Personal data empowerment: restoring power to the people in a digital age

Background Paper

Kathleen McGowan, Priya Vora, Matthew Homer and Jonathan Dolan

Blavatnik School of Government
University of Oxford

Pathways for Prosperity Commission
Technology & Inclusive Development

Kathleen McGowan, Priya Vora,
Matthew Homer *and* Jonathan Dolan

Background Paper 11
September 2018

The Pathways for Prosperity Commission on Technology and Inclusive Development is proud to work with a talented and diverse group of commissioners who are global leaders from government, the private sector and academia. Hosted and managed by Oxford University's Blavatnik School of Government, the Commission collaborates with international development partners, developing country governments, private sector leaders, emerging entrepreneurs and civil society. It aims to catalyse new conversations and to encourage the co-design of country-level solutions aimed at making frontier technologies work for the benefit of the world's poorest and most marginalised men and women.

This paper is part of a series of background papers on technological change and inclusive development, bringing together evidence, ideas and research to feed into the commission's thinking. The views and positions expressed in this paper are those of the author and do not represent the commission.

Citation:
McGowan, K., Vora, P., Homer, M ., and Dolan, J. (2018). *Personal data empowerment: restoring power to the people in a digital age.* Pathways for Prosperity Commission Background Paper Series; no. 11. Oxford, United Kingdom

www.pathwayscommission.bsg.ox.ac.uk
@P4PCommission
#PathwaysCommission

Cover image © Ishtan Tankha

FUTURE STATE

BLAVATNIK SCHOOL OF GOVERNMENT
UNIVERSITY OF OXFORD

Pathways for Prosperity Commission
Technology & Inclusive Development

Executive Summary

Data is the new frontier in the fight against poverty. Personal digital data is growing at exponential rates and, as such, offers an important new opportunity for enhanced decision-making at all levels. If governed well, data can close capability gaps in government, foster small business growth, and empower people to access life-enhancing services.

Yet the digital revolution could just as easily, and perhaps more naturally, exacerbate exclusion and inequality by fuelling jobless growth, deepening discrimination, undermining trust in critical institutions, and eroding social norms through breaches of data privacy and targeted disinformation campaigns. Both trends – the positive and negative – are playing out simultaneously. Ultimately, determining the mix of rules and investments required to harness personal digital data for the benefit of individuals and society, while mitigating the risks, requires balancing the interests of the state and the market. This is because the power dynamics of the data economy tend to favour market interests against those of the individuals who are collectively generating the data.

Different countries are taking very different approaches to personal data governance. Indeed, most governments have yet to choose a strategic direction to guide their emerging data economies. This paper examines how four countries are approaching the governance of personal digital data and the implications of those models on the ability of data to enhance the capabilities of the state, market, and individuals:

1. China: data supports socio-economic and political objectives of the state
2. Estonia: data enables citizens to more easily access public services
3. India: data empowers individuals and fuels marketplace competition
4. US: data facilitates corporate growth and innovation

Each model has emerging weaknesses and risks, and these nascent approaches have yet to fully reveal their impact on individual or societal wellbeing. However, we hope this paper provides policymakers with insights to navigate choices around where power to use data can be bestowed and how to facilitate such a vision. In evaluating four different country cases we highlight three important lessons:

1. The natural dynamics of the data economy skew power away from the very people who generate the data;
2. As such, the role of government is critical in ensuring that individuals can fully benefit from their personal data endowments; and,
3. The role of government goes beyond establishing policies, but also in designing technology systems and tools to make policies effective and data rights meaningful.

The transformational potential of personal digital data in offering new pathways out of poverty deserves much more attention. We call for further research around how data can serve as a new input into models of inclusive growth. We also encourage action to develop practical solutions for how data can bolster effective state institutions and individual empowerment.

Table of Contents

Introduction: from digital access to personal data

By the end of 2018, there will be more than 4 billion internet users globally, an increase of approximately 1 billion since 2015. According to Cisco's latest Visual Networking Index, there will be 3.5 networked devices per capita globally by 2021[1] and some estimates suggest that connected devices could grow to 125 billion by 2030 – an annual average growth rate of 12%.[2] This growth has been driven, in large part, by increasingly affordable smartphone and data plans as well as innovations in broadband access technologies. Importantly, some of the regions where internet usage has lagged historically are now seeing some of the fastest growth rates. Africa, for instance, is experiencing 20% year-on-year growth, and a number of countries in the region, including Benin, Sierra Leone, Niger, and Mozambique, have doubled internet usage in the last year.[3] The expanding population of internet users, as well as the increasing amount of time each user spends online daily, is dramatically increasing the amount of data produced by individuals.

Yet, this is only a sliver of digital transformation story. The rapid growth of the Internet of Things (IoT) and machine-to-machine communications will also play a highly formative role in what digital transformation means for individuals and their data. Data is more profuse than ever before and is increasing exponentially. The oft-cited 2016 IBM report, *10 Key Marketing Trends for 2017*, found that 90% of all data had been generated in 2015 and 2016 alone. Recent estimates suggest that more data was created in 2017 than in all previous years combined. These trends will only accelerate: fixed internet traffic is expected to more than double between 2017 and 2021, while mobile internet traffic is projected to more than quadruple over the same period. These global figures translate to 35 gigabytes of internet traffic per person by 2021 – a threefold increase since 2016.[4]

Within this sea of data, individuals are generating rich data histories about themselves, and their lives are increasingly digitally documented and analysed. Collectively, this produces personal digital data,[5] that can fall into three distinct categories: government data, regulated data, and private/social data. Government data includes tax details, licences, land rights, and other documentation or engagement with government bodies. Regulated data includes mobile phone call detail records,

[1] Cisco Visual Networking Index (VNI): Forecast and Methodology, 2016-2021. (Updated 15 September, 2017). Retrieved from www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html#_Toc484813970

[2] Howell, Jenalea. (24 October, 2017). Number of Connected IoT Devices Will Surge to 125 Billion by 2030, IHS Markit Says [press release]. Retrieved from https://technology.ihs.com/596542/number-of-connected-iot-devices-will-surge-to-125-billion-by-2030-ihs-markit-says

[3] Kemp, Simon. (30 January, 2018). Digital in 2018: World's Internet Users Pass the 4 billion mark. Retrieved from https://wearesocial.com/blog/2018/01/global-digital-report-2018

[4] Cisco, Visual Networking Index 2017

[5] Emerging approaches to governing personal data tend to focus on any data relating to individuals. For example, the EU's General Data Protection Regulation (GDPR) refers to personal data "as any information relating to an identified or identifiable natural person." California's new law refers to personal information as "information that identifies, relates to, describes, is capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household." India's draft law refers to personal data as "data about or relating to a natural person who is directly or indirectly identifiable, having regard to any characteristic, trait, attribute or any other feature of the identity of such natural person, or any combination of such features with any other information."

medical records, banking and insurance, and data from other sectors for which there is a distinct regulator. Private data includes browsing history, social media data, airline history and other data generated between an individual and a private provider.

Figure 1: Unlocking personal data

For the purpose of data governance, the distinction between these three kinds of personal data is important because the speed and method through which policy can influence each category is different. While government can assert its power over each category, it is far easier in terms of time and process to make government data available than it is for the other types of data.

Ultimately, the story of digitisation and its impact on people will not be about access and use of technology alone. It will also be a story about the rise of personal data as a potential socio-economic asset or liability for individuals and communities. This story is still being written in every country as policymakers grapple with fundamental questions:

- How should personal data be defined?
- Who has access to and control of it?
- How should personal data be valued?
- How can it be used?
- How do historically poor populations benefit from being increasingly data rich?

The answers to these questions will define whether personal data becomes a new driver of, or a constraint to, inclusive development.

