles and contextually specific embedding, and therefore, there might be significant differences in how the model may be operationalized in different jurisdictions.
[149] Digital self-determination is another data governance model that emphasises the importance of empowering data subjects in the management and control of their data, so that they can have a meaningful engagement with public and private sector data accessors and sharers (i.e., stakeholders who tend to be more powerful in data ecosystems).
[150] As explained by Ya Lan Chang, the view that Asian communitarianism places a priori the common good above that of the individual, including his or her rights and interests, is a misconception (Chang, 2022).
[151] Durkheim's theorising of social relations and social spaces is explained in Findlay (2020).

4. *Enhancing responsible data access*. The motivation for data-powerful players to engage in a governance project that needs them to divest power is the promise of greater data access through responsible and respectful relationships with data subjects. Recently, global regulatory discourse has directed attention to 'responsible AI', assuming that risk minimisation will be achieved if technology is safe, robust and accountably operated. This discussion fails to appreciate that AI currently cannot operate without data produced by humans (Shumailov et al., 2023). Therefore, responsible AI is as much dependent on responsible data use as it is on safe technology.

   As mentioned earlier, the model cannot rely on attempts to contain responsible access in order to limit the volume of data sharing, or even worse, to pretend protection and then merely enable selective access to benefit one sector on the basis of superior market preferences. Smart cities need sophisticated digital applications and a wealth of accessible data. Citizen/residents largely benefit from the cultures of convenience which digitized urban environments can offer. Rather than seeking to ensure social good or public benefit through the governance of mass data sharing, the research project emphasises the empowerment of data subjects in the realm of increasingly pluralistic and datafied public services. With a view to that purpose, the model advocates responsible and respectful data management, which will ensure greater data integrity and open up opportunities for responsible access and sharing in environments where data subjects have a say.

Grounding the model: the essential connection between co-creation/co-production and inclusive, participatory and collective regulation

The model is intended as an governance enterprise that holds promise for every type of smart city stakeholder, understood as representatives of the citizens, administration and business (Drapalova & Wegrich, 2022). The model is not built around notions of compulsion or sanction, but instead looks at regulation as a collaborative project. The model envisions that the proposed governance processes will progress through two dimensions that are interconnected, and to some extent interdependent. These dimensions move across a governance spectrum (from *command and control* to *collaboration/self-determination*) based on an expectation that while some more data-powerful potential stakeholders initially may require more compulsory motivations to participate, the mutualised benefits of the model will later become an incentive to progress to a more collaborative endeavour where data-related power is dispersed. Along with an a growing understanding of common interests in data management and access, choices about data control become meaningful, because they are informed, they emerge in safe spaces and they are shared in trusted relationships. At the risk of oversimplification, these dimensions can be expressed as

(1) Command-and-control, which as mentioned above may require some external motivation to convince the data-powerful and the data-vulnerable to trust the possibility of mutualised benefit, and
(2) Co-creation/co-production, consisting of three component mechanisms:
    a) Awareness and consciousness building;
    b) Strategy forming, and
    c) Governance implementation.

These dimensions of thinking about and experiencing how a dynamic governance model is activated and how it grows require consideration at all points in the model's staging.

The Model

The foundation stage

Information openness as key to empowerment, and expected hurdles to openness

At present, the processes of mass data sharing in smart cities rely on the selective creation and use of information infrastructures. Such information infrastructures, and the pathways they operate with, concentrate and polarise power in the hands of public sector and private market players, while regulators and vulnerable governance beneficiaries are largely excluded from influencing these information pathways. In the initial stage of the governance model, it is important to avoid the impression that the data-powerful participants (including external regulators) are yet again attempting to 'capture' the ill-informed and data-poor stakeholders. Effective engagement with disempowered market players in participatory self-regulation calls for a genuine commitment to genuine power dispersal.

In light of information deficits and market power asymmetries as manifested in the snapshots and examples analysed in Part B,[152] a key first step in the regulatory enterprise is to provide the less powerful stakeholder with advance access to essential, comprehensive, specific and reliable information managed by more powerful stakeholders in the data ecosystem. With regard to the specific examples of data sharing explored in Part B, this could involve making sure that users of the platform understand the possible and unknown consequences of having one's data shared with advertisers, such as in Snapshot E (Ruangguru), and that users of trackers such as those required to participate in the LumiHealth

[152] See especially snapshots F (DELIMa) and C (LumiHealth) for a discussion of these challenges.
[153] For example, see Snapshot A.

programme (Snapshot C) understand that while their information may not be sold to advertisers by the programme, wearing the tracker day and night will still result in creating a lot of biometric data about the users, which they themselves may inadvertently agree to share e.g. by syncing them with third-party apps, or which may be extracted by malicious actors in a data breach (Landi, 2021). While information deficits may have different consequences across various contexts such as healthcare, transport or education, it is fundamental to reveal information concerning data sharing and the uses of data in a manner capable of repositioning the data subjects/citizens and their communities alongside the stronger market players. Raising the market presence of otherwise vulnerable data market players will enable them, if they choose, to better contribute to inclusive participatory self-regulation, turning secretive and combative data protection posturing into more open data sharing as a precursor to mutualised regulatory responsibility.

Ignorance or confusion about the nature, purpose, and processes of data sharing is common among less powerful regulatory participants. Opening up knowledge about data access in a way that encourages shared participation and trust in meaningful inclusion (not just for the sake of appearances) requires creating an information infrastructure in smart city data sharing which flattens structural imbalances by encouraging bottom-up data management models (Delacroix & Lawrence, 2019). As an example, when super app service platforms intend to share data across different service agencies, users are usually offered a consent option governing such sharing, but the specific design of consent will usually be decided by the in a top-down manner by the more powerful stakeholder(s).[153] Obviously, this is more than a market structure issue. As a pre-condition in its development, the information access technology and pathways built into the governance model need to reflect a more equitable and user-driven format.[154]

It is anticipated that in the first stage of the governance model, there will be technical and operational impediments to information openness:

- Locating and identifying automatically produced data on the primary originators of the data, and as such, regulatory recipients;[155]
- Respecting data privacy, if the data is not anonymous when returned to data originators as information about access and use or not aggregated in bulk;
- Introducing AI-assisted technologies to notify regulatory recipients of data production, storage and use;
- Creating convenient paths of open access which recognize and respect commercially sensitive data that may attach to automatically produced personal data (anticipating that commercial users of original personal data will argue that through their processing of the data they have added commercial value to it);
- Ensuring internal privacy protections, that may have been created by rights-based governance modes, covering the identity of data subjects;
- Educating regulatory recipients on the use and utility of AI-assisted information technologies and their data pathways;
- Enabling regulatory recipients with simple tools to analyse the significance of automatically produced personal data; and
- To activate and enable an 'honest broker' third party/agency to ensure that conditions covering access are complied with, in the spirit of enforced self-regulation.

Incentives to mutuality and resolving data ownership issues

To achieve the transition of power tied to information asymmetries on data sharing, the governance model initially will progress along a governance spectrum from command and control to enforced self-regulation, and then finally collaboration based on appreciating the common benefits that co-production can achieve.[156] Put more directly in reference to the challenges identified in Part B,[157] the model initially will require some external stimulus for both public and private data holders to join the governance project and divest data-related power, so that a climate of mutualised interests can grow, focusing on responsible and respectful data access and sharing. This external stimulus could take the form of best practice commitments for private sector players, and directives as part of smart city administrative policy for the public sector.

Arguments about data ownership and who bears the responsibility and cost for establishing a more inclusive access and information framework will equally need to be settled at the 'command and control' stage of the governance model, between a relevant state agency and the private sector players in the data-sharing enterprise. Ideally, data

[154] An initial challenge lies in the current market reality, in which private and public service providers claim ownership of the personal data automatically produced through commercial and surveillance technologies, and will resist any possibility that its value as a market commodity may be reduced through more open access. Such data ownership determinations must not be a prohibitive pre-condition to the governance project.
[155] This term refers to stakeholders/data market players who participate in and are affected by mass data sharing. Priority is given to data subjects and their communities consistent with the model's commitment to power sharing.
[156] For a discussion of such a governance spectrum see Findlay (2013).
[157] For example, as noted in Part B, Malaysia, Indonesia and Singapore score high on Hofstede's power distance index (see fn 122 and corresponding text).
[158] An adequate description and prediction of the role of data trusts (or data stewards) in achieving comity among the data-powerful players goes beyond the scope of this work, but there is a growing body of literature on the role of data trusts in ensuring data-subject centric data governance more generally. See e.g. Delacroix & Lawrence (2019).

subjects (citizen/residents and their communities) should also be on notice of these points of contention and how they are being managed, but this may be difficult if at that point the data subjects are yet to be presented with an active role for their inclusion and participation in data management and control. These efforts to resolve contention will rely on processes of arbitration, so that fundamental disputes have access to an orderly and legitimate resolution at the hands of an honest third party. Experience from the operation of 'data trusts' no doubt would be helpful in these negotiations, where third party intervention adds legitimacy to resolutions in dispute.[158]

The governance model will need to counter the current market reality that ungoverned algorithmic intervention (such as using personal data for on-selling advertising without notice to the data subject) adds cash in the pockets of the data-powerful players who govern information about data and its flow. The alternative to approaching data as an exclusive commodity is a more universal and widespread recognition of the need to protect automatically produced personal information from market abuse and data subject discrimination. Once this message becomes a fundamental moral principle for shared data practice, other stakeholders in the data ecosystem will require designated locations and duties in an informed and inclusive decision-making interface.

Culture of exclusivity

As mass data sharing between private and public players in the delivery of essential services becomes the norm, so too does a culture of exclusivity linked to the absence of alternatives to access fundamental urban services made available via one app or platform. Once this culture takes hold, users are being denied a place at the decision-making table when data use and re-use is being determined. Users equally tend to be bypassed from information pathways and instead are resigned to accepting assurances from powerful data players that their interests are being recognised and respected.

As more and more users are being compelled to facilitate data mining that enables a whole range of ancillary surveillance potentials, the datafication driving surveillance systems across the city exacerbates information asymmetries and power dysfunctions. It is only through a governance model that is committed to 'feeding such data back to users, enabling them to orient themselves in the world' (Kennedy, Poell & Van Dijck, 2015) that data integrity, in terms of the data's original purposes, can be ensured.

Our conversation methodology consistently highlighted the risk of regulatory elitism in an ecosystem where data subjects/citizens and their communities are also hostage to technology and software that is introduced to their life-worlds as part of techno-colonialism.[159] To complement a move away from regulatory exclusivity and towards participant inclusion through information access, important external players such as government agencies and private sector business confederations must ensure that citizen/residents and their communities are institutionally included in the regulatory process; both in its crafting and implementation. These are a *priori* external market requirements if information access is to contribute to regulatory empowerment, as this model assumes. In addition, inclusion of data subjects in decision-making about data sharing should be considered as evidence of good administrative practice and corporate social responsibility (in parallel to motivations of self-interest envisioned in the model).

The foundation stage of the governance model, therefore, needs to reflect:

a) Access for data-powerless stakeholders to essential information managed and manipulated by data harvesters through the process of mass data sharing;
b) Incentives (via increased responsible access) for those currently monetising secondary data to participate in and contribute towards inclusive participatory self-regulation by enabling the data subjects to have meaningful choice;
c) Motivation (through responsible innovation) to move away from secretive and combative data protection posturing into more open data sharing as the precursor to regulatory responsibility; and to ensure that
d) The governance project progresses along a continuum (spectrum) — from command and control to enforced self-regulation — as the benefits of the latter become clear to those who may initially oppose regulatory openness.

The operational stage

While citizen/residents are central to this governance project, in many contexts analysed in this research they have been revealed to be:

[158] An adequate description and prediction of the role of data trusts (or data stewards) in achieving comity among the data-powerful players goes beyond the scope of this work, but there is a growing body of literature on the role of data trusts in ensuring data-subject centric data governance more generally. See e.g. Delacroix & Lawrence (2019).
[159] For example, as noted by the Indian Participant during the research roundtables, AI-based technology designed, trained and developed in the Global North countries typically fails to include all the relevant socioeconomic considerations in its design. Therefore, these systems tend to glitch, and create a lot of bias when deployed in the Global South. For example, in India, there has been a lot of caste discrimination based on people's names because of using technology developed in the Western hemisphere.

- Limited in their understanding of the data being shared or how it is shared;
- Not aware of the significance of this data, or how they can manage or govern its value;
- Less digitally literate than the stakeholders actively engaged in the public/private data sharing;
- Not engaged with the governance of this data, or well informed about the prospects of participation and any opportunities for conflict resolution, and
- Likely, even with more information about mass data sharing and its impact, not confident that their inclusion and participation in a governance strategy would make an appreciable difference to the quality of their urban life experience.

Therefore, a governance strategy that is designed to address these initial and prevailing disconnections should be explained and expressed in terms and a language that the most vulnerable stakeholders can relate to and understand.[160] It must have a simple structure. In keeping with the collaborative intention of the model, the duties and benefits of the three major regulatory recipients, i.e. the data subjects (citizen/residents) and their communities; public administrators, and private market service providers, should be clearly and openly expressed from the outset. It is the intention of the model that these three regulatory recipients should take 'ownership' of the specification of the model to their particular interests, concerns and contextual needs.

In this stage, the model works on the following collaborative commitments:

1. Co-creation through negotiating context-specific dialogue about the challenges posed by mass public/private data sharing, accepting the basic premise that *the benefits offered through smart city data platforms*[161] *should not come at the cost of data subject control over the integrity of their data; and*
2. Co-production of governance agendas that involve the creation and maintenance of responsible and respectful data relationships between the data subjects and the public/private data harvesters and sharers. Through these trusted arrangements in safe data spaces *where mutual benefits from responsible data access can be negotiated and operationalised, mass data sharing can provide collaborative advantage for all stakeholders in the data ecosystem, whether data is valued as a commodity, or a personal communication pathway.* Data integrity does not have to come at the expense of market profit, provided that the latter is produced through respectful engagement with data subjects.