Section I: Assessing the personal data opportunity

In theory, individuals are well-positioned in the data value chain to exert influence on how personal data is used. Individuals generate the supply of personal data – social media activity, browser histories, derived data – that fuels the data economy; individuals also represent significant demand for the products and services that make personal data so valuable – online ads, Facebook feeds, and so on. Of course, the data economy's value chain is complex and individuals are not the only actors at either end. The creation of personal data would not be possible without the platforms provided by technology companies, and governments and businesses also represent significant demand drivers.

The dynamics of data economies tend to skew power away from the individuals who generate personal data and toward platform companies and/or the state. This will continue to be the case unless there is government action to:

- recognise individuals' rights to their data
- create foundational technology infrastructure (such as payments, identity, data management)
- create tools that maximise individuals' ability to benefit from their data use.

Nevertheless, with the right technologies, policies, and tools in place for people to manage their personal data, individuals do have a significant opportunity to leverage their data as an asset. When this happens, personal data will become a critical driver of inclusive development, increasing access to economic opportunity and enabling individuals to participate more directly in the data-driven society. The following examples illustrate ways that personal data is already helping to empower individuals:

- **Financial services:** In China, personal data histories, such as mobile phone records and consumer behaviour, are helping individuals demonstrate their creditworthiness and gain access to lending in order to start or grow new businesses.
- **Health services:** In Estonia, the ability to access personal data digitally is helping individuals better manage their health by creating a single, consistent health history that can be shared with all healthcare providers. It is also improving access to supporting services like filling in prescriptions online, which is done nearly 100% of the time in Estonia.
- **Social inclusion:** Traditionally poor, marginalised populations have had limited voice in society. Yet, as affordable smartphones and data plans proliferate, many of these same people will become data-rich. Agency over that data presents an opportunity to more forcefully represent their wants and needs. In India, the Aadhaar card, which provides a unique digital ID based on biometric data, is now used by nearly all of India's 1.3 billion people. For instance, it offers opportunities for efficiency in delivering social welfare benefits.

Analysis of personal data in the aggregate represents an important opportunity for broader social impact as well. The ability to mine healthcare or financial databases can equip policymakers and service providers with information on economic trends, population movements, and so on. This can contribute to better-informed policies and product innovations which, in turn, can complement individual empowerment through personal data.

However, without the right policies in place and the appropriate technological complements, personal data can be more a liability for individual empowerment than an asset. In addition to the oft-cited concerns about security breaches, the rise of data, if not governed well, can create inequities in participation, agency and choice. These risks are described as part of the framework that follows.

A framework for assessing governance of personal data

Despite, or perhaps because of its promise as a potential driver of socio-economic growth, personal digital data is increasingly the subject of complex and urgent questions regarding data ownership, security, ethics and much more. This paper explores different means by which four countries – China, Estonia, India and the US, – have approached the issue of personal data rights and the underlying technological solutions to facilitate their respective visions – their approaches to data governance. The four case studies analyse the power dynamics of the data economy through the perspective of individuals. The case studies serve as points of comparison from which to draw lessons learned and identify risks.

For purposes of this paper, we define data governance as: the combination of policies, technological systems and tools that determine how personal data is treated; and how they affect an individual's ability to participate in the data economy, exercise agency over their data, and make meaningful choices with regard to service providers.

> I. Participation refers to the extent that people are active users of digital tools. Participation is driven by factors of availability, accessibility, affordability, value and trust. High availability, accessibility and affordability of digital tools can lead to high levels of participation. High trust and value of digital tools can also support and sustain participation. The greater the sustained participation in using digital tools, the more data is generated and ultimately derived, making it increasingly important to define the vision of agency over one's personal data.
>
> Low participation risks the growth of "data deficits" that exacerbate existing inequities between those communities that are digitising and those who are not. Connectivity-constrained communities with low digital literacy risk becoming a new "digital underclass" that lags still further behind connected communities. Digital technology can also deepen existing discrimination through data and tools that enable more precise segmentation of consumers. Without effective strategies to ensure that the benefits of connectivity and participation in the digital economy are broadly available to all in society, digitisation will have the perverse effect of causing deepening socio-economic inequities.

II. Agency refers to individuals having meaningful control over their personal data. Understanding agency in digital economies involves the consideration of bestowing to individuals the *right* to access and make decisions over how their personal data is shared and used. Importantly, it also involves the *means* by which individuals are able to consent to the use of their data, revoke consent for the use of their data, track the use of their data, and move their data from one provider to another.

Low levels of individual agency over personal data are directly related to the growing concern around privacy and the misuse of data. With data increasingly concentrated with a smaller set of actors, these commercial and government 'data lakes' become targets for malicious attacks. Also, the concentration of data makes it easier for companies or governments to surveil the population.

III. Choice refers to how competitive a digital marketplace is and whether it enables new entrants.

Choice is naturally constrained by the monopolistic tendency inherent to digital platform businesses that results in the emergence of "data dominators," incentivised to hoard rather than share data. By design, commercial platforms provide products to consumers and capture the data generated by service use. Data use patterns tend to initially be used to target ads or improve the platform's services, thus creating a self-reinforcing cycle where more use of a platform creates more value in using the platform. Because digital platform businesses thrive on collecting and analysing data, they are pursuing ever-new strategies to expand their data reserves. Facebook and Netflix develop tactics to keep customers on their platform longer, while Amazon and Google venture into vastly new product lines to broaden the data they collect. This appetite for data is only growing as artificial intelligence makes user data more valuable, which in turn makes the platform's algorithms more powerful. In the US and China, this has manifested itself in entire industries being quickly dominated by a handful of commercial platforms. In their dominance, these companies have the power to suppress competition and result in potentially dangerous concentrations of economic and political power. Because data is a strategic asset to any business, there is a natural incentive to resist efforts to open or share data for the sake of competition, consumer rights, or security.

Extra risks for emerging economies

There are unique factors in developing countries that exacerbate risks in those markets. For example, data deficits are more likely to emerge in economies or communities that have low purchasing power and, therefore, represent the least attractive segments for commercial players to serve. Indeed, of the more than 3 billion people not yet online, the mobile and technology community actively talk about reaching only the next 1 billion. Research shows that those online have greater access to information, jobs, and markets to improve their social and economic standing. Thus the persistent gap between those online and those offline is exacerbating and reinforcing inequities.

Developing countries also have weaker institutions not well-poised to keep up with technology trends. It is clear from the India experience that technology, even when designed for inclusion as a public good, must be partnered with: policy, law and regulation to establish consumer protection and privacy; enforcement mechanisms to enforce the rules; and a thriving civil society to keep the government in check.

Consumers in developing countries are structurally more vulnerable to data capture and over-consent. Low-income consumers in particular use tools and apps that are more susceptible to data capture. The cost of computers and scarcity of reliable power leads more people to access the internet through a mobile device than through a desktop. In India, 80% of users access the internet through a mobile channel. In Africa, where internet access is significantly lower, 64% of users rely on a mobile device for internet access. The mobile device, whether handheld or tablet, has revolutionised the cost and convenience of accessing the internet. However, this has simultaneously led to a user interface dominated by apps. Apps such as Alibaba, WhatsApp and Facebook have become portals through which a user can access a variety of services, generating further data for the parent company. In contrast, desktop users can have more fragmented approaches to accessing services.

Section II: Global landscape of data governance

A number of factors have brought into focus the need for clearer data governance policies and regulations, including:

- the rapid expansion of internet access into new markets and with new populations
- the exponential growth of connected devices
- emergent public demand for greater transparency and accountability in the management of individuals' data.

There is also an appreciation of the current opportunity to rethink how personal data can be practically managed.

This is a quickly evolving area, and no country has established a single, comprehensive vision for how to govern individuals' data. Even in countries where data use by the state, industry, or by citizens is advanced, there are fragmented approaches to overall data governance. For instance, the World Bank estimates that 50% of countries that have a national ID card system lack any data protection legislation.[6] Similarly, nearly 80% of the UN Conference on Trade and Development's (UNCTAD) member countries have electronic transaction legislation while only half have consumer protection legislation. The global landscape is highly diverse in terms of data governance approaches and priorities: 107 countries have some data protection and privacy laws; but fewer than 40% of countries in Africa and Asia have such laws.[7]

Despite this diversity, in more advanced data economies, four distinct paradigms of personal data governance are discernible, each with very different implications for the natural asymmetries of power and information inherent to data economies.

- One model, which can be described as *activist*, focuses primarily on preventing harm caused by the misuse of personal data. This approach, exemplified the EU's General Data Protection Regulation (GDPR), which came into force in May 2018, shifts the burden for maintaining the privacy and security of personal data from the users of online services to the providers of the services. This approach establishes strict rules for the ways firms collect and handle personal data, and sets out steep fines for violations. While a potentially powerful deterrent to the intentional or accidental misuse of personal data, this solution assumes that state institutions have the capacity to police and enforce the rules. While this capacity may be sufficiently developed in the EU and elsewhere, many governments in the developing world lack the institutional 'teeth' to enforce strict rules. Indeed, many countries simply do not have the potential market size to make the compliance burden of GDPR-like frameworks cost-effective from the perspective of

[6] Domingo, A., Segovia, I. and Enriquez, AM. (2018). Digital Identity: The current state of affairs [working paper]. BBVA Research, p. 28.

[7] UN Conference on Trade and Development (UNCTAD) Cyber Law Tracker: Overview. (2018). Retrieved from: http://unctad.org/en/Pages/DTL/STI_and_ICTs/ICT4D-Legislation/eCom-Global-Legislation.aspx

service providers. While this approach establishes accountability and some transparency in data use, shifting some power over personal data back into the hands of the person who generated the data, it does not effectively address the opportunity to 'unlock' data to spur innovation.

- A second model, predominant in the US, applies little or no regulation to commercial service providers on the internet. This *laissez-faire* or market-centric approach allowed for the rapid growth of the internet economy and high rates of participation. However, the absence of a holistic governance framework has left personal data vulnerable to misuse. It has also absolved internet service providers of taking responsibility for the use of customers' data by third parties. By effectively allowing service providers to 'lock away' their customers' personal data, the provision of major services has quickly become dominated by a handful of US mega-companies. Individuals are left with increasingly fewer alternatives as these data dominators stifle competition in the market.

- China's *statist* approach to governing the internet and the data produced by users inside the country illustrates a third model: one in which the state benefits significantly from the rapid digitisation of the country's economy and society. Through strict control and censorship of online content, and strident protectionism of domestic internet companies, China has effectively cordoned off its cyberspace from the larger World Wide Web, enabling data produced online to be used to further the government's social, political and economic objectives.

- A fourth model seeks to build on the activist approach by bestowing individuals with rights to their data as well as fostering systems, protocols and tools that enable them to access and use the data they produce for their own benefit. This approach of data *empowerment* seeks to shift power in the data economy back to the individuals who generate the data and drive innovation. It does this by 'unlocking' personal digital data and enabling users to safely and practically share their data with service providers. The contours of the data empowerment model are emerging most clearly in India, where the draft data protection laws envision GDPR-like protections and rules. A complementary suite of digital platforms and tools designed for the public good form the foundation for broad, user-centric participation in the country's growing data economy.

Section III: Personal data governance case studies

Table 1: Case study comparison

	China	Estonia	India	US
Genesis	Socialist market economy opened door to internet → liberalisation of telecom market → rapid proliferation of digital tech → statist approach to preserve internet sovereignty	Post-Soviet overhaul of systems → introduction in parallel of X-Road platform and Digital ID → political buy-in to principles governing data society	Welfare reform → Digital ID → broader public goods digital infrastructure → Data Empowerment and Protection Architecture (DEPA)	Federalist model for overall governance → early leader in information society policymaking (eg Privacy Act of 1974 → advocacy for multi-stakeholder internet governance → data governed by market forces and info privatisation
Objectives	Maintain internet sovereignty to exert social, political, and economic control; leverage state control of data	Protect and empower users with regard to their data, provide efficient e-government services, and ensure interoperability of distributed databases	Protect and empower users with regard to their data, while preserving prerogatives of the state	Protect private sector agency in internet
Policy approach	Protectionist policies, evolving privacy standards which will increase expectations on private sector but unlikely to apply to government	Application of domestic digital principles (once-only, interoperability) within EU policy framework	Rights-based, with exceptions for state, emerging data sovereignty	Complex patchwork of state and federal laws, predominantly focused on data privacy and security more than consumer agency
Technology approach	Fragmented, lacking shared information platforms or national unique ID system	Standards and tools for users	Standards and tools for users	Fragmented, lacking shared information platforms (with a few regulated data exceptions)
Types of data addressed in model	Government - Yes Regulated - Yes Private - Yes	Government - Yes Regulated - Yes Private - No	Government - Starting Regulated - Eventually Private - Partially	Government - Yes Regulated - Yes Private - Yes
User participation	Barriers to inclusion are low but central state control of data limits dynamic participation	Nearly universal for some key services (eg tax filing) and growing	Rapidly growing as a result of Public Goods Digital Infrastructure	Barriers to inclusion are low but complexity of policies and institutions that govern data presents challenges to dynamic participation

User agency	Trust and user-centric policy and technology are limited as a result of statist model	State-run user portal to access and manage personal data generated by government, and increasingly regulated data from specific sectors	Intermediary approach combined with user tools	User-centric policies limited to specific regulated data, particularly in the financial sector – Consumer Financial Protection Bureau (CFPB) Consumer Protection Principles and ability to access credit reports. Limited user-centric policies or technology for government data and private data
User choice	Limited by data dominator firms and centralised state control of data	Focus on e-government services has limited impact on consumer choice	Public goods digital infrastructure supports development of new products and services	Limited by data dominators, some emergent private sector solutions
Concerns/risks	Overreach of the state, particularly surveillance and censorship, data protectionism	Model has limited risks to individual, but has not translated to broader inclusive development	Data protectionism, enforcement capacity, overreach of the state	Inconsistent model limits individual agency, power skews to data dominators with exception of a few regulated industries

Case Study 1: China

Problem and raison d'etre: The internet's arrival in China in 1994 was a natural consequence of the government's move towards a socialist market economy in the 1980s and early 1990s. While this socio-political transition increased economic freedom and opened new markets, the Communist Party of China continued to actively protect against competing values and political ideologies, while aggressively promoting Chinese enterprises over foreign competitors. In 1997, as internet use continued to spread, the Ministry of Public Security issued comprehensive regulations guiding its use. Among other things, it prohibited internet use for cultivating resistance to laws or the government or undermining Party values and social order.

Foundations for data action: These regulations, paired with internet censorship and surveillance technologies, are the Chinese government's effort to exert internet sovereignty – the idea that each country has the right to control its domestic internet space. In China, this control focuses on three main goals: 1) social control – monitoring online speech and guarding against anti-government campaigns; 2) information control – requiring localising data and censoring information, particularly if it casts a negative light on the government; and 3) economic protectionism – preferencing Chinese companies that are subject to domestic regulations.

Since the initial regulations were issued in 1997, the efforts to control foreign companies and international content has been referred to as the 'Great Firewall'. Efforts to monitor and manage domestic internet use has been done under the authority of the Golden Shield Project, an integrated multi-layer network operated by china's Ministry of Public Security since 2003. Originally, the Golden Shield Project was envisioned to be a comprehensive database-driven surveillance system that could access every citizen's data and connect that with national, regional and local security data. However, the initial vision for the project evolved as the internet expanded faster than expected and the liberalisation of the telecommunications market brought about rapid changes in technology, complicating the integration of local, regional and national data. The project evolved: "from content control at the gateway level to individual surveillance of users at the edge of the network."[8]

Building on this statist model of internet governance, the Chinese government is currently seeking to leverage technological advances in big data and artificial intelligence to further the digital transformation of its socio-political systems such as its Social Credit System (SCS) launched in 2014 and expected to be nationwide by 2020.