Co-creation is the overarching commitment, and must therefore be in evidence from the model's inception. Co-creation is also the manner in which inclusion is verified – whether data subjects have had having a voice from the model's outset. Co-production describes how co-creation is materialised into a governance strategy which is collaborative, inclusive and participatory.

Each of the model's three components/operational mechanisms[162] (i.e., 1. awareness and consciousness building; 2. strategy forming, and 3. governance implementation) involves a conflict resolution 'pressure valve' to enable parties to express concerns over trust and mutuality as soon as these arise, so as not to threaten these crucial factors in the model's success.

Conflict resolution approach: conflict empowerment[163]

Smart city urban development planning relies on expansive surveillance technologies and the mass sharing of consequent data sources. While the operational benefits resulting from both these streams include integrated service provision and improved predictive planning, the impact on individual citizens and communities is significant and too often under-recognised.

When these citizen impacts are identified by planners, the tendency may be to:

- Build firewalls around certain data, thereby restricting its availability for mass sharing based on privacy concerns or other specific data protection principles, or
- Mobilise 'trust' and re-assurance through campaigns encouraging platform use designed to emphasise the positive

[160] For example, in areas affected by digital and/or actual literacy (such as India or Indonesia), information available online or even in writing (hardcopy) cannot be thought of as sufficient to correct information (and therefore power) asymmetries, even if that would ostensibly meet the legal requirements. As an alternative approach to information openness, the Indian Participant in the research roundtables suggested the use of model cards.

[161] While Part B was primarily interested in mass data sharing through the amalgam of public sector service delivery agendas and private sector information platforms, there are other examples of citizen/resident data sharing across a range of fundamental urban infrastructures that may also be suitable to the application of the model.

[162] These operational mechanisms are components of the model at this stage, but are referred to as 'mechanisms' to emphasise their action-orientation. Mechanisms and components might be used interchangeably in the description of the model.

[163] It is anticipated (as with any governance model reliant on trust) that when trust and mutuality is strained or challenged, the stakeholders need an immediate opportunity to identify, confront and if possible, resolve issues that endanger trust. Stakeholders should have the opportunity to activate conflict resolution conversations at any point where trust is at risk.

collective advantages of moderated data sharing against any negative impacts to data protection, and creating third-party agencies to 'broker' data use and enable citizens to have a sense of independent oversight (such as limited personal data protection institutions or consumer complaints agencies).

Designed as they are to minimize the negative 'fall-out' from mass data sharing and surveillance, neither of these approaches works from fundamentals of individual human dignity and citizen empowerment. Equally, the latter two values are usually missing from regulatory initiatives where smart cities are concerned;[164] even those that appear people-centric, are usually driven by the need to ensure efficiency, profitability, coverage, time-management and other operational priorities. To address this lack, the 'Conflict/Empowerment Approach' to regulation works from a central focus of human dignity and citizen inclusion.

What is a conflict/empowerment model of regulation? The model grows from general conflict theory assumptions that society, cities and communities are structured around frameworks and relationships of power imbalance (Hård, 1993). Because of the technologized nature of smart city urban spaces, and the information deficits they present for their inhabitants not reached by the small elite of urban planners, this power disparity is exacerbated through apprehension, misunderstanding and disaffection, creating structures of obligation and dependency which will eventually generate conflict as power shifts, and information is centralized.

To counter the power asymmetries between stakeholders, co-creation and co-production should also be at the heart of conflict resolution. No stakeholder should have the dominant role in identifying or resolving conflict, beyond the obvious responsibility of the data-powerful to make information available so that conflict may be fully understood by all parties. While conflict resolution can be viewed as a 'mechanism', we advance it here as an opportunity (in terms of structures and communication pathways) to identify where trust relationships may become strained within the operational mechanisms, and require addressing and recreating through conflict resolution processes that act as a pressure valve for maintaining trust throughout the process. As such, conflict resolution should be available in all operational mechanisms of the model.

It is important, in the spirit of contextual respect for different capacities of comprehension revealed in the challenges analysis, that there should be regular opportunities for raising concerns by data subjects/citizens and their communities, and having these negotiated progressively. Therefore, dispute resolution frameworks that are created to balance power asymmetries should be built into the governance model at the earliest opportunity. Only then will inclusive self-regulation grow to its potential for market power dispersal. In the proposed model, stakeholders will determine the nature, form and frequency of dispute resolution opportunities that best suit the governance challenges they are addressing, and consistent with the power dispersal pathways they operate with.

If power differentials in smart cities are structural, and the conflicts they will produce as individual citizens and communities experience social exclusion through technological oversight are inevitable, then the question for regulators is how to address or manage these instances and sites of conflict. A conventional state-sponsored or market-manipulated approach is to minimize conflict and neutralize its dynamics and effects as it occurs (Flyvberg, 2001, p. 108). A 'trust overlay' approach is one soft regulatory intervention for conflict minimization.

In a conflict/empowerment approach, it is imperative not to divert ownership, explanation, management and resolution of conflict away from citizens and communities. Paternalist regulation uses conflict minimization as another way of disempowering the citizen, by taking control of the conflict and of its evolution. In paternalist terms the citizen is a data object and as such needs to be provided with institutional arrangements and processes for conflict identification and resolution. In so doing, the regulatory intervention does not deal with the source of conflict from the individual citizen perspective which may be rooted in apprehensions about surveillance, technology and data usage (Christie, 1977).

If conflict identification and resolution is primarily returned to individual citizens and their communities, conflicts can actually become a positive force and at the same time an extremely useful evaluator of the impact of smart city development on urban users without a voice in decision-making linked to planning.

The approach requires:

- Contextual identification of the urban space conditions that may lead to conflict before conflict emerges. This requires continual, informed and responsive conversations between citizens (particularly vulnerable citizen groups) and public sector (urban) administrators about specific planning technologies, operations and placement. These conversations need to be informed by earlier conflict sites and the dynamics of their evolution. At this pre-emptive stage potential conflicts can be mediated, and conflict-generating technological developments can be moderated through building true trust relationships. Citizens need to be provided with the information they believe may

[164] This observation is not to deny the desire for social good as a secondary consequence of efficient urban development. However, paternalist approaches to engaging with citizens' concerns about the changing conditions of their lived world do not always produce resilient and lasting outcomes.

ameliorate power dependencies and social exclusion, on which conflict feeds. Information sharing at this stage is imperative if fear and perceived risk are to be addressed.
- If a conflict is identified, the community affected by it must be given a structured opportunity to manage its development and resolution. The public sector administrator should do no more than facilitate citizen-centred resolution initiatives.
- Most importantly, the 'learning from experience' dimension of conflict as a tool for social bonding as much as a force for disruption needs recognition.[165]

Change as it effects the lived world will always be disruptive. Conflicts will accompany this disruption and as their connection is both organic and mechanical, planners need to utilize conflict as a positive force for social cohesion as suggested by Flyvberg (2001, p. 108). On the organic level, it is important to recognize and reiterate the forces of strong social bonding that will be strained through change.

At the same time, disruption in an economic and technological sense nowadays has positive connotations. If smart city regulation is to benefit a citizen constituency, with human dignity and citizen empowerment at its heart, the city must be seen to serve the citizen when its operations become contested. This is not a process of achieving compromise or risk abatement. Instead, as made clear by the vigorous contestation of Quayside, the Sidewalk Labs controversial, sensor-studded 'Google city' experiment in Toronto, conflict may be a context for empowering both the citizen and the regulator to create a city not so 'smart' that it has lost the citizens it is designed to serve (Teale, 2020). Once a conflict has been predicted and talked through, or identified and resolved via citizen ownership, information sharing, and planning adjustment, the city will be able to readjust to fit better the needs of its citizens.

The model's components/mechanisms

The proposed governance model has three general components/mechanisms: a) Awareness and consciousness building; b) strategy forming, and c) governance implementation. These components/mechanisms require the three principal stakeholder groups to participate and interact in co-creating and co-producing the governance objectives they themselves determine to achieve equitable data management. The first two components/mechanisms emphasize co-creation, while the third focuses on co-production. Different components/mechanisms will require different duties and levels of commitment from regulatory recipient groups. The first and the second will be closer to the *command and control* end of the governance spectrum.

Awareness and consciousness building

Earlier in this Part there has been a discussion of why a*wareness raising and consciousness building is necessary for data subjects and their communities (citizen/residents)*. Because in the early stages of the governance project trusted relationships and appreciation of mutual benefits will not have yet emerged, a targeted campaign launched and run by a third party which the given stakeholder group aligns with (such as civil society organisations) may help in achieving awareness. The campaign should emphasise the need for and benefit from integration and participation in the governance process. The motivation for active inclusion should be the offer of genuine participation in the management and control of personal data, as well as its valuing. For the private sector service providers, industry groupings and representatives with an interest in good governance and best practices could take on the campaigning for this stakeholder group. The motivation for this group in particular should be a genuine potential for greater data access and higher quality data achieved through user confirmation. As for the public administration stakeholders, they should already be familiar with and sensitive to co-creation and co-production governance strategies (considering that these increasingly figure in contemporary policies and actions of public organizations as described by Røiseland, 2023, and that co-creation is viewed as a new public administration paradigm in some countries, as per Torfing, Sørensen & Røiseland, 2019), and view the strengthening of the legitimacy of administrative authority as the main motivation for participation.

Strategy forming

Once the three stakeholder groups have understood the reasons behind the model and its potential benefits, it is necessary to create safe (trusted) data spaces in which relationships of trust can emerge and develop, and where discussions about mutualising benefits through power dispersal may commence. Safe data spaces will initially be the responsibility of the data-powerful players who commit to listening to the data subjects' concerns, and at least offering information about data use as requested or otherwise needed to stimulate engagement. This stage will involve respectful negotiations in which the data-powerful offer empowerment through information dissemination concerning data sharing, and opportunities to discuss the nature of data use and its valuing. In return, the data-disempowered need to pledge constructive engagement, active data validation, and to discuss enhanced respectful and responsible data access. If mutual interests/benefits can be agreed, the contextual specifics of the governance strategy (prioritising challenges and conditions for their addressing) can be the foundation of co-creating governance mechanisms and processes.

[165] There is an interesting discussion of how ethics as 'friction points' in the ecosystem can point to areas of strain and conflict in decision-making between ethical compliance and a company's market imperatives in Findlay, Seah & Wong (2023), chap. 3 & 4.

Governance implementation

At this stage, use-cases will be required as test beds to see whether negotiation and mutualising can produce the motivational benefits offered to stakeholders in the other two components/mechanisms. An example can be seen in the open banking movement, where customers were initially offered (by banks and government regulators) qualified information on data use, and sufficient information about financial data to enable data mobility (Remolina, 2019). Use-cases in the health and education orbits from Part B might be sufficiently universal and context-specific to localise the model and to present tangible data sharing experiences, for example in super apps. It is in this component/mechanism that most conflict occasions could be anticipated, highlighting the importance of resorting to the conflict/empowerment approach (as described in the preceding subsection). In order to extrapolate the model to other potential stakeholder groups and data ecosystems, it could be useful to carry out an evaluation of the successes and shortcomings manifested in these use-cases. In addition, if tensions arise in the relationships and expectations between any stakeholder groups, intervention from external mediators like data stewards, or through the creation of a data trust/data cooperatives could be effective.

Activating the model: divesting data-power

The normative underpinning of the model in action is the recognition that significant power imbalances in governance projects are not compatible with equitable participation and inclusion. In order to become involved in this governance project, each and all stakeholders will need sufficient data-related power to do so. The public and private sector players should have sufficient information about data access and use, or at least the opportunity to become aware of these. Data subjects will need empowering through information and inclusion so that they do not feel intimidated or by the prospect of participation, or captured by the more powerful stakeholders.

At the risk of being criticised for 'data dumping' on otherwise unaware or insufficiently informed data subjects, and thereby not altering much the existing power asymmetries over data, the activation of the model through the divesting of data power to citizen/residents will involve two simple yet significant directions. The data accessors, harvesters and users should be required to:

- Inform data subjects when their data has been accessed and used (including situations of initial use, and re-use);
- Explain in simple terms what was the purpose of the data use/re-use.

In this respect the data subject may not need access to all the data that has been used and re-purposed. Information about how and why data is used may suffice to provide the citizen/resident with a greater potential to negotiate control from an informed position.

In the process of activating the model against specific challenges, stakeholders assume responsibility for power dispersal in direct proportion to the extent to which they presently have control over data, and are interested in increased data access. This may sound like a lofty aspiration, but it has some very pragmatic features:

- For data-rich stakeholders, data sharing is a commercial enterprise
- Data, while being ephemeral, needs replenishing, growth and sharing if its market value is to be maintained
- The commercial value of data is impacted by its accuracy and integrity
- Throughout its lifecycle, Big Data moves further away from its source, and as such moves further away from open validation
- Data subjects and their communities are prolific producers of data that is the essence of data marketing
- It is often forgotten that data subjects and their communities are the most accurate measures of data accuracy and integrity
- Much data can be accessed and marketed without data subject knowledge
- *but* the closer data is to the data subject, the higher the quality of that data, and
- Access to better quality data depends on the trust of data subjects and their communities.