SCS is more a fragmented ecosystem of initiatives unified by similar goals and strategies rather than a fully integrated tool for social control.[9] It is representative of the Chinese government's aspirations to leverage data in all of its major development efforts, and demonstrates the government's appetite for experimentation in how data can be used to strengthen the regime's objectives.

In many respects, the foundations of China's data governance regime are still a work in progress, and three foundational issues will go a long way in determining the path moving forward: data privacy regulations; the state's relationship with private data dominators; and investment in centrally managed information platforms.

Policy, regulation and institutions: Together, the Great Firewall and the Golden Shield Project underpin a model that is characterised by a high degree of state control over the digital ecosystem, how users can engage in it, and how their data is used. In many respects, these factors have conspired to create a Chinese internet that runs parallel to the rest of the world with, for instance, the number of Google, Amazon, and Facebook users dwarfed by their domestic Chinese counterparts in search (Baidu), e-commerce (Tmall.com, operated by the Alibaba Group), and social media (WeChat).[10]

Importantly for the Chinese data governance model, the dominance of the domestic market by Chinese firms has created a relationship of mutual dependence between the government and the private technology industry. Companies benefit from the protectionist policies of the Great Firewall but are increasingly expected to march in lockstep with the government's programmes. The government, meanwhile, is increasingly dependent on these private companies' platforms to access citizens' data and monitor users' behaviour online.

---

[8] The Great Firewall of China: Background. (2011). Torfox: A Standford Project. Retrieved from https://cs.stanford.edu/people/eroberts/cs181/projects/2010-11/FreedomOfInformationChina/the-great-firewall-of-china-background/index.html

[9] Creemers, R. (22 May, 2018). China's Social Credit System: An Evolving Practice of Control. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3175792

[10] Richter, F. (16 August, 2017). China's Parallel Online Universe [blog]. Retrieved from www.statista.com/chart/10706/online-services-in-china/

This interdependence is the defining characteristic for data rights and data usage in the country. The SCS is the most notable example of this relationship and how it shapes the government's approach to data. Chinese legal scholar, Rogier Creemers, describes the SCS as "a set of mechanisms providing rewards or punishments as feedback to actors, based not just on the lawfulness, but also the morality of actions, covering economic, social and political conduct".[11] The inputs into those mechanisms are both public and private and are applied to citizens, government officials and businesses. Ultimately, the Chinese government leverages government data (for example, tax records), regulated data (such as financial sector information), and private data (for example, social media information and networks) to further SCS and other such programmes.

On the public side, SCS has necessitated data-sharing across previously siloed branches of government. On the private side, while the central government provided a top-level design for the SCS programme, it also provided licences to eight private companies to experiment with social credit scoring in an attempt to identify and refine effective, scalable scoring models. Furthermore, private companies are increasingly signing memoranda of understanding with the government's National Development and Reform Commission (NDRC), which runs the State Credit Information-Sharing Platform, to establish joint rewards and punishments as well as sharing social credit data. As of September 2017, the NDRC claimed to have 37 such data-sharing agreements in place.[12]

Data privacy: Not surprisingly, SCS and other data-driven government programmes, such as Healthy China 2030 which seeks to build online diagnostic and treatment services using 'big data' and artificial intelligence, have increased public demand for data privacy and further clarity of individuals' data rights. Until recently, China has preferenced experimentation with new technologies and data collection tools like facial recognition over strict privacy and data protection laws. And, not surprisingly given its interdependence with the private technology sector, it had previously made only minor regulatory moves to protect individuals' data rights. However, the Personal Information Security Specification, which took effect in March 2018, is a key regulatory effort that responds to growing public concerns about their data. The specification addresses collection, storage, use, sharing, transfer and disclosure of personal data – detailing the thresholds for individuals' consent and puts into place greater restrictions on secondary uses of data. It remains to be seen exactly how the Personal Information Security Specification will be implemented and enforced and what its impact on China's overall data governance model will be. It signals a meaningful change in how much freedom companies have in handling individual data.[13] Europe's GDPR served as a model for the Chinese specification, but the China rules are intentionally designed to be more permissive, particularly on definitions of consent and exemptions to consent – giving the government space to continue its collaboration with companies in key fields like big data and artificial intelligence.[14]

---

[11] Ahmed, Shazed, A. and Lang, B. (6 April, 2018). Who's really responsible for digital privacy in China? [blog]. Retrieved from www.merics.org/en/blog/whos-really-responsible-digital-privacy-china

[12] ibid.

[13] Chorzempa, M., Triolo, P., and Sacks, S.. (June 2018). China's Social Credit System: A Mark of Progress or a Threat to Privacy? Peterson Institute for International Economics. Retrieved from https://piie.com/system/files/documents/pb18-14.pdf

[14] Sacks, S. (9 March, 2018). China's Emerging Data Privacy System and GDPR [blog]. Center for Strategic and International Studies. Retrieved from www.csis.org/analysis/chinas-emerging-data-privacy-system-and-gdpr

State/Data dominators relationship: In addition to the domestic pressures, including public demand for greater data privacy and regulatory efforts like the Personal Information Security Specification, the increasingly global aspirations of Chinese companies may reshape the cosy data-sharing relationship between the government and private data dominators. Partly in an effort to distance themselves from the Chinese state's efforts to collect data on its citizens, companies – including Alibaba, Tencent, and Baidu – have actively participated in developing and implementing domestic data protection policies. Despite this, as well as public commitments by these companies to meet GDPR standards, EU regulations are still determining whether Chinese firms are adequately meeting GDPR requirements.[15] The balance between global aspirations of Chinese firms and the goals of the Chinese government, may further determine what Chinese data governance looks like moving forward.

The implementation of the Personal Information Security Specification will further define the relationship between state and the data dominators. Currently, policy analysts suggest that the emerging concept of personal data privacy in China is more likely to apply only to the way companies handle personal data; the government will operate under a distinct framework governed by national and cybersecurity laws that would actually expand access to personal data.[16]

Technology architecture: Surprisingly, despite the significant government investments in data systems, the Chinese data governance model does not yet have a single solution for linking one's data footprint to a unique identifier. This fragments China's data landscape and represents a key challenge to rolling out SCS nationwide. While IDs are increasingly linked to digital tools like cell phones and apps, and identity data sets are currently held across many systems, physical ID cards are still required for many activities.[17]

Centrally managed information platforms: While efforts like SCS have started to push for cross-government data-sharing and have blurred lines between public and private data, the overall landscape of data management systems in China remains highly fragmented. National and local government agencies are often reluctant to share data as it can be a source of political influence, and many of the largest handlers of data – for example, Ministry of Public Security, NDRC, and the Bank of China – all maintain their own separate databases. Competition is also making companies increasingly hesitant to share valuable commercial data. Ultimately, there is no central government repository for handling data from industry and government.[18] This, combined with the lack of a unique digital identifier, represents a significant challenge to implementing a data governance model that would make it possible to fully realise the vision for programmes like SCS. As recently as June 2018, the Chinese State Council publicly reaffirmed its commitment to nationwide rollout of SCS by 2020.[19] Without the investment in such central platforms, there will always be a gap between the vision for China's data governance model and its practice.

---

[15] Chorzempa, Triolo and Sacks. (June 2018). China's Social Credit System: A Mark of Progress or a Threat to Privacy?

[16] Sacks, S. (25 April, 2018). What the Facebook Scandal Means in a Land without Facebook: A Look at China's Burgeoning Data Protection Regime [blog]. Center for Strategic and International Studies. Retrieved from www.csis.org/analysis/what-facebook-scandal-means-land-without-facebook-look-chinas-burgeoning-data-protection

[17] Chorzempa, Triolo and Sacks. (June 2018). China's Social Credit System: A Mark of Progress or a Threat to Privacy?