To conclude, a comment on how universality and contextual specificity of governance challenges play out with regard to data disempowerment is due. The challenges which are identified in the orbits and snapshots comprising Part B have both contextual and universal features. Indeed, a significant hurdle for any governance strategy intending to incorporate both dimensions of challenges is neither to drift into irreconcilable subjectivity, nor to ignore the importance of context by imposing some overarching, top-down policy that may turn out to favour the traditionally privileged stakeholders in data governance. This model assumes that a contextual commitment to data integrity, and a universal ascription to the normative frame of Rule of Law will enable a balance between subjective and objective regulatory directions. As an example, information deficit may be considered as a contextual (subjective) variable, and information openness as a universal governance requirement. If information openness is not a feature of data relationships, it will be one cause of an information deficit. Depending on the degree of digital literacy, community awareness and civil society action around data use and re-use in any particular context, information deficit can play a crucial role in data disempowerment. The degree of this disempowerment and the extent to which it is recognised as a challenge to data integrity by data

subjects will depend not only on limited data openness (universal challenge), but on the extent to which data subjects understand their data and value participation in its management and control (contextual challenge). Following on from an appreciation of context and universality, real-life applications of the model will advance a priority of citizen-centricity in a governance project where value of data is measured in proportion to its quality and integrity.

Working in parallel with regulatory and governance endeavours focused on safe (trustworthy) data spaces, the following *activation plan* lists the conditions for co-creation/co-production to enforce a vibrant and effective data access and management regime. Each focus of activation should not be viewed as discrete, but rather as working as a part of an inter-operative scheme.

Activation plan for the model: the 7 foci

INCLUSION
It is important to identify the stakeholders who have an interest in the nature of the data over which they may have claims, the relationships between stakeholders and their duties/responsibilities to each other.[166]

EDUCATION
Once stakeholders are identified and included, they need to be informed and educated about their data (or data in which they have an interest) which is held/used/intended to be used by other stakeholders in that space.[167]

ENGAGEMENT
As a self-regulatory strategy that depends on the engagement of stakeholders in open communication and negotiation over data, constant and open communication is essential.[168]

MOTIVATION
Even if the data spaces are safe, inclusive and operate with respectful engagement, this will not guarantee all stakeholder participation. It is important that when parties are considering the benefits of participating, potential stakeholders make clear to each other what they can offer as a consequence of participation.[169]

INTEGRITY
The model envisages a process that enables data integrity and the protection of that integrity as individual data producers control and manage their data.[170]

ACCOUNTABILITY
Governance policy that applies the model will succeed or fail based on whether trusted relationships concerning data use can be established and maintained. Transparency around the storage, provision and use of data is an important factor in establishing and maintaining trust.[171]

SUSTAINMENT
The model alone cannot provide a process for curing all pre-existing data control inequalities. Considering the possible resistance to open data relationships in the minds of some stakeholders, sustainability needs commitment. Communities and markets engaged in the governance model must develop trusted data relations (whether these be commercial or social) that will perpetuate more sustainable market arrangements and social bonds where data is concerned.[172]

---

[166] The identification process can be a communal exercise, but is primarily the responsibility of stakeholders with most power over data in that space, and who wish to use that data for any secondary purpose. Once these are identified, it is necessary to maintain a register of inclusion to ensure that data-subject interests are adequately recognized.

[167] The duty to inform and educate rests with stakeholders who hold/use/intend to use or reuse such data. A log should be kept by the stakeholders on how they have discharged their duty.

[168] If the engagement is either not positive or respectful, the data space is not safe. Engagement, therefore, is crucial to the achievement of respective data interests through data control and management.

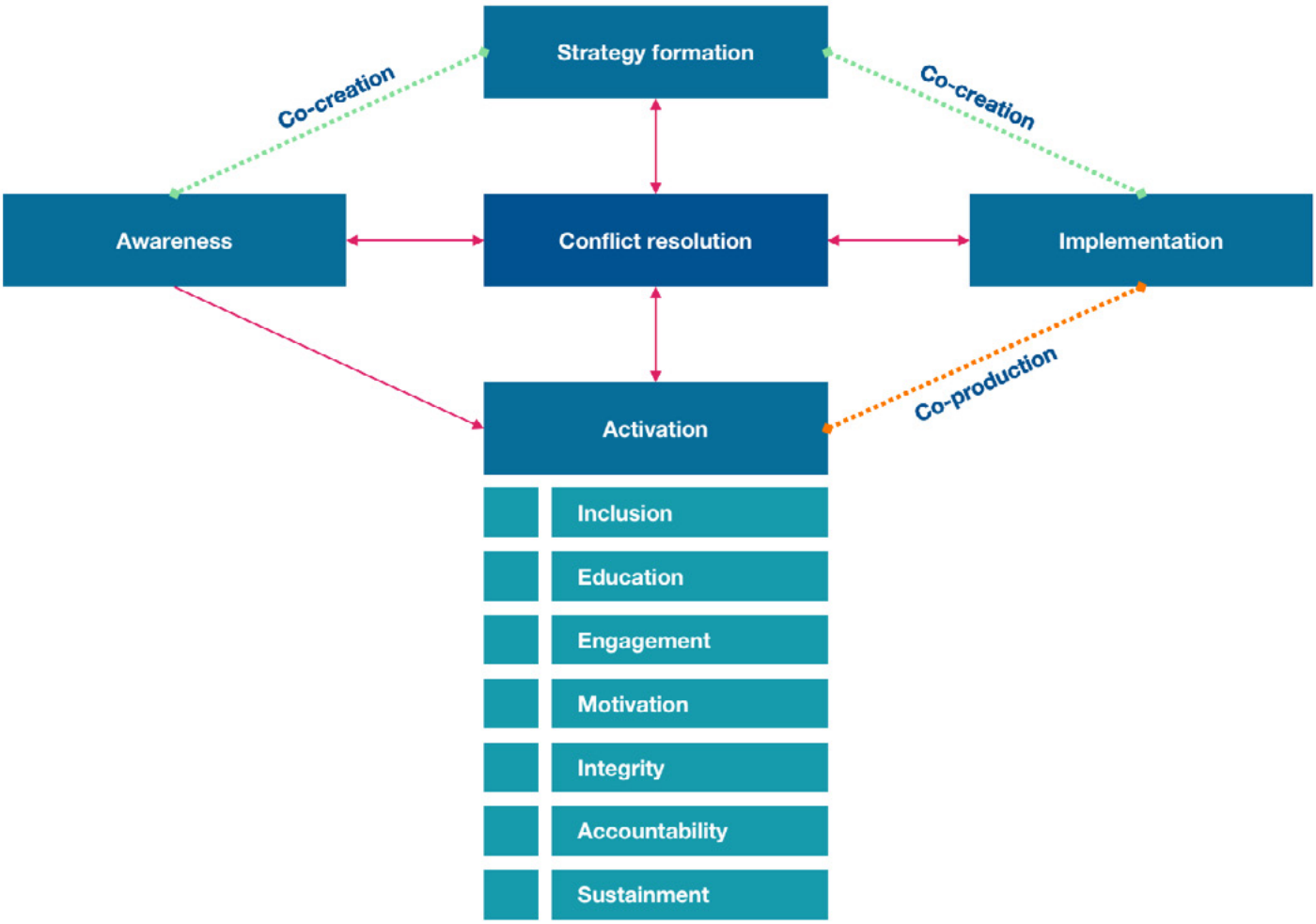
[169] Particularly for vulnerable data-subjects who have up until this point been disempowered, ignored or largely lacking trust in regulatory mechanisms, finding motivation for participation may be challenging. In terms of benefits of participating, the least of what the data storers/sharers can offer to data-subjects is information about the personal data they share and how it has been, or will be used. In turn, if the data subjects are willing to trust the data storers/providers, they may be willing to authenticate their data and open up further access to personal data for agreed uses.

[170] This is a fundamental and prevailing pre-condition for the model. A market consequence of data/data-subject integrity is the potential for subject validation of personal data and thereby an increase in the integrity of data down pathways of access. This consequence of the open storage, provision and use of data not only protects the interests of data subjects in their data, but will improve the efficacy of data as it is then accessed and negotiated.

[171] However, openness alone will not always ensure trust between stakeholders. In fact, unless good data use practices are agreed, openness about data use may damage trust. A lack of openness on problematic data storage, provision and use could endanger the possibility of trust when data is transacted. Trust, once established, must be continually nurtured and confirmed.

[172] Most importantly, the 'learning from experience' dimension of conflict which arises during the evolution of the model as a tool for social bonding as much as a force for disruption needs recognition. Once a conflict has been predicted and talked through, or identified and resolved via information sharing, interest compromise and mutual respect, the relationships essential for sustaining the model will become more trusting.

The Model: a graphic representation



Implications of the Report

Every day, everyone living in a 'smart city' generates data which then is used in the delivery of administrative and commercial services that make the city function. Without such data flows, the advantages of digitized urban environments would not be as they are. However, citizen/residents in smart cities, who are often identified as the central focus of smart city policy, are largely unaware of how, where, when, why, how and who uses and shares their data, and are therefore less likely to have a say in decisions on data access and control.

Following a substantial research project and its detailed reporting, it is useful to reflect on the implications of what was assumed, confirmed and projected. In the next few pages some thoughts are presented that will enable a compact understanding of why the research was carried out, and what it uncovered in the way of challenges for citizen/resident data. The first part of the Report, Part A, set out what needs to be known, i.e. the questions the research work has aimed to address. Are large pools of data being shared between public and private sector agencies in the running of certain smart cities? If so, does this sharing create particular governance challenges if the interests of citizen/residents in the access, use and control of this data are to be recognised and ensured? Answering these research questions has required both contextual and generalized evaluation in terms of good governance expectations and relevant Rule of Law principles. This empirical exercise is undertaken in Part B of the Report. Finally, assuming that it is citizen/resident interests that data sharing should ultimately benefit in terms of their urban life experience, what implications from the challenges and concerns presented in Parts A and B should good governance address? Part C draws down these issues (recognising their contextual and universal sources) and suggests a model in which citizen/residents can take a key role in governing their data when it is shared in their cities. The implications from Part C remain to be activated and evaluated in dealing with how data sharing evolves as a central feature of digitized urban life-experience.

Particularly in this research work, we have endeavoured to make a contribution to the knowledge on the governance challenges involved in the sharing of citizen data between the public and the private sectors by analysing selected orbits of platform-based public service delivery in three SE Asian jurisdictions. This Report will inform research and policy activity in the field of good governance of data sharing in smart cities more generally.

This brief concluding reflection is broken down into three directions:

- Firstly (employing the thinking presented mainly in Part A of the Report), it explains the stakes involved in smart city mass data sharing, shows how the research assumptions about mass data sharing, surveillance and governance challenges could be confirmed through research and provides some broad themes that emphasise why this is an important, and as yet largely underdeveloped field of governance not only for those who contribute and use data to drive smart cities, but also those who want urban life experiences that are inclusive, informed and respectful regarding how data is accessed, used and shared;

- Next (and mainly drawing on Part B), it details the ways in which ideals of citizen/resident centricity in urban development are challenged by mass data sharing, and identifies broad governance concerns arising from data arrangements and relationships linked to public service delivery in different contexts, as well as more common emerging themes;

- Finally (and mainly going back to Part C), it discusses what can be done about insufficiently governed data sharing, i.e., in what ways can governance policy and practices help revitalise citizen/resident interests in the accessing and sharing of their data while ensuring quality of urban life experiences. In particular, its explains what the proposed model says we can do about those governance challenges, and also signals the most important take-away observations that have implications for how we view data as a community resource, not just for commercial and administrative benefit, but also for enhancing individual and collective urban life experience. In particular, through its emphasis on co-creation and co-production, the model advocates governance practices and policies that have citizen/residents involved from the outset.

1. Governing Public-private data sharing

The research was undertaken to better understand how data-based surveillance and mass data sharing in smart cities presents challenges and requires governance. At the outset it became clear that while the interests of citizen/residents were regularly identified as key motivators for smart city expansion, when it came to the data they produced, there was little recognition of their preferences and priorities in determining how such data was shared. Accepting the importance of citizen/resident interests as a starting point (the citizen-centric frame), the research explored many different instances and situations of data sharing, which in turn revealed common governance concerns. From there it speculated on how different interests in accessing, using and managing data were competing, and if these could be drawn together in a governance model based on trusted, respectful and responsible data relationships.

If smart cities are developed for the benefit of citizen/residents, and good data governance should recognize openness in access and use, the question then arises whether it is sufficient for data subjects to trust and rely on the appropriate

practices of big public and private data sharers without having sufficient knowledge of data sharing, or the opportunity to participate actively in its governance. The answer becomes more complicated considering that there are regular instances identified in the report of data access, use and sharing that either proceeds without effective governance commitments across the data ecosystem, or is not fully understood and monitored even by the big data players.

Further, no one doubts that digitized urban infrastructures, and the data on which these rely, produce convenient, efficient, and even beneficial outcomes for citizen/residents. Indeed, some might think as a result that losing data autonomy is a small price to pay for such advantages. Complicating any such simple trade-off, it became apparent in the research that on many occasions data sharing resulted in governance challenges, although it was neither necessary nor essential for the achievement of beneficial urban outcomes for citizen/residents to be factored out of the equation.

This research confirms that as a consequence of the reliance in smart cities on digital infrastructure, data created by citizen/residents is being shared at scale. This affects the life experiences of those citizen/residents not only in terms of privacy and data protection, but also having to trust the often-compulsory platform-based service delivery. Much of this sharing occurs without the direct knowledge or informed consent of the citizen/residents as data subjects. The governance challenges raised by data sharing of this nature are significant and recurrent, the more so because public and private commercial and administrative users access and operate on data that traditionally may have been contained within separated data pools, but which are currently more often than not being merged in urbanized, "smart" city contexts. Consequently, the traditionally disparate public and private governance domains associated with those data pools are coming together, raising concerns as to whether and how the impacts of that sharing on the data subjects (the citizen-residents) will be governed. For instance, do Rule of Law considerations apply beyond public sector data use and transfer into this merging of public and private law governance operations?

If good data governance is to ensure responsible, transparent and accountable data use, then data subjects should not be overlooked when it comes to data protection and integrity. The research reveals major power imbalances between citizen/residents and the large public and private data users when it comes to controlling access and use. Many of the examples forming the contextual understandings of data arrangements and relationships in the Report make it clear that despite producing and contributing much of the data that fuels digital urban commerce and administration, data subjects often lack even the most basic information about how their data is being shared. At the same time, conventional and pre-existing governance regimes that focus on data rights or privacy are not effectively addressing this information deficit when it comes to mass data sharing. That is because such regimes and laws (such as the PDPA laws in Singapore and Malaysia, for example) were not originally designed for mass and repeat data sharing and merging between the private and private sectors. Thus, while personal data protection is appreciated in most of the smart city locations studied as important indicia of responsible data use, the research set out to transcend protective agency approaches directed at individualised personal data use, to explore how and why personal data integrity was being compromised through mass data sharing that lacked openness for citizen/residents. Indeed, many of the governance challenges identified in the Report are not privacy or rights issues, but more fundamental considerations about participation and engagement in how cities function, and who uses data to determine how city life experiences play out.

An additional, but no less important implication for smart city data sharing objectives is related to data integrity and quality. Powerful public and private sector agencies share data for the management of commercial and administrative services on a vast scale. However, without governance practices to better ensure the quality of data as it is reused from the original purpose to many secondary ones, even these more empowered stakeholders regularly risk the economic and social value of such sharing.

2. The challenges to citizen interests from mass data sharing

Our research makes it clear that there are particular governance challenges flowing from public and private data sharing at scale and the consequent merging of governance spheres, values and management styles traditionally associated with the public and the private, but absent the oversight of all stakeholders and interested parties. While some of these challenges will be specific to the context in which data relationships generate and develop, others will be more general in nature and recurrent over various settings, even if they may still manifest in different ways between and even within the analysed jurisdictions. Addressing these challenges in a citizen-centric commitment is a crucial step towards the goals of openness, transparency, accountability and responsibility that underpin good governance of data sharing in smart cities. Part C argues that in a governance model based on a co-creation/co-production endeavour and emphasising data-subject inclusion, the mutual benefits arising for all stakeholders in terms of access and quality will justify some necessary power dispersal across data arrangements.

Challenge 1: transparency

One recurrent challenge identified was consistently little transparency and information openness about the way data is used and shared in smart city programmes and initiatives. Throughout the empirical research, information deficits as far as citizen/residents are concerned were identified across all analysed snapshots and orbits, though of differing degrees. In some contexts and snapshots, even the more powerful stakeholders did not appear to have sufficient information about what user data was being collected, and how it was being used and shared across what is likely a constellation

of actors who could be expected to have little incentive to reveal this information on their own account for governance purposes.[173] More generally, it was objectively difficult for the research team to gather detailed information on the flows and uses of data due to complex layering of data access and use by platforms. By the same token, individual platform users remain unlikely to develop a clear picture of the flows of his or her data on similar platforms in a non-professional capacity. Compounding such obscurity, the research literature on the data sharing in smart cities is still limited. Larger research initiatives could broaden and deepen the scope of enquiry first revealed in this report, especially if they are able to benefit from data openness.

Additional commercial, administrative and cultural factors such as a high distance culture, and prevailing economic considerations, especially where platform apps are being developed without competitive resourcing and consumer choice, may make it unlikely that fundamental and pervasive information deficits would be challenged and addressed spontaneously in relations between the stakeholders. Motivations for data power dispersal across data arrangements, such as market reputational advantage and data evaluation, need greater investigation and development. In emerging economies such as some of those covered in this report, digital, platform-based public services such as education and telemedicine might be the only way for some governments to reach remote, often underserved populations. This growing tech-dependency in fundamental life experience administration emphasizes the need for additional mechanisms to ensure transparency around public and private data sharing based on solid incentives to do so, such as suggested in the governance model.

Challenge 2: accountability

Without the fundamental value that is sufficient information openness, a number of other principles of good governance are undermined or even absent. Where there is insufficient information about data use and sharing, it is difficult to demand accountability for it from those public and private agencies who operate and control the sharing agenda. Many privacy policies linked to platforms described and analysed in the empirical part of this research failed to provide a complete and comprehensive overview of data uses, and tended to be characteristically silent on the matter of who the data would be shared with to ensure the functionality of the platform, let alone for any reuse of that data. Although the typically long and complex privacy policies of platforms directed to the mass user have often been criticized as unhelpful, with some commentators referring to the tactic of overwhelm, and even extortion of user consent for an appearance of legitimacy, this continues to be the prevailing model linked to informational obligations of online platforms in the field of data protection and privacy. Pending further research, this suggests that similar governance challenges might well be identified across orbits of public service delivery outside of the scope of this research, such as transportation, communications and digital finance, for instance.

But beyond privacy policies, one of the snapshots analysed in the Report involved a deficit of fundamental information on who has actual control over citizen data shared with a platform.[174] This raises the question how to ensure accountability for sharing in such cases, where responsibility is masked or obscure. Moreover, accountability for data sharing may be additionally limited by various factors, such as the lack of alternative apps offering a similar public service to the citizen, the presence of particular cultural norms, or low levels of education (both in terms of digital and actual literacy) among its users, which may be found in some jurisdictions. In any case, it needs emphasis that awareness for the sake of empowering the citizen/residents as data subjects is the cornerstone of the model proposed in Part C of the Report, reflecting the belief that an alternative governance strategy must emphasize effective informing for the purpose of ensuring accountability, not avoiding it. The accountability challenges boiled down to determining who is accountable to whom and for what data sharing and its extent. Without basic information informing these questions which would be assisted by openness and a richer research base, the challenges identified remain suggestive, but no less crucial.

Challenge 3: participation

As the platforms which offer the delivery of services rely on the users (citizen-residents) contributing data, it is reasonable to expect that these users should be included in the governance of that data. However, it appears as though once such data has been collected by the platform, it is assumed that the data subject has lost any stake in it and thus he or she is typically excluded from any relevant decision-making. To give one example, one could point to decision-making that involves devising rules on how to attribute points for activity in a tracker-based health app, which can otherwise be reframed as ultimately deciding on how to convert data for monetary awards. In one of the snapshots, app users were resentful when the rewards-for-data rules were changed, which could be linked to a lack of understanding of the rationale and criteria for the modification of the rule.[175] Such disquiet may be worse in areas where users appear to be particularly attracted to the given platform by the promise of monetary rewards and could potentially interfere with a broad uptake of important public policy programmes. The greater availability of information in a co-production- and co-creation-based governance model would help address the information deficit, making citizen/residents better positioned to decide if and to what extent they would like to become involved in the governance of sharing of their data.

[173] For example, see Snapshot F.
[174] See Snapshot B.
[175] See Snapshot C.

Challenge 4: power asymmetry

All snapshots have featured the challenge of power asymmetry between stakeholders, shaped not merely by the embedding of private tech companies in the sphere of platform-based service delivery, but also by the data itself. As a problem by no means limited to these Asian locations, the power of tech companies in the sharing domain will in many cases be virtually unchecked in a data governance sense. These often global mega corporations control the technological infrastructure, the data, and the information about how the data is used and shared. In particular, they decide how and to what extent that information will be shared publicly. By the same token, they influence at least some public perceptions of how the data is used, for example around the controversial issue of what constitutes the sale of user data through applying an unorthodox definition of such a sale.[176] It appears that an alternative governance strategy must rely on the objective of data power dispersal through co-creation and co-production leading to an empowering of data subjects through information and participation. In this vein, the model in Part C again emphasizes the fundamental importance of addressing information deficits, of assisting in the maintaining of trusted relationships between stakeholders, and the easy availability of a conflict resolution mechanism to address stakeholder differences.

Ensuring a more balanced and better informed participation in the governance of shared data in smart cities is consistent with a vision of an inclusive urban development. In some of the jurisdictions studied the organic development of such participation is potentially impeded when the public sector enters into preferred and often singular arrangements with big tech providers, enshrined in private law contracting. The implications of this market narrowing in the provision of platform-based services go beyond more efficient and equitable access to these services. If private law contracting is not counter-balanced by governance obligations that ensure stakeholder participation in the widest sense, it may result in the loss of transparency, accountability and inclusion.

3. Why is governance going to help?

Currently smart city data sharing is proceeding either ungoverned, or inadequately governed/regulated by strategies typically designed for a more individualist and static approach to data arrangements. Likewise, the analysed jurisdictions tend to over-rely on public/private compliance without the informed and actual involvement of data subjects. Yet, an integrated and representative governance strategy could offer opportunities to solve these challenges. Part C has balanced the need for specific direction in governance policy and practice against a commitment to these stakeholders as owners in the data sharing ecosystem.

'A model' format in Part C was chosen so as not to stand in the way of the data stakeholders determining the essential constituents of shared data governance policy, while at the same time emphasizing and giving support to crafting data relationships that are fit-for-purpose in addressing context-specific challenges. This recognition of contextual determination is not in any way meant to undermine the recurrent considerations which define good governance against the challenges enunciated in Part B. For instance, the model demands that the power asymmetries at the heart of all data sharing arrangements be levelled through a commitment by powerful data players to inform and enable citizen/resident inclusion and participation from the outset. Along with achieving power dispersal, all stakeholders need to commit to establishing and maintaining trusted data arrangements within safe data spaces. While these aspirations might appear as a tall order (relating back to the thinking in Part A and the problems revealed in Part B), it is possible to achieve the mutualising of interests among stakeholders by ongoing negotiations within respectful and responsible data sharing environments.

Trust in data relationships and safety within and across the spaces in which data is accessed, shared and controlled will have profound implications for the evolution of smart city data use in a way that benefits stakeholder interests and perpetuates use/control arrangements encouraging getting better data. If trust is not a central plank in a respectful and responsible data governance model, the data subjects will become alienated and their engagement inhibited through information deficits. This in turn will have a negative impact on data access, quality and the integrity of data sharing infrastructure. The Report implies how a proliferation of data sharing platforms relies in no small part on the reputation these platforms generate. It would be difficult to envisage, even when other options are foreclosed, how platform use would be positively approached by citizen/residents, if trust in the platform was based on nothing more than ignorance. In these circumstances any eventual data breach could bring the platform into disrepute and make data access and sharing impossible to maintain. If trust is missing, then the implications for the sustainability of commercial and administrative data sharing ecosystems are at risk.

[176] See Snapshot F, where Google was one of the platform's key partners.

The model relies on citizen/resident participation. Participatory governance strategies have been plagued either by apathy, or ignorance as to the necessity and value of such participation. Worse still, some governance policies may inadvertently convince the citizen/resident of the futility of their engagement as data subjects by displaying only a token commitment to such engagement (falling short of co-production). Mindful of these hurdles, the model advocates community education as to the importance and benefits of meaningfully addressing citizen interests, and the reasons why powerful stakeholders may be inclined to relinquish some data control (i.e., mainly to achieve quality data access and reputational trust). Additionally, the model injects conflict resolution 'pressure valves' at junctures where trusted data relationships and data spaces may come under strain.

The implications of either not governing mass data sharing specifically and appropriately, or trying to do so without prioritizing citizen/resident interests should be an urgent concern for future research. Part C speculates that the loss of real opportunities for generating trusted, respectful and responsible data arrangements, coupled with an increased likelihood of data breaches and the dampening of data-subject confidence are matters for wider economic and social concern.

Bibliography

AEPD & EDPS (2021, 27 Apr). 10 misunderstandings related to anonymisation. European Data Protection Supervisor. https://edps.europa.eu/data-protection/our-work/publications/papers/aepd-edps-joint-paper-10-misunderstandings-related_en

Alban, D., Eynaud, P., Malaurent, J., Richet, J. & Vitari, C. (2019). Information systems management: Governance, urbanization and alignment. Wiley

Alder, S. (2022, 24 Oct). Meta facing scrutiny over use of Meta Pixel tracking code on hospital websites. The HIPAA Journal. https://www.hipaajournal.com/meta-facing-scrutiny-over-use-of-meta-pixel-tracking-code-on-hospital-websites/

Alfredson, T. & Cungu, A. (2008). Negotiation theory and practice: A review of the literature. FAO. https://www.fao.org/3/bq863e/bq863e.pdf

Al-Shahab, F. (2020, 18 Jun). Why does Malaysia's MoE discriminate against local technology and digital content? DNA. https://www.digitalnewsasia.com/insights/why-does-malaysias-moe-discriminate-against-local-technology-and-digital-content

Ang, B., Ho, T. & Jayakumar, S. (2023, 1 Mar). Commentary: The good, bad and unknowns of letting Singapore's civil servants use ChatGPT. Today. https://www.todayonline.com/commentary/commentary-good-bad-and-unknowns-letting-singapores-civil-servants-use-chatgpt-2119016

Apple Newsroom (2020, 16 Sept). Singapore and Apple partner on national health initiative using Apple Watch. Apple. https://www.apple.com/sg/newsroom/2020/09/singapore-and-apple-partner-on-national-health-initiative-using-apple-watch/

Arantes, J. (2023). Educational data brokers: using the walkthrough method to identify data brokering by edtech platforms. Learning, Media and Technology, pp. 1-14. https://vuir.vu.edu.au/43665/

Arbi, I.A. (2020, Apr 21). Jokowi's millennial staffer, Ruangguru CEO resigns from State Palace. The Jakarta Post. https://www.thejakartapost.com/news/2020/04/21/jokowis-millennial-staffer-ruangguru-ceo-resigns-from-state-palace.html

Arlinta, D. (2023, 2 Mar). 'PeduliLindungi' reincarnates into 'Satu Sehat', features expanded. Kompas.id. https://web.archive.org/web/20230302104826/https://www.kompas.id/baca/english/2023/03/02/pedulilindungi-reincarnates-into-satu-sehat-features-expanded

ART GSC (Analysis and Research Team, General Secretariat of the Council of the EU). (2023, 24 Apr). ChatGPT in the public sector – overhyped or overlooked? Research Paper. https://www.consilium.europa.eu/media/63818/art-paper-chatgpt-in-the-public-sector-overhyped-or-overlooked-24-april-2023_ext.pdf