[18] ibid.

[19] China speeds up effort in building a social credit system [website]. (7 June, 2018). Retrieved from http://en.people.cn/n3/2018/0607/c90000-9468262.html

In November 2017, China's National Internet Finance Association (NIFA), the country's regulatory body for online financial services, launched a personal credit information platform to serve online lending firms. This was in an effort to fill the gap in financial information sharing and to cut down on fraudulent loan applications.[20]

Impact on user participation, agency and choice: Despite the incomplete story of the Chinese data governance journey, the statist model has led to significant scale in data usage. Nearly 34 million credit codes have been issued to newly registered businesses, 71 central government departments and provincial governments have been connected to the national credit information-sharing platform,[21] tens of millions of people have voluntary signed up for SCS, and nearly 10 million Chinese citizens have already been restricted from buying airplane and train tickets due to low social credit scores.[22] As these efforts continue to increase, and the relationship between public institutions, private companies and citizens evolves, the impact this data governance model has on individuals will change. However, a few conclusions can be drawn about its current and future forms, assuming that the state's vision for data usage is realised.

1. Participation: The extent that the Chinese data governance model enables individuals to participate in the digital economy relies on two primary factors:

- *Inclusion in the system:* While data systems in China remain fragmented, the barriers to inclusion – affordability and accessibility of services – appear low. Internet penetration continues to increase rapidly, with nearly 800 million users and smartphone ownership expected to reach 690 million by 2019. Inclusion in government programmes such as SCS have few to no barriers.

- *The ability to create and transact within the system:* The model is limiting. The government's central control of data, as well as the dominance of a few private actors in collecting data, does little to enable individuals to leverage their data for their own benefit. China ranked last in 2017 on Freedom House's Freedom on the Net ranking.[23]

2. User agency: While elements of the Chinese model promote participation, the two key elements of user agency – trust and control – are very low. The increasing public demand for stronger privacy protections and greater clarity on data usage are reflective of the low level of trust in the system. In many respects, SCS is held up by the Chinese government as a tool for repairing trust in a society that has a trust deficit, a system to create standards and set expectations in the relationship between government, business, and people.

[20] Tham, E. (27 November, 2017). China online finance regulator launches credit rating platform. Reuters. Digital.

[21] China speeds up effort in building a social credit system [website]. (7 June, 2018). Retrieved from http://en.people.cn/n3/2018/0607/c90000-9468262.html

[22] Karsten, J. and West, DM. (18 June 2018) China's social credit system spreads to more daily transactions. Brookings. Retrieved from www.brookings.edu/blog/techtank/2018/06/18/chinas-social-credit-system-spreads-to-more-daily-transactions/

[23] Freedom of the Net 2017. (June 2016 – May 2017). Freedom House. Retrieved from https://freedomhouse.org/report/table-country-scores-fotn-2017

However, the opacity of data collection, sharing and usage systems has raised significant public concern. As the Personal Information Security Specification takes effect, individuals may begin to have more clarity on their data rights. However, without consent, a more permissive approach to data use will continue to minimise the level of control individuals have over their data. Furthermore, as technologies like facial recognition become more pervasive, the extent of data collection and the passive way individuals create that data will only serve to reinforce their lack of control.

3. Choice: Much like Western markets, China's data landscape is characterised by a small number of private data dominators. This, combined with the state's efforts to tightly control data systems, significantly limits individuals' choice. While programmes like SCS are framed as serving 'public good' by the government, that doesn't translate to consumer choice in the statist data governance model.

Case Study 2: Estonia

History and context: The seeds of Estonia's digital transformation can be traced back to the early 1990s, following the collapse of the Soviet Union and restoration of the country's independence. During its first decade of independence, Estonian policymaking was shaped by two key political forces which would have a formative influence on the country's digital transformation: the desire to transition from the post-Soviet era through replacement of existing systems, instead of upgrades; and the desire to 'leapfrog' the West.[24] For a young government, emboldened by the ICT successes of its Scandinavian neighbours and endowed with a legacy of Soviet-trained telecommunications research and development experts, investments in technology represented a natural opportunity to realise these political goals and think aspirationally about the future. In many respects, policymaking during the 1990s was playing catch-up with fast-moving technological developments. However, to the government's credit, they cultivated the space for creativity and risk-taking and set the foundation for future innovation in ICT policymaking.[25]

The 1990s were characterised by aspirational thinking and technological advancement informing policy ex post. The 2001 introduction of two key pillars of Estonia's digital transformation – the data exchange layer known as X-Road, and the mandatory national digital ID – mark the beginning of a more intentional strategic direction that has turned the country into a global leader in the provision of e-government services. It was at this point that the government decided it should offer its own secure national platform for digital services instead of looking to private technology companies to do so.[26] At the same time, the government also started the national databases from scratch, providing systematic and unique numeric identifiers – including personal identification codes for citizens, business registration numbers for businesses, and land titling information management.

[24] Kattel, R. and Mergel, I. (2018). Estonia's digital transformation: Invisible vs hiding hand. Oxford University Press, forthcoming volume, p. 5-6.

[25] ibid.

[26] Heller, Nathan. 'Estonia, the Digital Republic' The New Yorker 18 and 25 December 2017.

Since joining the EU in 2004, Estonia has developed its data governance model within the context of EU rules and regulations. Data protection rules in the EU were previously set out in a Directive – a legal act to be incorporated into national law in all Member States – which Estonia complied with since joining the EU. In May 2018 the EU GDPR came into force. Regulations are directly applicable and binding on all Member States. This means that Estonia has rules on data protection almost identical to all other EU Members. Some procedural and other aspects shall be further clarified in national law, which means amendments to the Estonian data protection legislation. These amendments are still in the process of being drafted, but the GDPR as such applies fully in Estonia as in the rest of the EU. The main principles of the GDPR are that data shall be gathered only for specific purposes, in proportion with need, and based on legitimate grounds such as fulfilling tasks of public administration. Emphasis is put on the processes and responsibilities for data handling and it is in this context that changes are introduced, to strengthen oversight.

In recent years, Estonia has started to expand the boundaries of its digital transformation. This started in 2014 with the creation of its e-residency programme, which enables citizens of another country to become residents virtually and benefit from the digital services of Estonia. Estonia is now looking to export its X-Road infrastructure, particularly across Europe.[27] It also used its presidency of the EU from July to December 2017 to advocate for the idea of freedom of movement of data across the EU. The EU has four essential freedoms in its internal marketplace: free movement of goods, capital, people, and freedom to establish and provide services. Estonia has made the case that data movement become the fifth such freedom.[28]

Description of the model: The distributed nature of Estonia's data management is central to its governance model. In this model, the government's role is not to invest heavily in expensive, monolithic data systems, but rather to ensure coordination, convenience, security, and citizens' agency over their data. Each government agency, for instance, is responsible for developing their own ICT strategies and data systems. The central government manages the Population Database – which provides a single unique identifier for all citizens and residents – and issues identity cards that provide legally binding identity assurance and electronic signature capabilities.[29] The same approach applies to all core registries of the country – for example, businesses, land, vehicles, and so on).

In this decentralised model, government systems are not allowed to store the same data in more than one place. It is prohibited by law to create a public database with data that already exists in another database. In the case of basic personal data, for instance, it doesn't need to be held anywhere other than the Population Database. Other systems simply need access to the unique identifier data in the Population Database. The distribute data model also provides some level of data protection as no single place can hold all of an individual's data, greatly lowering the risk of massive breaches seen elsewhere.[30] The distributed architecture also allows each government agency to develop ICT systems that meet their unique needs.

[27] Heller, N. 'Estonia, the Digital Republic' The New Yorker 18 and 25 December 2017.