Asaduzzaman, M. & Virtanen, P. (2016). Governance theories and models. In A. Farazmand (ed.), Global encyclopedia of public administration, public policy, and governance, vol. 65, pp. 1–13. Springer International

Associated Press (2023, 20 Sept). George RR Martin and John Grisham among group of authors suing OpenAI. The Guardian. https://www.theguardian.com/books/2023/sep/20/authors-lawsuit-openai-george-rr-martin-john-grisham

Aziz, F. (2022, 18 Aug). Two e-learning co-curriculum modules now on Education Ministry's DELIMa. The Star. https://www.thestar.com.my/news/nation/2022/08/18/two-e-learning-co-curriculum-modules-now-on-education-ministrys-delima

Barr, K. (2023, 16 Mar). GPT-4 is a giant black box and its training data remains a mystery. Gizmodo. https://gizmodo.com/chatbot-gpt4-open-ai-ai-bing-microsoft-1850229989

Basyir, M. (2022, 30 Mar). TI-M: MySejahtera ownership, supervision issue worrying. New Straits Times. https://api.nst.com.my/news/nation/2022/03/784559/ti-m-mysejahtera-ownership-supervision-issue-worrying

Bayamlıoğlu, E. & Leenes, R. (2018). The 'rule of law' implications of data-driven decision-making: a techno-regulatory perspective. Law, Innovation and Technology, 10(2), pp. 295-313

Bedford, O., & Chua, S.H. (2017). Everything also I want: An exploratory study of Singaporean kiasuism (fear of losing out). Culture & Psychology, 24(4), pp. 491–511

Bell, J. & Lichère, F. (2022). Contemporary French administrative law. CUP

Bender, E.M., Gebru, T., McMillan-Major, A. & Mitchell, M. (2021). On the dangers of stochastic parrots: Can language models be too big? In Conference on Fairness, Accountability, and Transparency (FaccT '21), March 3–10, 2021, Virtual Event, Canada. ACM, New York, NY, USA, https://dl.acm.org/doi/pdf/10.1145/3442188.3445922

Berners-Lee, T. (2017, 12 Mar). Three challenges for the web, according to its inventor. World Wide Web Foundation. https://webfoundation.org/2017/03/web-turns-28-letter/

Bevir, M. (2012). Governance: A very short introduction. OUP

Black, J. (2002). Critical reflections on regulation. Australian Journal of Legal Philosophy, 27, pp. 1-36

Boo, S.-L. (2020, 12 Aug). How MySejahtera protects your data and does more than contact tracing. Code Blue. https://codeblue.galencentre.org/2020/08/12/how-mysejahtera-protects-your-data-and-does-more-than-contact-tracing/

Boo, S.-L. & Batumalai, K. (2022, 30 Mar). Khairy: Government owns MySejahtera app's IP, personal data, source code. Code Blue. https://codeblue.galencentre.org/2022/03/31/khairy-government-owns-mysejahtera-apps-ip-personal-data-source-code/

Borji, A. (2023, 28 Feb). A categorical archive of ChatGPT failures. ArXiv preprint, https://arxiv.org/pdf/2302.03494.pdf

Braithwaite, J. (1982). Enforced self-regulation: A new strategy for corporate crime control. Michigan Law Review, 80(7), pp. 1466-1507. https://repository.law.umich.edu/cgi/viewcontent.cgi?article=4245&context=mlr

Burman, A. & Sharma, U. (2021, Apr). How would data localization benefit India? Carnegie India. https://carnegieendowment.org/files/202104-Burman_Sharma_DataLocalization_final.pdf

Cao, P. (2022, October 29). Evidation Health: A platform that pays you to track your health. Harvard Business School, Digital Innovation and Transformation blog. https://d3.harvard.edu/platform-digit/submission/evidation-health-a-platform-that-pays-you-to-track-your-health/

Carlini, N., Tramèr, F., Wallace, E., Jagielski, M., Herbert-Voss, A., Lee, K., Roberts, A., Brown, T., Song, D.,; Erlingsson, U., Oprea, A. & Raffel, C. (2021). Extracting training data from large language models. Proceedings of the 30th USENIX Security Symposium, August 11–13, 2021, pp. 2633-2650. https://www.usenix.org/system/files/sec21-carlini-extracting.pdf

Carver, J. (2003). A new basis for governance effectiveness research. Board Leadership, 2003(67), pp. 4-5

Chang, Y.L. (2022). Communitarianism, properly understood. Canadian Journal of Law & Jurisprudence, 35(1), pp. 117-139

Cheah, K.Y.L. (2023, 27 Feb). Press release: Transparency and protection of privacy crucial for personal data collected through the MySejahtera application. Malaysian Bar. https://www.malaysianbar.org.my/article/news/press-statements/press-statements/press-release-transparency-and-protection-of-privacy-crucial-for-personal-data-collected-through-the-mysejahtera-application

Cheng, C. Y. & Wee, S. X. R. (2023). Kiasu (fear of losing out): An indigenous psychological construct in Singapore and its impact. International Perspectives in Psychology: Research, Practice, Consultation, 12(2), pp. 65-74

Chen, Y.-L. & Shin H.B. (2019). Neoliberal urbanism, contested cities and housing in Asia. Palgrave Macmillan

Chia, O. (2023a, 14 Feb). Civil servants to soon use ChatGPT to help with research, speech writing. The Straits Times. https://www.straitstimes.com/tech/civil-servants-to-soon-use-chatgpt-to-help-with-research-speech-writing

Chia, O. (2023b, 23 May). Public officers can use ChatGPT and similar AI, but must take responsibility for their work: MCI. The Straits Times. https://www.straitstimes.com/tech/public-officers-allowed-to-use-chatgpt-and-other-ai-but-must-take-responsibility-for-work-mci

Chik, W. (2021). The future of personal data protection law in Singapore. In G. K. Y. Chan & M. Yip (eds.) AI, data and private law: Translating theory into practice. Bloomsbury

Chin, C. (2022, 22 Aug). DELIMa improved to offer more personalized experience, says Radzi. The Star. https://www.thestar.com.my/news/education/2022/08/22/delima-improved-to-offer-more-personalised-experience-says-radzi

Christie, N. (1977). Conflicts as property. The British Journal of Criminology, 17(1), pp. 1–15. http://www.jstor.org/stable/23636088

Cohen, D. (2022). Any time, any place, any way, any pace: Markets, EdTech, and the spaces of schooling. Environment and Planning A: Economy and Space, 56(1), pp. 270-287. https://journals.sagepub.com/doi/10.1177/0308518X221084708

Cohen, J. (2019). Between truth and power. OUP

Constine, J. (2019, 22 Feb). Facebook will shut down its spyware VPN app Onavo. Tech Crunch. https://techcrunch.com/2019/02/21/facebook-removes-onavo/

Cornish, F., Breton, N., Moreno-Tabarez, U., Delgado, J., Rua, M., de-Graft Aikins, A. & Hodgetts, D. (2023). Participatory action research. Nature Reviews Methods Primers, 3(1). https://eprints.lse.ac.uk/118822/

Cotterrell, R. (2018). Law, emotion and affective community. Scholarly Paper ID 3212860, SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3212860

Cotterrell, R. (1999). Transparency, mass media, ideology and community. Cultural Values, 3, pp. 414-426

Couldry, N. & Mejias, U. (2019). The costs of connection: How data is colonizing human life and appropriating it for capitalism. Stanford University Press

Couldry, N. & Mejias, U. (2023). The decolonial turn in data and technology research: what is at stake and where is it heading? Information, Communication & Society, 26(4), 786–802

Cox, D. (2018, 3 Sept). Watch your step: why the 10,000 daily goal is built on bad science. The Guardian. https://www.theguardian.com/lifeandstyle/2018/sep/03/watch-your-step-why-the-10000-daily-goal-is-built-on-bad-science

Criddle, C. (2022, 31 Aug). Edtech companies breaking UK data laws, privacy campaigners claim. Financial Times

Custers, B. & Ursic, H. (2016). Big data and data reuse: a taxonomy of data reuse for balancing big data benefits and personal data protection. International Data Privacy Law, 6(1), pp. 4-15

Cyphers, B. (2020, 19 Mar). Google says it doesn't 'sell' your data. Here's how the company shares, monetizes, and exploits it. Electronic Frontier Foundation. https://www.eff.org/deeplinks/2020/03/google-says-it-doesnt-sell-your-data-heres-how-company-shares-monetizes-and.

Dacadoo (2023). Health risk quantification. https://www.dacadoo.com/products-services/dacadoo-health-risk-quantification/

Daim, N. & Radhi, N.A.M. (2023, 23 Feb). MoE finalising post pandemic Digital Education Policy. New Straits Times. https://www.nst.com.my/news/nation/2023/02/882727/moe-finalising-post-pandemic-digital-education-policy

Delacroix, S., & Lawrence, N. (2019). Bottom-up data trusts: Disturbing the 'one size fits all' approach to data governance. International Data Privacy Law, 9(4), pp. 236–252

De Pascalis, F. (2022). The journey to open finance: Learning from the open banking movement. European Business Law Review, 33(3), pp. 397–420

Devita, E. (2017, 28 Nov). Langkah menuju '100 Smart City' [Steps towards 100 Smart Cities]. Ministry of Communication and Information Technology. https://www.kominfo.go.id/content/detail/11656/langkah-menuju-100-smart-city/0/sorotan_media

Dickens, A. (2021). The right to health implications of data-driven health research partnerships. Thesis submitted for the degree of PhD in Law. University of Essex. Https://repository.essex.ac.uk/31194/1/PhD-%20FINAL%20VERSION%20(w.%20corrections).pdf

Doctorow, C. (2023, 16 October). Why Big Tech, cops, and spies were made for one another: The American surveillance state is a public-private partnership. The Intercept. https://theintercept.com/2023/10/16/surveillance-state-big-tech/

Drapalova, E., Wegrich K. (2020). Who governs 4.0? Varieties of smart cities. Public Management Review, 22(5), pp. 668-686

Drazewska, B. (2023a, 4 Mar). Smart cities should be people-centric, but is that really the case? The Business Times, p. 2

Drazewska, B. (2023b). Private-public data governance in Indonesia's smart cities: promises and pitfalls. In M. Findlay, L.M. Ong, L.M. & W. Zhang. (eds). Elgar companion to regulating AI and Big Data in emerging economies (pp. 59-77). Elgar

Drexl, J., Hilty, R., Desaunettes-Barbero, L., Greiner, F., Kim, D., Richter, H., Surblyte, G., & Wiedemann, K. (2016). Data ownership and access to data - position statement of the Max Planck Institute for Innovation and Competition of 16 August 2016 on the Current European Debate. https://www.researchgate.net/publication/318000556_Data_Ownership_and_Access_to_Data_-_Position_Statement_of_the_Max_Planck_Institute_for_Innovation_and_Competition_of_16_August_2016_on_the_Current_European_Debate

Developer Portal (2023). Getting started with Pair. https://www.developer.tech.gov.sg/products/categories/productivity-tools/pair/getting-started

Durkheim, E. (1982) [1895]. The rules of sociological method (Lukes, S. ed.; Halls W.D. tr.). The Free Press

East Ventures. (2023, 5 Apr). Every student has the same learning opportunities: Iman Usman, Co-Founder & COO of Ruangguru. https://east.vc/leadership/iman-usman/

EDMW (2023). New LumiHealth is pathetic [Online forum thread]. Hardwarezone. https://forums.hardwarezone.com.sg/threads/new-lumihealth-is-pathetic.6931737/

Ess, C. (2005). 'Lost in translation'?: Intercultural dialogues on privacy and information ethics (Introduction to special issue on Privacy and Data Privacy Protection in Asia). Ethics and Information Technology, 7, pp. 1–6

Elangovan, N. & Tan, Y. L. (2021, 7 Jan). Some TraceTogether users upset with Govt's revelation on police access to data, say they'll use it less. Today. https://www.todayonline.com/singapore/some-tracetogether-users-upset-govts-revelation-police-access-data-say-theyll-use-it-less

Elangovan, N. (2023, 10 Feb). Questions remain over Tin Pei Ling's role at Grab; firms should be more careful in hiring MPs, say analysts. Today. https://www.todayonline.com/singapore/tin-pei-ling-grab-companies-hiring-mp-2105521

Eliot, L. (2023, 27 Jan). Generative AI ChatGPT can disturbingly gobble up your private and confidential data, forewarns AI ethics and AI law. Forbes. https://www.forbes.com/sites/lanceeliot/2023/01/27/generative-ai-chatgpt-can-disturbingly-gobble-up-your-private-and-confidential-data-forewarns-ai-ethics-and-ai-law/?sh=7ca1294c7fdb

Engels, A. (2020). Singapore Government & Apple partnership: a privacy concern? Nextpit. https://www.nextpit.com/singapore-government-apple-partnership-privacy-concern

Ethnologue & the World Bank (2021). The countries with the most linguistic diversity [Graph]. Statista. https://www.statista.com/chart/3862/countries-with-the-most-spoken-languages/

Eyert, F., Irgmaier, F., Ulbricht, L. (2018). Algorithmic social ordering: Towards a conceptual framework. In G. Getzinger (ed.), Critical issues in science, technology and society studies: Conference proceedings of the 17th STS Conference Graz 2018, 7th-8th May 2018 (pp. 48-57). Verlag der Technischen Universität Graz. https://www.econstor.eu/bitstream/10419/191592/1/f-21526-full-text-Eyert-et_al-Algorithmic-v3.pdf

Feathers, T., Fondrie-Teitler, S., Waller, A., & Mattu, S. (2022, 16 Jun). Facebook is receiving sensitive medical information from hospital websites. The Markup. https://themarkup.org/pixel-hunt/2022/06/16/facebook-is-receiving-sensitive-medical-information-from-hospital-websites

Findlay, M. (2013). Regulating regulation — Who guards the guardian. In M. Findlay (ed.), Contemporary challenges in regulating global crises (International Political Economy series) (pp. 227–247). Palgrave Macmillan

Findlay, M. (2020). Property abandoned: Rights, wrongs and forgetting Durkheim. In P. Drahos P., G. Ghidini and H. Ulrich (eds.), Kritika: Essays on intellectual property: vol IV (pp. 100-120). Elgar

Findlay, M. (2023). Depending on AI, SMU CAIDG Research Paper No. 1/2023, SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4321460

Findlay, M. & Henham, R. (2007). Integrating theory and method in the comparative contextual analysis of trial process. In M. McConville. & W.H. Chui (eds.) Research methods for law (pp. 134-162). Edinburgh University Press

Findlay, M., Seah, J & Wong, W. (2023). AI and Big Data: Disruptive regulation. Elgar

Fischer, A. & Streinz, T. (2022). Confronting data inequality. Columbia Journal of Transnational Law, 60(3), pp. 829–956

Flyvberg, B. (2001). Making social science matter. Why social inquiry fails and how it can succeed again. CUP

Forbrukerrådet (2020). Out of control: How consumers are exploited by the online advertising industry. https://storage02.forbrukerradet.no/media/2020/01/2020-01-14-out-of-control-final-version.pdf

Foundry. (2023, 14 Sept). Foundry Mixer Healthcare, Indonesian Minister of Health Budi G. Sadikin: 'Together we leapfrog Indonesia's healthcare transformation'. https://www.prnewswire.com/apac/news-releases/foundry-mixer-healthcare-indonesian-minister-of-health-budi-g-sadikin-together-we-leapfrog-indonesias-healthcare-transformation-301927595.html

Fowler, G.A. (2023, 26 Sept). Your Instagrams are training AI: There's little you can do about it. The Washington Post

FTC (2013, 1 Feb). Path social networking app settles FTC charges it deceived consumers and improperly collected personal information from users' mobile address books. https://www.ftc.gov/news-events/news/press-releases/2013/02/path-social-networking-app-settles-ftc-charges-it-deceived-consumers-improperly-collected-personal

Gal, U. (2023, 8 Feb). ChatGPT is a data privacy nightmare. If you've ever posted online, you ought to be concerned. The Conversation. https://theconversation.com/chatgpt-is-a-data-privacy-nightmare-if-youve-ever-posted-online-you-ought-to-be-concerned-199283

Garante per la Protezione dei Dati Personali (2023, 30 March). Provvedimento del 30 marzo 2023 [9870832] (in Italian). Available at https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9870832

Gellman B., & Soltani, A. (2013, 14 October). NSA collects millions of e-mail address books globally. Washington Post. https://www.washingtonpost.com/world/national-security/nsa-collects-millions-of-e-mail-address-books-globally/2013/10/14/8e58b5be-34f9-11e3-80c6-7e6dd8d22d8f_story.html

GEM (Global Education Monitoring) Report UNESCO (2023, 6 Sept). Minister of Education, Indonesia, Nadiem Makarim. YouTube. https://www.youtube.com/watch?v=W0EQAIw9YBs

Gilbert, J. (2023, 18 Sept). Padu: Governance by data. New Straits Times. https://www.nst.com.my/business/2023/09/956583/padu-governance-data

Google for Education (nd.) Privacy & Security Center. https://edu.google.com/intl/ALL_sg/why-google/privacy-security/

Google Play. (2023a). SATUSEHAT Mobile. https://play.google.com/store/apps/datasafety?id=com.telkom.tracencare&hl=en&gl=US

Google Play (2023b). MySejahtera. https://play.google.com/store/apps/details?id=my.gov.onegovappstore.mysejahtera&hl=en_SG&gl=US

Google Play (2023c). Ruangguru. https://play.google.com/store/apps/details?id=com.ruangguru.livestudents&hl=en&gl=US

Granovetter, M. S. (1973). The strength of weak ties. American Journal of Sociology, 78(6), pp. 1360-1380

Greenleaf, G. (2010). Limitations of Malaysia's Data Protection Bill. 104(1) Privacy Laws & Business International Newsletter, pp. 5-7

Gurman, M. (2023, 2 May). Samsung bans staff's AI use after spotting ChatGPT data leak. Bloomberg. https://www.bloomberg.com/news/articles/2023-05-02/samsung-bans-chatgpt-and-other-generative-ai-use-by-staff-after-leak

GVH (2023, 25 Jul). GVH investigates how Microsoft informs consumers about its new search service. https://gvh.hu/en/press_room/press_releases/press-releases-2023/gvh-investigates-how-microsoft-informs-consumers-about-its-new-search-service

Hack for Public Good. (2023). Pair. https://hack.gov.sg/2023-prototypes/pair/

Hakim, R.N. & Galih, B. (2020, 15 Apr). Surat stafsus milenial Jokowi yang dinilai berpotensi korupsi…. [Jokowi's millennial special staff letter assessed as potentially corrupt...]. Kompas. https://nasional.kompas.com/read/2020/04/15/09401311/surat-stafsus-milenial-jokowi-yang-dinilai-berpotensi-korupsi

Haque, M.S. (2007). Theory and practice of public administration in Southeast Asia: Traditions, directions, and impacts. International Journal of Public Administration of Public Administration, 30(12-14), pp. 1–63

Harding, A. (2018). Five-foot ways as public and private domain in Singapore and beyond. Journal of Property Planning and Environmental Law, 10(1), pp. 36-55

Hård, M. (1993). Beyond harmony and consensus: A social conflict approach to technology. Science, Technology, & Human Values, 18(4), pp. 408–432. http://www.jstor.org/stable/690002

Haus, M. (2018). Governance and power. In H. Heinelt (ed.) Handbook on Participatory Governance (pp. 53-76). Elgar

Hayat, M. A. (2007). Privacy and Islam: From the Quran to data protection in Pakistan. Information & Communications Technology Law, 16(2), pp. 137-148

Hern, A. (2021, 30 May). Gadgets have stopped working together, and it's becoming an issue. The Guardian. https://www.theguardian.com/technology/2021/may/30/gadgets-have-stopped-working-together-interoperability-apple

Hern, A. (2019, 18 Apr). Facebook uploaded email contacts of 1.5m users without consent. The Guardian. https://www.theguardian.com/technology/2019/apr/18/facebook-uploaded-email-contacts-of-15m-users-without-consent

Hii, B. H. & Sofwan, M. b. M. (2022). Teachers' perception towards DELIMa learning platform for mathematics online teaching and learning in Sibu district. International Journal of Academic
Research in Business and Social Sciences,12(11), pp. 1888–1908

Hillman, V. (2019). Student agency in a data-driven educational ecosystem. Journal of Design and Science. https://jods.mitpress.mit.edu/pub/bw0s06i8/download/pdf

Hirst, P. (2000). Democracy and governance. In J. Pierre (ed.) Debating governance: Authenticity, steering, and democracy (pp. 13-35). OUP

Hooper, L., Livingstone, S., & Pothong, K. (2022). Problems with data governance in UK schools: Google Classroom and ClassDojo. Digital Futures Commission, 5Rights Foundation. https://eprints.lse.ac.uk/119736/

Horibe, M. (2017). Privacy culture and data protection laws in Japan. 39th International Conference of Data Protection and Privacy Commissioners, Hong Kong, 25-29 September 2017. https://www.ppc.go.jp/files/pdf/290928_en_horibespeech.pdf

Hufty, M. (2011). Investigating policy processes: The Governance Analytical Framework (GAF). In U. Wiesmann & H. Hurni (eds.), Research for Sustainable Development: Foundations, Experiences, and Perspectives (pp. 403-424) . NCCR North-South / Geographica Bernensia

HRW. (2021a, 10 Feb). Privacy snapshots: Ruangguru. https://features.hrw.org/features/StudentsNotProducts/files/privacy_snapshots/Privacy%20Snapshot%20-%20Indonesia%20Ruangguru.pdf

HRW (Human Rights Watch). (2021b). Privacy snapshots: DELIMa. https://features.hrw.org/features/StudentsNotProducts/files/privacy_snapshots/Privacy%20Snapshot%20-%20Malaysia%20DELIMa.pdf

HRW (Human Rights Watch). (2021c). Privacy snapshots: Zoom. https://features.hrw.org/features/StudentsNotProducts/files/privacy_snapshots/Privacy%20Snapshot%20-%20Global%20Zoom.pdf

HRW (Human Rights Watch). (2021d). Privacy snapshots: Microsoft Teams. https://features.hrw.org/features/StudentsNotProducts/files/privacy_snapshots/Privacy%20Snapshot%20-%20Global%20Microsoft%20Teams.pdf

HRW (Human Rights Watch). (2022, 25 May). 'How dare they peep into my private life?' Children's rights violations by governments that endorsed online learning during the Covid-19 pandemic. https://www.hrw.org/report/2022/05/25/how-dare-they-peep-my-private-life/childrens-rights-violations-governments

Hussin, R. & Kamal, A. (2023, 3 Jan). Clarify MySejahtera deal once and for all. Code Blue. https://codeblue.galencentre.org/2023/01/03/clarify-mysejahtera-deal-once-and-for-all-rais-hussin-ameen-kamal/

IDEAS (2022, 27 Apr). MySejahtera episode poses questions about data privacy. https://www.ideas.org.my/ideas-mysejahtera-episode-poses-questions-about-data-privacy/

Idris, I.K. & Pitaloka, D. (2022, 19 Aug). Indonesia's 'Super Apps': will they be another waste of state budget? The Conversation. https://theconversation.com/indonesias-super-apps-will-they-be-another-waste-of-state-budget-188702

IEEE (2016, 2 Sept). On the use of AI – the dependency dilemma [interview with Jeff Robbins]. IEEE Technical Community Spotlight. https://technical-community-spotlight.ieee.org/ai-ethical-dilemma/

IMDA & PDPC (2020). Model Artificial Intelligence Governance Framework, 2nd ed. https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf

ISL (Internet Safety Labs). (2022). K-12 EdTech safety benchmark: National findings report. Part 1. https://internetsafetylabs.org/resources/reports/2022-k12-edtech-safety-benchmark-national-findings-part-1/

Ismail, N. & Yong Cieh, E. L. (2013). Beyond data protection: Strategic case studies and practical guidance. Springer

Italy, France & Germany (2023, 20 Nov). An innovation-friendly approach based on European values for the AI Act: Joint non-paper by IT, FR and DE. https://logistic-natives.com/an-innovation-friendly-approach-based-on-european-values-for-the-ai-act.html

Jibiki, K. (2022, 13 Jan). Jokowi taps Indonesia's corporate stars for policymaking input. Nikkei Asia. https://asia.nikkei.com/Politics/Jokowi-taps-Indonesia-s-corporate-stars-for-policymaking-input

Jones-Jang, S.M., Park, Y.J. (2023). How do people react to AI failure? Automation bias, algorithmic aversion, and perceived controllability. Journal of Computer-Mediated Communication, 28(1), pp. 1-8. https://academic.oup.com/jcmc/article/28/1/zmac029/6827859

Jung Han, H. (2020, 13 Aug). An algorithm shouldn't decide a student's future. Politico. https://www.politico.eu/article/an-algorithm-shouldnt-decide-students-future-coronavirus-international-baccalaureate/

Kapong, K. B. (2023, 15 Oct). Digitisation of textbooks among tech interventions to enhance education in Sarawak. Borneo Post. https://www.theborneopost.com/2023/10/15/digitisation-of-textbooks-among-tech-interventions-to-enhance-education-in-sarawak/

Kelly, H. (2021, 15 Jul). Lots of apps use your personal contacts. Few will tell you what they do with them. The Washington Post. https://www.washingtonpost.com/technology/2021/07/15/contacts-sharing-privacy/

Kelly, H. (2021, 22 Jul). A priest's phone location data outed his private life. It could happen to anyone. The Washington Post. https://www.washingtonpost.com/technology/2021/07/22/data-phones-leaks-church/

KEM (Kementerian Pendidikan Malaysia [Ministry of Education of Malaysia]). (2013). Malaysia Education Blueprint, 2013-2025. https://www.moe.gov.my/menumedia/media-cetak/penerbitan/dasar/1207-malaysia-education-blueprint-2013-2025/file

KEM. (2020a). DELIMa. https://sites.google.com/moe-dl.edu.my/delimaeng/home?authuser=1

KEM. (2020b). DELIMa: Dasar Privasi (Privacy Policy) [in Malay]. https://sites.google.com/moe-dl.edu.my/public/dasar-privasi

KEM. (2021). DELIMa: Terma Dan Syarat (Terms and Conditions) [in Malay]. https://d2.delima.edu.my/login

Kemendikbud (Ministry of Education and Culture). (2020, 15 Apr). Kemendikbud gandeng swasta siapkan sistem belajar daring [Ministry of Education and Culture collaborates with private sector to prepare online learning system]. https://www.kemdikbud.go.id/main/blog/2020/03/kemendikbud-gandeng-swasta-siapkan-sistem-belajar-daring

Kementerian Kesehatan (2023a). Kebijakan Privasi SATUSEHAT Mobile | SATUSEHAT Mobile Privacy Policy. https://faq.kemkes.go.id/faq/kebijakan-privasi-satusehat-mobile-i-satusehat-mobile-privacy-policy

Kementerian Kesehatan (2023b). Syarat dan Ketentuan Penggunaan SATUSEHAT Mobile | Terms and Conditions of Use of SATUSEHAT Mobile. https://faq.kemkes.go.id/faq/syarat-dan-ketentuan-penggunaan-satusehat-mobile-i-terms-and-conditions-of-use-of-satusehat-mobile

Kennedy, H., Poell, T. & van Dijck, J. (2015). Data and agency. Big Data & Society, 2(2), pp. 1-7

Keping, Y. (2018). Governance and good governance: A new framework for political analysis. Fudan Journal of the Humanities and Social Sciences, 11, pp 1–8