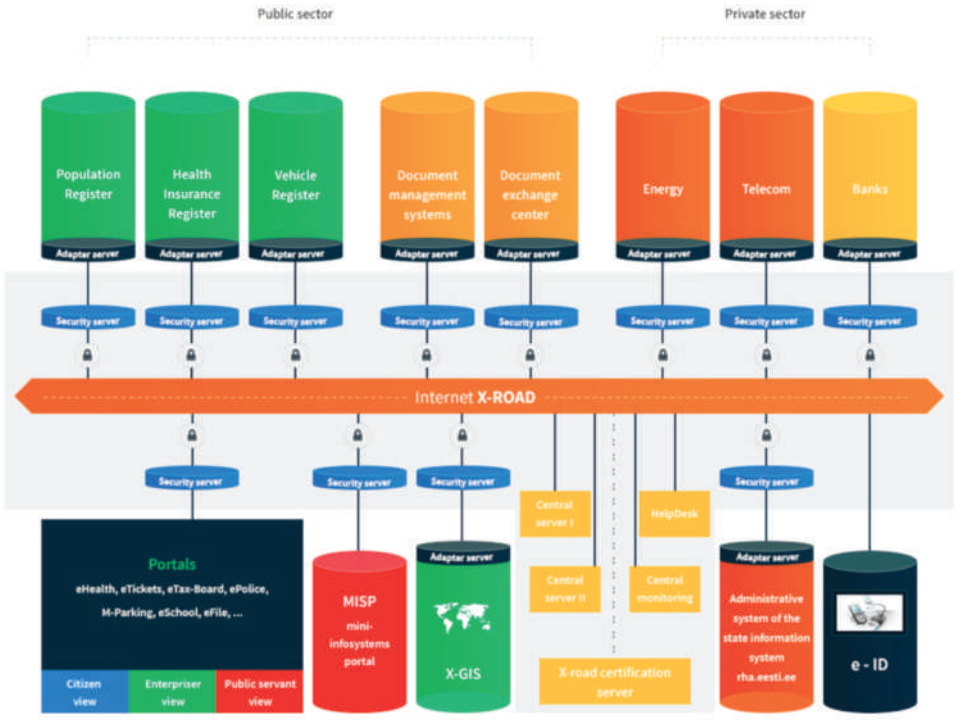
[28] E-Estonia. (2017) Estonia – making the case for the free movement of data. September 2017. Retrieved from https://e-estonia.com/estonia-making-the-case-for-the-free-movement-of-data/

[29] Herlihy, P. (31 October, 2013). Government as a data model: What I learned in Estonia [blog post]. Retrieved from https://gds.blog.gov.uk/2013/10/31/government-as-a-data-model-what-i-learned-in-estonia/

[30] ibid.

Of course, this distributed model would lead to highly fragmented systems without the central government's investment in X-Road. In X-Road, the government created a platform for interoperability between decentralised databases and a data exchange layer that can be used by public and, increasingly, private entities. Ultimately, X-Road provides two critical functions: interoperability; and enforceable data standards to ensure the security of the system.[31]

Figure 2[32]



This model has a number of positive implications for businesses and citizens:

• *Only-once policy:* business and citizens only have to supply their information once to government agencies and participating businesses. The X-road enabled interoperability, paired with the digital ID card, creates a system where, for instance, individuals do not need to prepare a loan application because all of their existing financial data – income, debt, savings – are accessible and verifiable to a potential lender. The same is true for health information, where healthcare providers can access data on a patient's medical history.[33]

• *Data management and privacy:* A core principle of the Estonian system is that individuals own all the personal data they produce or that is recorded about them. Citizens and residents can access all of their personal data online through the State Portal. There are more than 2,000 services integrated through X-Road and more than

[31] Kattel, R. and Mergel, I. (2018). Estonia's digital transformation: Invisible vs hiding hand. Oxford University Press, forthcoming volume, p. 3

[32] Herlihy, P. (31 October, 2013). Government as a data model: What I learned in Estonia [blog post]. Retrieved from https://gds.blog.gov.uk/2013/10/31/government-as-a-data-model-what-i-learned-in-estonia/

[33] Heller, N. 'Estonia, the Digital Republic' The New Yorker 18 and 25 December 2017.

900 connected organisations, public registries and databases. You can log in to the portal using your ID card, and view all your data, correct errors, and manage who has access to what information.[34] For example, an individual can make a particular medical file accessible to some of his or her doctors while keeping it private from others. Also, each time an authority figure such as a police officer, doctor or government official, looks at an individual's secure data online, it is recorded and visible to the person concerned. Looking at an individual's data without a reason is a criminal offense.[35] Estonia has set up a Data Protection Inspectorate that acts as ombudsman and preliminary court to assess whether an individual's data protection rights have been violated, and to issue legally binding decisions.[36]

• *Data integrity:* Following a series of cyber attacks in 2007, Estonia became the first country to develop a blockchain solution at the national level. KSI Blockchain was designed in Estonia and deployed to ensure that no data could be changed or manipulated by anyone and that authenticity of data can be verified.[37] X-Road facilitates more than 500 million transactions per year (as of 2015), none of which have a supporting traditional paper trail. The ability to deploy blockchain or similar technologies to increase verifiability of data has contributed significantly to overall trust in the system. Estonia has also taken steps to create backup systems for added security, and creating a 'data embassy' in Luxembourg in 2017 that follows the same international laws as physical embassies.

Foundations of approach/preconditions: The design principles that characterise the Estonian model – namely the once-only principle, interoperability and security of distributed architecture, and the customisation of technology solutions (as opposed to buying off-the-shelf systems) – are made possible through the three foundational technology components (discussed above): X-Road; the compulsory national digital ID system; and the use of technologies like blockchain to ensure a strong focus on data integrity. The model has roots in strong political will and cross-sector networks that have developed (formally and informally) through the movement of talent between government and industry. Yet, the most significant guiding principles of the model are not codified as laws or regulations. Rather, the principles have been adopted and supported by political leadership.

This is not to say that the country lacks policy and regulatory frameworks for its digital transformation. Working within the relevant EU frameworks, it has consistently demonstrated an innovative culture in its policymaking. The current #Krattlaw debate – Estonia's digital shorthand for a new type of legal entity, including artificial intelligence, algorithms, and robots – is yet another example of Estonia looking to the future. The proposed legislation would enable algorithms to buy and sell services.[38]

---

[34] Herlihy, P. (31 October, 2013). Government as a data model: What I learned in Estonia [blog post]. Retrieved from https://gds.blog.gov.uk/2013/10/31/government-as-a-data-model-what-i-learned-in-estonia/

[35] Heller, N. 'Estonia, the Digital Republic' The New Yorker 18 and 25 December 2017.

[36] Jackson, E. (January 2015). The right mix: How Estonia ensures privacy and access to e-services in the digital age. Estonian World | How Estonians See It [blog]. Retrieved from http://estonianworld.com/security/right-mix-estonia-ensures-privacy-access-e-services-digital-age/

[37] E-Estonia. Safety and Security: KSI Blockchain. Retrieved from https://e-estonia.com/solutions/security-and-safety/ksi-blockchain/

[38] Heller, N. 'Estonia, the Digital Republic' The New Yorker 18 and 25 December 2017.