Khalaf, R. (2018, 20 Nov). Google has a responsibility to protect DeepMind data. The Financial Times. https://www.ft.com/content/83e1e46c-ebf0-11e8-8180-9cf212677a57

Kim, E. (2023, 1 Feb). Microsoft warns employees not to share 'sensitive data' with ChatGPT. Business Insider. https://www.businessinsider.com/chatgpt-microsoft-warns-employees-not-to-share-sensitive-data-openai-2023-1

Kitiyadisai, K. (2005). Privacy rights and protection: Foreign values in modern Thai context. Ethics and Information Technology, 7(1), pp. 17-26

Klein, L.F. & D'Ignazio, C. (2020). Data feminism. MIT Press

Komljenovic, J. (2021). The rise of education rentiers: Digital platforms, digital data and rents. Learning, Media and Technology, 46(3), pp. 320-332

Kong. (2022, 7 Jan). 1 month of LumiHealth: How to get vouchers faster (review). https://sgunlocked.com/1-month-of-lumihealth/

Kong, D. & Yoon, K. (2018). Modes of public governance: A typology toward a conceptual modeling. World Political Science, 14(1), pp. 145-167

Kovacs, A. (2020, 28 May). When our bodies become data, where does that leave us? DeepDives. https://deepdives.in/when-our-bodies-become-data-where-does-that-leave-us-906674f6a969

Kovacs, A. & Jain, T. (2021). Informed consent - said who? A feminist perspective on principles of consent in the age of embodied data. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3788322

Krein, A. (2020, June). The screens that ate school. The Monthly. https://www.themonthly.com.au/issue/2020/june/1590933600/anna-krien/screens-ate-school

Kuosmanen, E., Visuri, A., Risto, R. & Hosio, S. (2022). Comparing consumer grade sleep trackers for research purposes: A field study. Frontiers in Computer Science, 4, pp. 1-16. https://www.frontiersin.org/articles/10.3389/fcomp.2022.971793/full

Kwet, M. (2017). Operation Phakisa Education: Why a secret? Mass surveillance, inequality, and race in South Africa's emerging national e-education system. First Monday, 22(12). https://firstmonday.org/ojs/index.php/fm/article/view/8054/6585

Kwet, M. (2019, 13 Mar). Digital colonialism is threatening the Global South. Al Jazeera. https://www.aljazeera.com/opinions/2019/3/13/digital-colonialism-is-threatening-the-global-south

Landi, H. (2021, 13 Sept). Fitbit, Apple user data exposed in breach impacting 61M fitness tracker records. https://www.fiercehealthcare.com/digital-health/fitbit-apple-user-data-exposed-breach-impacting-61m-fitness-tracker-records

Law, J. (2003). Ordering and obduracy. Centre for Science Studies, Lancaster University, Lancaster. http://www.comp.lancs.ac.uk/sociology/papers/LawOrdering-and-Obduracy.pdf

Law, J. (1993). Organizing modernity. Wiley Blackwell

Lessig, L. (2002). Privacy as property. Social Research, 69(1), pp. 247-269

Li, C. & Lalani, F. (2022). How to address digital safety in the metaverse. World Economic Forum. https://www.weforum.org/agenda/2022/01/metaverse-risks-challenges-digital-safety/

Li, C. & White, K. (2023) Metaverse privacy and safety. World Economic Forum. https://www.weforum.org/reports/privacy-and-safety-in-the-metaverse/

Liew, J.T., (2021). Tech: Can MySejahtera be turned into a super app, and should it? The Edge Malaysia. https://theedgemalaysia.com/article/tech-can-mysejahtera-be-turned-super-app-and-should-it

Lilly & Co (2019, 8 Aug). Lilly, Evidation Health and Apple study shows personal digital devices may help in the identification of mild cognitive impairment and mild Alzheimer's Disease dementia. PR Newswire. https://www.prnewswire.com/news-releases/lilly-evidation-health-and-apple-study-shows-personal-digital-devices-may-help-in-the-identification-of-mild-cognitive-impairment-and-mild-alzheimers-disease-dementia-300898719.html

Lindh, A. & Nolin, J. (2016). Information we collect: Surveillance and privacy in the implementation of Google apps for education. European Educational Research Journal, 15(6), pp. 644-663

Litman-Navarro, K. (2019, 17 Jun). We read 150 privacy policies. They were an incomprehensible disaster. The New York Times. https://www.nytimes.com/interactive/2019/06/12/opinion/facebook-google-privacy-policies.html

Loasana, N. A. (2023, 3 Mar). Privacy concerns arise as COVID-19 app repurposed. Jakarta Post. https://www.thejakartapost.com/indonesia/2023/03/02/concerns-over-privacy-as-covid-19-tracing-app-repurposed.html

Lopez-Claros, A., Dahl, A.L., & Groff, M. (2020). Global governance and the emergence of global institutions for the 21st century. CUP

Lubben, S. J. (2013). Separation and dependence: explaining modern corporate governance. Seton Hall Law Review, 43(3), pp. 893-908

Lee, J. (2023, 26 Sept). Report: Govt to repurpose MySejahtera app as mechanism for handing out targeted subsidies. Malay Mail. https://www.malaymail.com/news/malaysia/2023/09/26/report-govt-to-repurpose-mysejahtera-app-as-mechanism-for-handing-out-targeted-subsidies/92947

LeVasseur, L., Edwards, Z., Alexanyan, K., Maler, E. & Spalding, S. (2021, 4 May). Me2B alliance product testing report: School mobile apps student data sharing behavior. https://internetsafetylabs.org/wp-content/uploads/2021/05/school-apps-data-sharing-behavior-spotlight-report-final.pdf

Loo, J. (2022). Private-public partnerships in Singapore - KAS scoping study (unpublished manuscript on file with authors)

LumiHealth. (2023a, 24 Jul). Privacy Policy. https://www.lumihealth.sg/privacy-policy

LumiHealth. (2023b, 24 Jul). Terms of Use. https://www.lumihealth.sg/terms-of-use

Lumihealth. (2023c). Program Participation Agreement (PPA). Version retrieved 13 October 2023 (on file with authors)

LumiHealth. (2023d). FAQ. https://www.lumihealth.sg/faq#programme-overview-lh3_1

Lynn, L.E., Heinrich, C.J., & Hill, C.J. (2000). Studying governance and public management: Challenges and prospects. Journal of Public Administration Research and Theory: J-PART, 10(2), pp. 233-261

Muhammad Ashraf, M. (2023, 23 Jan). Gender-based Abuse on the Metaverse: The New Internet is Being Coded on a Toxic Palimpsest. Bot Populi. https://botpopuli.net/?post_type=post&p=6887

Malay Mail. (2023, 3 Oct). Digitalisation agenda among Education Ministry's priorities in Budget 2024, says minister. https://www.malaymail.com/news/malaysia/2023/10/03/digitalisation-agenda-among-education-ministrys-priorities-in-budget-2024-says-minister/94241

Bernama (2023, 27 Feb). Govt aiming to make MySejahtera a 'public health super app', Malaysia Now. https://www.malaysianow.com/news/2023/02/27/govt-aiming-to-make-mysejahtera-a-digital-public-health-super-app

Mance, H. (2023, 30 Jan). Shoshana Zuboff: 'Privacy has been extinguished. It is now a zombie'. https://www.ft.com/content/0cca6054-6fc9-4a94-b2e2-890c50d956d5

Markoff, J. (2015). Machines of loving grace. HarperCollins

Martinus, M. (2022) Smart City and privacy concerns during COVID-19: Lessons from Singapore, Malaysia, and Indonesia. In T. Phan and D. Damian (eds). Smart cities in Asia: Regulations, problems, and development (pp. 33-47). Springer

Mau, S. (2019). The metric society: On the quantification of the social (tr. by S. Howe). Polity Press

MCI (2024). Digital society. https://www.mci.gov.sg/what-we-do/digital-singapore/digital-society/

McLean, J.M. (2019). For a law of public contract per se: An intervention from liberal contract theory. Oxford Journal of Legal Studies, 39(4), pp. 856–877

Meijer, A. J., Gil-Garcia, J. R. & Rodriguez-Bolivar, M. P. (2016). Smart city research: Contextual conditions, governance models, and public value assessment. Social Science Computer Review, 34(6), pp. 647-656

Menary, R. (2007). Writing as thinking. Language Sciences, 29(5), pp. 621-632

MHLG (Ministry of Housing and Local Government). (2018). Malaysia Smart City Framework (MSCF): Executive Summary. https://www.kpkt.gov.my/kpkt/resources/user_1/GALERI/PDF_PENERBITAN/FRAMEWORK/FRAMEWORK_SMART_CITY_EXECUTIVE_SUMMARY.pdf

MIC (Ministry of Information and Communications of the Socialist Republic of Vietnam). (2020, 26 May). Facebook launches campaign to assist Vietnam in developing digital economy. https://english.mic.gov.vn/Pages/TinTuc/tinchitiet.aspx?tintucid=142092

MICI (Ministry of Communications and Information) and Smart Nation Singapore (2023, 4 Dec). National Artificial Intelligence Strategy 2.0 to uplift Singapore's social and economic potential: press release. https://www.smartnation.gov.sg/media-hub/press-releases/04122023/

Microsoft (2023). Privacy Statement. https://privacy.microsoft.com/en-US/privacystatement

Min, A.C. (2023, 18 Jul). 4,000 civil servants using government Pair chatbot for writing, coding. The Straits Times. https://www.straitstimes.com/singapore/4000-civil-servants-using-government-pair-chatbot-for-writing-coding

Mohad, M. (2022, 6 Oct). Whither accountability for government procurement? BFM morning brief podcast. https://www.bfm.my/podcast/morning-run/morning-brief/corruption-procurement-mysejahtera-kpisoft-malaysia-littoral-combat-ship-letter-of-support

Mok, A. (2023, 11 Jul). Amazon, Apple, and 12 other major companies that have restricted employees from using ChatGPT. Business Insider. https://www.businessinsider.com/chatgpt-companies-issued-bans-restrictions-openai-ai-amazon-apple-2023-7

Mozilla (2023, November). Ask Microsoft: Are you using our personal data to train AI? https://foundation.mozilla.org/en/campaigns/microsoft-ai/

Mulyanto, R. (2022, 8 August). Indonesia launches healthcare data integration platform as it continues digital transformation. https://www.healthcareitnews.com/news/asia/indonesia-launches-healthcare-data-integration-platform-it-continues-digital

MySejahtera (2022). Polisi Privasi (Privacy policy) [in Malay] https://mysejahtera.moh.gov.my/ms/tentang-mysejahtera/polisi-privasi

Narayanan, A. & Kapoor, S. (2022, 6 Dec). ChatGPT is a bullshit generator. But it can still be amazingly useful. Substack. https://aisnakeoil.substack.com/p/chatgpt-is-a-bullshit-generator-but

Nasr, M., Carlini, N., Hayase, J., Jagielski, M., Cooper, A. F., Ippolito, D. , Choquette-Choo, C. A., Wallace, E., Tramèr, F., & Lee, K. (2023, 28 Nov). Scalable extraction of training data from (production) language models. ArXiv preprint. https://arxiv.org/pdf/2311.17035.pdf

Neff, G. & Nafus, D. (2016). Self-tracking. MIT Press

Nethercote, M. (2023). Platform landlords: Renters, personal data and new digital footholds of urban control. Digital Geography and Society, 5, pp. 1-15

Ng, A. (2021, 2 Sept). What does it actually mean when a company says, "we do not sell your data"? The Markup. https://themarkup.org/the-breakdown/2021/09/02/what-does-it-actually-mean-when-a-company-says-we-do-not-sell-your-data

Nissenbaum, H. (2009). Privacy in context. Stanford University Press

News Desk (2020, 28 Mar). Telkomsel provides students with 30GB of free internet for distance learning. The Jakarta Post. https://www.thejakartapost.com/life/2020/03/28/telkomsel-provides-students-with-30gb-of-free-internet-for-distance-learning.html

Nguyen An Luong, D. (2020, 28 Oct). Vietnam's ambitious politicians: In Facebook we trust. Fulcrum. https://fulcrum.sg/vietnams-ambitious-politicians-in-facebook-we-trust/

Norton, T. B. (2016). The non-contractual nature of privacy policies and a new critique of the notice and choice privacy protection model. Fordham Intellectual Property, Media and Entertainment Law Journal, 27(1), pp. 181-210. https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=1641&context=iplj

OGP (Open Government Products). (2023). Hack for Public Good 2023 Demo Day. Youtube. https://www.youtube.com/watch?v=mgxE3IPE4WY (9:20-12:40)

OAIC (Office of the Australian Information Commissioner) (2022, 21 Dec). Australian Privacy Principles Guidelines, Chapter B: Key Concepts (ver. 1.4)

O'Flaherty, K. (2021, 9 May). How private is your Gmail, and should you switch? The Guardian. https://www.theguardian.com/technology/2021/may/09/how-private-is-your-gmail-and-should-you-switch

OpenAI. (2023a). GPT-4 technical report. https://cdn.openai.com/papers/gpt-4.pdf

OpenAI. (2023b). Text generation. https://platform.openai.com/docs/guides/text-generation

Otto, B., ten Hompel M. & Wrobel S. (2022). Designing data spaces. The ecosystem approach to competitive advantage. Springer

PAC (Laporan Jawatankuasa Kira-Kira Wang Negara). (2022). Pembangunan dan perolehan aplikasi MySejahtera di bawah Kementerian Kesihatan Malaysia, Kementerian Kewangan dan Jabatan Perdana Menteri, Dr.18 Tahun 2022 [Procurement and development of the MySejahtera application under the Malaysian Ministry of Health (KKM), Ministry of Finance (MOF) and the Prime Minister's Department (JPM)] [in Malay]. https://www.parlimen.gov.my/pac/review/docs-262-324.pdf

Park, B.-G., Child Hill, R. & Saito, A. (2012). Locating neoliberalism in East Asia : Neoliberalizing spaces in developmental states. Wiley

Paska Darmawan, J. (2016). The potential drawbacks of forced data localization in Indonesia. Center for Digital Society. https://cfds.fisipol.ugm.ac.id/2016/04/07/the-potential-drawbacks-of-forced-data-localisation-in-indonesia/

Perez, O. (2002). Using private-public linkages to regulate environmental conflicts: The case of international construction contracts. Journal of Law and Society, 29(1), pp. 77-110

Pierre, J. (2000). Introduction. Understanding governance. In J. Pierre (ed.) Debating governance: authenticity, steering, and democracy (pp. 1-10). OUP

Pineau, E. & Hummel, T. (2023, 30 Nov). Stop using WhatsApp, get Paris-made alternative, French PM tells ministers. Reuters. https://www.reuters.com/world/europe/stop-using-whatsapp-get-paris-made-alternative-french-pm-tells-ministers-2023-11-29/

Pisa, M. & Polcari, J. (2019). Governing Big Tech's pursuit of the "next billion users". Center for Global Development. https://www.cgdev.org/publication/governing-big-techs-pursuit-next-billion-users

Piso, Z., L. Goralnik, J. C. Libarkin, & M. C. Lopez. (2019). Types of urban agricultural stakeholders and their understandings of governance. Ecology and Society, 24(2), pp. 18-33

Plantin J.C., Lagoze C., Edwards P., & Sandvig C. (2016). Infrastructure studies meet platform studies in the age of Google and Facebook. New Media & Society, 20(1), pp. 293-310

Powles, J. & Hodson, H. (2017). Google DeepMind and healthcare in an age of algorithms. Health and Technology, 7, pp. 351–367

Prabowo, H. (2023, 6 Mar). Menakar potensi penyalahgunaan data 'superapp' SatuSehat [Measuring the potential of SatuSehat 'superapp' data misuse]. Tirto. https://tirto.id/menakar-potensi-penyalahgunaan-data-superapp-satusehat-gDb1

Ramli, R. (2023, 10 Oct). Opening remarks: Launch of the 28th Malaysia Economic Monitor. https://www.ekonomi.gov.my/sites/default/files/2023-10/Speech_MEM-launch_10Oct23.pdf

Random. (2022) Hi everyone! Is lumihealth or healthy365 better (in terms of getting rewards for walking)? Thank you!! [Online forum thread]. Seedly. https://seedly.sg/posts/hi-everyone-is-lumihealth-or-healthy365-better-in-terms-of-getting-rewards-for-walking-thank-you/

Random. (2023). Hi! Is Lumihealth better / easier in terms of earning rewards or using Healthy365? Thanks :) [Online forum thread]. Seedly. https://seedly.sg/posts/hi-is-lumihealth-better-easier-in-terms-of-earning-rewards-or-using-healthy365-thanks/

r/askSingapore. (2023). Lumi Health vs Healthy 365 [Online forum thread]. Reddit. https://www.reddit.com/r/askSingapore/comments/10dg4qb/lumi_health_vs_healthy_365/

Rayda, N. (2023, 23 Jul). Indonesian tribe feels its tradition is under threat from smartphones; requests internet signals be halted. CNA. https://www.channelnewsasia.com/asia/indonesia-baduy-tribe-remote-community-technology-smartphone-internet-3644716

Remolina, N. (2019). Open finance: Regulatory challenges of the evolution of data sharing arrangements in the financial sector. SMU CAIDG Research Paper No. 05/2019, SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3475019

Remolina, N. & Findlay, M. (2021, 22 Apr). The paths to digital self-determination - A foundational theoretical framework (April 22, 2021). SMU CAIDG Research Paper No. 03/2021, SSRN. https://ssrn.com/abstract=3831726

Rennie, E., Schmieder, K., Thomas, J., Howard, S.K., Ma, J., & Yang, J. (2019). Privacy and app use in Australian primary schools: insights into school-based Internet governance. Media International Australia, 170(1), pp. 78-89

Riles, A. (2020). Building platforms for collaboration: A new comparative legal challenge. In M. Corrales Compagnucci, N. Forgó, T. Kono, S. Teramoto & E. P. M. Vermeulen (eds.), Legal Tech and the New Sharing Economy (pp. 15-20). Springer

Rinke, A. (2023, Nov 21). Exclusive: Germany, France and Italy reach agreement on future AI regulation. Reuters. https://www.reuters.com/technology/germany-france-italy-reach-agreement-future-ai-regulation-2023-11-18/

Risse, M. (2023). Political theory of the digital age: Where artificial intelligence might take us. CUP

r/Lumihealth (2022). Lumihealth beyond the two year programme [Online forum thread]. Reddit. https://www.reddit.com/r/Lumihealth/comments/w0967z/lumihealth_beyond_the_two_year_programme/

Rodriguez Bolivar, M. P. (2019). Setting foundations for the creation of public value in smart cities. Springer

Rokom. (2023, 28 Feb). Besok PeduliLindungi resmi bertransformasi menjadi SATUSEHAT Mobile [Tomorrow PeduliLindungi will officially transform into SATUSEHAT Mobile]. Communications & Public Services Bureau of the Indonesian Ministry of Health. https://sehatnegeriku.kemkes.go.id/baca/rilis-media/20230228/2042474/besok-pedulilindungi-resmi-bertransformasi-menjadi-satusehat-mobile/

Roystonn, K., AshaRani, P. V., Devi, F., Wang, P., Zhang, Y., Jeyagurunathan, A., Edimansyah, A., Tudor Car, L., Siow, A.C. & Subramaniam, M. (2023). Exploring views and experiences of the general public's adoption of digital technologies for healthy lifestyle in Singapore: a qualitative study. Frontiers in Public Health, 11, pp. 1-12

Røiseland, A. (2023). For all seasons? Exploring the policy-context for co-creation. Public Money and Management, pp. 1-9. https://www.tandfonline.com/doi/epdf/10.1080/09540962.2023.2206046?needAccess=true

Ruangguru (2023a). About Ruangguru. https://www.ruangguru.com/about-us

Ruangguru (2023b, 6 Jun). Privacy policy. https://www.ruangguru.com/privacy-policy

Salim, S. (2022, 31 Mar). Info, data in MySejahtera app fully owned by govt, says Khairy. The Edge Malaysia. https://theedgemalaysia.com/article/all-data-and-information-obtained-result-use-mysejahtera-sole-property-government-%E2%80%94-khairy

Savov, V. 2020 (16 Sept). Singapore to pay citizens for keeping healthy with Apple Watch. Bloomberg. https://www.bloomberg.com/news/articles/2020-09-16/singapore-to-pay-citizens-for-keeping-healthy-with-apple-watch#xj4y7vzkg

Sharon, A. (2023, Jun 22). Malaysia's Omnibus Act: Streamlining data sharing for efficient governance. OpenGov Asia. https://opengovasia.com/malaysias-omnibus-act-streamlining-data-sharing-for-efficient-governance/

Sharwood, S. (2020, 16 Sept). Singapore to pay its citizens to wear Apple Watches. The Register. https://www.theregister.com/2020/09/16/singapore_apple_watch_deal/

Shumailov, I., Shumaylov, Z., Zhao, Y., Gal, Y., Papernot, N. & Anderson, R. (2023, 13 May). The curse of recursion: Training on generated data makes models forget. ArXiv Preprint. https://arxiv.org/pdf/2305.17493v2.pdf

Silalahi, M. (2017, 5 Sept). Telkomsel digital advertising tumbuh signifikan [Telkomsel digital advertising grows significantly] https://mix.co.id/marcomm/brand-communication/telkomsel-digital-advertising-tumbuh-signifikan/

Smart Nation (2023, 28 Feb). Implication if civil servants use ChatGPT (PQ reply by SMS Janil Puthucheary), Fourteenth Parliament of Singapore – 28 Feb 2023, First Session. https://www.smartnation.gov.sg/media-hub/parliament/28022023/

Singer, N. (2017, 13 May). How Google took over the classroom. The New York Times. https://www.nytimes.com/2017/05/13/technology/google-education-chromebooks-schools.html

SMRC (2018, 25 Jan). Share of respondents in Indonesia who are of the opinion that homosexuality is forbidden by religion, as of December 2017 [Graph]. Statista

Statistics Indonesia. (2023, 31 Aug). Share of the rural population owning a mobile phone in Indonesia from 2013 to 2022 [graph with description]. Statista

Stenvall, J., Laitinen, I., Yeoman, R., Thompson, M. & Mueller Santos, M. (2022). Public values for cities and city policy. Palgrave Macmillan

Strauss, D. & Foster, P. (2023, 23 Nov). Can AI improve UK public sector productivity? The Financial Times. https://www.ft.com/content/86a5c3d8-e980-42b0-a1a0-48c3a1306367

Srnicek, N. (2016). Platform capitalism. Polity Press

Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. Big Data & Society, 4(2), pp. 1-14. https://journals.sagepub.com/doi/full/10.1177/2053951717736335

Teale, C. (2020, 13 May). 'Techlash' at Sidewalk Labs could mean smart city rethink. https://www.smartcitiesdive.com/news/sidewalk-labs-techlash-toronto-community-engagement/577799/

Teng, R. (2022, 4 Aug). Power to the people: Making digital services citizen-centric. GovInsider. https://govinsider.asia/intl-en/article/power-to-the-people-making-digital-services-citizen-centric

Teo, J. (2023, 13 Jul). Healthier SG care plans will be more personalized over time with use of apps: Ong Ye Kung. The Straits Times

The Culture Factor (2023). Country comparison tool. https://www.hofstede-insights.com/country-comparison-tool?countries=indonesia%2Cmalaysia%2Csingapore

Thouvenin, F. & Tamò-Larrieux, A. (2021). Data ownership and data access rights: Meaningful tools for promoting the European Digital Single Market? In M. Burri (ed), Big Data and Global Trade Law (pp. 316- 339). CUP

Tong, X.X. & Tay, E.S. (2023). Relevance of MySejahtera application in post-pandemic era: Legal regulations on data ownership and privacy. In Y. C. Adam and A. W. Abdull Manaf (eds.), Proceedings of the International Conference on Law and Digitalization (ICLD 2022), pp. 110–122. Atlantis Press

Torfing, J., Sørensen, E., & Røiseland, A. (2019). Transforming the public sector into an arena for co-creation: Barriers, drivers, benefits, and ways forward. Administration & Society, 51(5), pp. 795–825

Van Daalen, O. (2023, 21 Mar). Presentation at the discussion series about Generative AI 'The impact we generate' organized by the AI Media and Democracy Lab

UNICEF. (2022). Digital Transformation in the East Asia and Pacific region. https://www.unicef.org/eap/media/14086/file/%20EAPR%202022%20DX%20Annual%20Report.pdf

Van der Ploeg, I. (2012). The body as data in the age of information. Routledge

Van Zeeland, J. (2019, 12 Feb). Data is not the new oil. Towards Data Science (Medium). https://towardsdatascience.com/data-is-not-the-new-oil-721f5109851b

Veale M, Zuiderveen Borgesius F. (2022). Adtech and Real-Time Bidding under European Data Protection Law. German Law Journal, 23(2), pp. 226-256

Verhulst S. G. (2023). Operationalizing digital self-determination. Data & Policy, 5, pp. e14-1–17

Viljoen, S. (2021). A relational theory of data governance. Yale Law Journal, 131(2), pp. 573-654

Walsh, T. (2022, 13 Dec). Everyone's having a field day with ChatGPT – but nobody knows how it actually works. The Conversation. https://theconversation.com/everyones-having-a-field-day-with-chatgpt-but-nobody-knows-how-it-actually-works-196378

Walters, Trakman & Zeller (2019). Data protection law: A comparative analysis of Asia-Pacific and European approaches. Springer

Warzel, C. & Ngu, A. (2019, 10 Jul). Google's 4,000-word privacy policy is a secret history of the Internet. The New York Times. https://www.nytimes.com/interactive/2019/07/10/opinion/google-privacy-policy.html?

Wiggers, K. (2023, 15 Mar). Interview with OpenAI's Greg Brockman: GPT-4 isn't perfect, but neither are you. TechCrunch. https://techcrunch.com/2023/03/15/interview-with-openais-greg-brockman-gpt-4-isnt-perfect-but-neither-are-you/

Williamson, B. & Hogan, A. (2020). Commercialisation and privatisation in/of education in the context of Covid-19. Education International. https://eprints.qut.edu.au/216577/

Williamson, B. (2021). Meta-Edtech. Learning, Media and Technology, 46(1), pp. 1-5

WIPO (2023). Predicting subjective recovery from acute events using consumer wearables. WO/2023/044052. https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2023044052

Wirtz, B.W., Weyerer, J.C. & Sturm, B.J. (2020). The dark sides of artificial intelligence: An integrated AI governance framework for public administration. International Journal of Public Administration, 43(9), pp. 818-829

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. Regulation & Governance, 12, pp. 505–523

Yeung, K. (2022). The New Public Analytics as an emerging paradigm in public sector administration. Tilburg Law Review, 27(2), pp. 1–32

Zahiid, S.J. (2022, 27 Mar). Anwar raises concerns over planned MySejahtera sale to firm allegedly owned by cronies. Malay Mail. https://www.malaymail.com/news/malaysia/2022/03/27/anwar-raises-concerns-over-planned-mysejahtera-sale-to-firm-allegedly-owned/2049822

Zhang P., Godin S.D., Owens M.V. (2019). Measuring the validity and reliability of the Apple Watch as a physical activity monitor. Journal of Sports Medicine and Physical Fitness, 59(5), pp. 784-790

Zhuo, T.Y., Huang, Y., Chen, C. & Xing, Z. (2023, 22 Feb). Exploring AI ethics of ChatGPT: A diagnostic analysis. ArXiv preprint, https://arxiv.org/pdf/2301.12867.pdf