Impact and use: As of 2017, Estonia's digital transformation had resulted in approximately 350 million digital signatures from a population of only 1.3 million people. Nearly 100% of all medical prescriptions are done online, as are nearly all income tax declarations, and 30% of all votes – nationally and locally – are cast digitally. The country now counts 35,000 people as e-residents, and the government estimates that its digital infrastructure helps save 2% of GDP annually.[39] All this progress has translated into individual data empowerment:

1. Participation: Despite the notable progress made in the provision of e-government services, Estonia is still grappling with how the 'Digital Republic' can lead to a more inclusive society. For instance, Estonia has the highest gender pay gap in Europe and a Gini coefficient that is higher than the EU average.[40] While Estonia's e-government has flourished over the last two decades, reaching near universal adoption for some use cases (for example, tax declaration), the same cannot be said for ICT as an industry. There are, of course, notable exceptions, such as the Estonian team that helped to create Skype. However, for the most part, the ideas and creation process in the ICT space have reached only a small network of Estonian elites working across government and the private sector. In other words, the digital transformation in Estonia has focused on delivery of service as opposed to increasing the size of the pie for citizens to engage in the data economy as creators and collaborators. The Estonian ICT sector provides around 6.2% of the country's gross domestic product (GDP) and represents 12% of exports.[41] Also, the agency Estonian citizens have over their own data has saved time and added convenience. However, the country's digital transformation has been limited in creating broad-based opportunities for expanded economic participation.

2. Agency: The Estonian data governance model fosters a high degree of individual agency over data. Estonia's model has cultivated two key aspects of agency – trust and control. The successful provision of e-government services has been built on citizens' trust in the government's intent and ability to keep their information secure. With online tax declarations and medical services reaching near-universal adoption in Estonia, it is clear that the steps the government has taken – technically (X-Road and blockchain systems), legislatively (Personal Data Protection Act), and behaviourally (transparency in instances of security breaches) – has helped build that trust. User control of data, as discussed above, is likewise a key feature of Estonia's overall governance model. Individuals are able to access, correct and manage their data virtually, and all in one place.

3. Choice: Given that the story of Estonia's digital transformation is predominantly one of e-government service provision, the concepts of choice have been limited. The policymaking culture has been innovative, yet that has translated more narrowly into specific systemic innovations for data management (such as X-Road's facilitation of distributed databases) as opposed to cultivating a broader ICT innovation eco-system. Furthermore, the effective but narrow focus on e-government services has necessarily had a dramatic impact on increasing competition or expanding products offered in the marketplace.

[39] Kattel, R. and Mergel, I. (2018). Estonia's digital transformation: Invisible vs hiding hand. Oxford University Press, forthcoming volume, p. 3

[40] ibid.

[41] Statistics Estonia. (2018). Exports of services by economic activity. (August 2018). Retrieved from www.stat.ee/413465

Case Study 3: India

Problem and raison d'etre: Over the past decade, India has taken significant steps to enable broader participation in the digital economy. As a result, hundreds of millions of people are now generating personal data histories for the first time. However, the country has lacked a data governance framework for the digital era and is now trying to establish a framework that aligns policy, regulation, institutions and technology architecture. The goal is to establish a governance framework for personal data that balances the rights of individuals with those of the state. Technologists are working on a related effort to craft systems that safeguard privacy while 'liberating' data to fuel private sector innovation and improve government service delivery.

This effort is still in progress and there is a vibrant public debate about how personal data should be treated by the law. However, the emerging approach appears to be two-fold: (1) establish individual rights related to personal data, while also asserting rights for the state; and (2) add technology standards and protocols to policy so that consumers can actively assert the rights they are afforded by law.

Because of the country's new digital infrastructure, some in India are betting that individuals will be able to translate their data into greater economic prosperity. One of India's internet pioneers, Nandan Nilekani, believes that being "data rich" can unlock new pathways to prosperity for the poor. The data governance framework that India is building will test whether this theory will be true in practice.

Foundations for data action: The ability of hundreds of millions of Indians to now participate in the digital economy is made possible by several important developments:[42]

1. Digital identity has reached more than 1.2 billion people, and other foundational infrastructure designed for the public good (referred to as the "India Stack") has enabled many Indians to access the digital economy for the first time.
2. The growing availability of cheap mobile phones and network connectivity has resulted in approximately 1.2 billion mobile phone connections (with over 350 million smartphones) and 462 million internet users, although India imposes state-issued internet shutdowns more than any other country.[43]
3. The banking sector has tried to advance financial inclusion, with more than 582 million unique bank accounts now open.
4. Innovative services and products have created compelling use cases – such as ecommerce – for deeper engagement in the digital economy, including 375 million social media users.

[42] Varma, P. (9 February 2018). India's Platforms Leapfrog [presentation posted online]. Retrieved from www.slideshare.net/ProductNation/indias-platforms-leapfrog-by-dr-pramod-varma

[43] Segal, A. (22 August 2018). The Link Between More Internet Access and Frequent Internet Shutdowns [blog]. Council on Foreign Relations | Net Politics and Digital and Cyberspace Program. Retrieved from www.cfr.org/blog/link-between-more-internet-access-and-frequent-internet-shutdowns

Policy, regulation and institutions: India's emerging governance framework for personal data can be seen into two main areas of activity. The first is policy, regulatory and institutional activity, which most notably includes efforts to establish a personal data protection law. Under the framework that is emerging, personal data will be governed through specific individual rights as well as technology standards and tools that make it possible for them to easily exercise those rights. India's draft Personal Data Protection Bill, released in July 2018, represents India's attempt to establish such rights the digital age.[44] Among other things, it proposes a Data Protection Authority, establishes the concept of a data fiduciary, and gives individuals certain rights related to their data. The draft Bill gives people the right to: (1) confirmation and access; (2) correction; (3) data portability; and (4) be forgotten. Individuals are able to exercise these rights through data fiduciaries, which the draft Bill defines as "any person, including the State, a company, any juristic entity or any individual who alone or in conjunction with others determines the purpose and means of processing of personal data".

Technology architecture: A technology architecture is under development to complement whatever final data protection law emerges. The Data Empowerment and Protection Architecture (DEPA)[45] has two primary components: (1) an Electronic Consent Framework;[46] and (2) a Digital Locker Technology Framework.[47] Both efforts are housed in the Ministry of Electronics and Information Technology, but have been developed with close participation by India's technology community. According to those who have developed DEPA, the goal "is to break the tension between: (a) maintaining privacy and (b) using the data for good. Rather than having to balance between them, DEPA aims to provide a third option – enabling safe and trusted sharing of data in which privacy is preserved". The Electronic Consent Framework is intended to improve consent in the digital realm and ensure that individuals are able to gain more control over their consent, whereas the Digital Locker Technology Framework establishes standards and tools for users to gain access to and manage their data after they have provided consent.

Impact on user participation, agency and choice: The technology architecture being put in place envisions specific tools to enable individuals to make use of their data rights. This includes the concept of an 'account aggregator' that would intermediate consent when users engage with companies and provide them tools to actively manage their consent.[48] The framework also envisions licensing digital locker providers that will provide dashboards for users to manage their data. India's Ministry of Electronics and Information Technology is starting by making government data available through digital lockers, but also expects to extend this to data collected by regulated firms such as banks, and eventually to other forms of private sector data as well.

---

[44] The Personal Data Protection Bill (2018). Retrieved from http://meity.gov.in/writereaddata/files/Personal_Data_Protection_Bill%2C2018_0.pdf

[45] iSpirit. (August 2017). Data Empowerment & Protection Architecture (DEPA) [presentation posted online]. Retrieved from www.slideshare.net/ProductNation/data-empowerment-protection-architecture-depa

[46] Electronic Consent Framework: Technology Specifications, Version 1.1. Retrived from http://dla.gov.in/sites/default/files/pdf/MeitY-Consent-Tech-Framework%20v1.1.pdf

[47] Digital Locker Technology Framework: Version 1.1. Retrieved from http://dla.gov.in/sites/default/files/pdf/DigitalLockerTechnologyFramework%20v1.1.pdf

[48] Reserve Bank of India. (23 February 2018). Master Direction- Non-Banking Financial Company - Account Aggregator (Reserve Bank) Directions. Retrieved from www.rbi.org.in/Scripts/NotificationUser.aspx?Id=10598&Mode=0

Given that the foundations are still being put in place, current impact is difficult to ascertain. Yet, the potential impact of India's emerging governance framework for personal data can be considered by looking at the capabilities it is establishing across the following categories:

1. Participation: The ability for nearly all Indians to participate in the country's digital conomy has been made possible through newly established digital infrastructure that enables even poor and marginalized populations to participate. For India's emerging data governance approach specifically, the country's two-pronged strategy – consumer rights combined with technology tools and standards – has intentionally focused on making it possible for users to actively manage their data. On their own, consumer rights have the potential of being nothing more than unenforceable high-minded principles, but when combined with user-facing tools, they have the ability to equip consumers with the authority to exercise certain rights over their data, and also the capability to do so. This approach has the potential to enable users to participate much more actively in the management of their personal data.

2. Agency: While India's approach does not envision giving users exclusive ownership over their data, it does imagine users having much greater control in at least two important ways: when consumers provide consent for an entity to collect an or share personal data; and after their data has already been collected. Using these tools, the emerging framework envisions giving users the ability to manage their data wherever it is stored, and also to proactively use their data to improve their lives.

3. Choice: India's digital infrastructure has been designed with the intention that public and private sectors alike will build solutions that leverage it. The country's new payments protocol, the Unified Payments Interface (UPI), provides an example of this, with the protocols being available for others to use. In this instance, WhatsApp and Google have both launched popular products using India's UPI. This approach of digital infrastructure as a platform for innovation could result in greater user choice. Also, the ability of users to take their data and transfer it to whatever entities they wish could promote a marketplace with more options.

India's digital identity system (Aadhaar) has by far seen the most use of any aspect of the India Stack. This is likely due to a variety of reasons, including the fact that it has been made mandatory for many government services, and it has become increasingly difficult for Indian residents to participate in the digital economy without it. Compare the 1.19 billion Aadhaar users with the 4.8 million users of India's first Digital Locker, the DigiLocker. Or the fact that the UPI facilitated more transactions in 18 months than credit cards had in 18 years. This suggests that, while many Indians are now using the more foundational and transactional functionality of the India Stack, fewer are seeking to proactively manage their data. This could change as new electronic consent tools are rolled out, and as more compelling cases emerge for users to do so. However, it also shouldn't be surprising that the India Stack's more advanced functionality would be used first by early adopters, and take some time before being used more broadly, especially by those who are new to the digital realm.

Broader implications: How India's emerging efforts in this space will impact on users and other stakeholders will take some time to fully understand, particularly since the framework is still in development. The public discourse currently emphasises the rights of the state versus the rights of the individual, with the role of the private sector featuring less prominently. It is still too soon to assess the overall risks associated with this approach, but several factors are worth keeping in mind. The first is whether India can rely on consumer rights and tools alone. It will be important to carefully consider the role of the Data Protection Authority in enforcing consumer rights. The second is whether these approaches are enough to overcome consumer apathy and foster greater user concern over their data.

While still nascent, even pre-emergent in some instances, India's approach is to establish a set of data rights that balance the interests of the state and the individual, while simultaneously equipping consumers with tools to more proactively manage their data. If successful, this approach could help level the information and power asymmetries that exist between individuals and those who use their data.

Case Study 4: The U.S.

Problem and raison d'etre: At the beginning of the 21st century, the internet experienced a rapid transition from a tool for the military, academic communities, and specific commercial functions, to a platform for the masses to communicate, transact, and participate. During this transition, two competing ideas about how the internet should be governed came to the fore. On one hand, the US and European governments and the private sector advocated for a cross-sectoral, multi-stakeholder approach which would make consensus-based decisions. On the other hand, governments – notably China and Russia – advocated for an intergovernmental approach that would put most decision-making in the hands of national governments. During much of this period, the US government was in the unusual position of vocally supporting the multi-stakeholder approach, while retaining oversight of internet naming and addressing conventions. Ultimately, in October 2016, the US turned over that authority to the Internet Corporation for Assigned Names and Numbers (ICANN), a non-profit coalition of governments, private sector actors, and civil society as a way to demonstrate (among other things) its commitment to the idea that national governments alone should not control the operation of the internet.

This position, combined with its advanced domestic technology sector and the federalist structure of US governance overall, has had a formative impact on the data governance model in the US – a model characterised by information privatisation, deference to market forces, and a complex relationship between state and national policies and institutions.

Foundations for data action: From the 1970s through the end of the 1990s, as the data economy slowly emerged, the US Congress was a leading force globally in developing privacy and data security laws – notably (among many other laws) the Privacy Act of 1974. This Act governs the

collection, maintenance, use and dissemination of information about individuals that is maintained in federal agency systems. US Congress was also instrumental in the Health Insurance Portability and Accountability Act (HIPAA) of 1996, which included a privacy rule that established national standards to protect individuals' medical records and other personal health information.

However, from the early 2000s, and aligning with the era of mass adoption of the internet, legislative activity of the federal government in the US slowed. Most federal legislative action related to data privacy post-September 11, 2001, came in the form of amendments to existing laws aimed at expanding the government's right to access personal data. The Health Information Technology for Economic and Clinical Health Act (HITECH Act) of 2009, which strengthened HIPAA privacy protections, and the Consumer Financial Protection Bureau's Consumer Protection Principles are notable exceptions to that trend. Nevertheless, as the internet and digital services rapidly proliferated, states stepped in to play a formative role in shaping data governance in US.

Policy, regulation and institutions: The US model of data governance currently lacks a single, unifying national law regulating the management of personal data. Instead, the US model is characterised by a patchwork system of federal and state laws and regulations that can overlap, complement or contradict one another. In addition, many governmental agencies and industry groups have developed guidelines that do not have the force of law, but are part of self-regulatory 'best practices' frameworks. Such frameworks are increasingly being used by regulators for accountability and, in some instances, enforcement.[49]

Within this patchwork, there are a number of federal privacy laws that regulate the treatment of personal data but these are primarily focused on particular categories of regulated data, such as financial or health information. There are also broad consumer protection laws, including the Federal Trade Commission Act, that prohibit unfair or deceptive practices related to the security of personal information and the Children's Online Privacy Protection Rule for protecting children's online privacy.[50]

Historically, new laws and regulations have been reactive to technological advances or privacy breaches. However, gradually more states are developing prescriptive and preventative laws, particularly in the area of data security. California is a national leader in an overall data protection framework. It was the first state to issue a data breach notification law in 2003. The Massachusetts Regulation 201 CMR 17.00 is the clearest example of a state taking the preventative approach to data breaches, prescribing in great detail "technical, physical, and administrative security protocols aimed at protecting personal information." Also, as of March 2018, all states, the District of Columbia, and a number of US territories had adopted laws requiring that individuals be notified of security breaches involving personal data.[51]

[49] Jolly, I. (1 July 2017). Data Protection in the United States: Overview. Thomas Reuters Practical Law. Retrieved from https://content.next.westlaw.com/Document/I02064fbd1cb611e38578f7ccc38dcbee/View/FullText.html

[50] ibid.

[51] ibid.

Most recently, California has again exhibited leadership in data governance with the California Consumer Privacy Act of 2018. This Bill is an effort to protect data security and regulate how companies handle regulated and private data. It also takes explicit steps to give consumers more agency. When it comes into force at the beginning of 2020, it will give individuals the right to request a record of the types of data an organisation holds about them, what is being done with their data, and how it is being shared with, and by, third parties. It will also require business websites to provide consumers with a clear way to object to the sale of their data and will give individuals the right to erasure of their data.[52]

Recent data breaches and mounting public concern have also started to encourage renewed action at the federal level. In April of 2018, Senators Ed Markey and Richard Blumenthal introduced the Customer Online Notification for Stopping Edge-provider Network Transgressions (CONSENT) Act, which would require edge-providers to notify users if their data will be collected when they subscribe, establish an account, purchase, or begin receiving a service. It would also "…require edge-providers to obtain express consent from users before using, disclosing, or permitting access to any of the personal information collected. This is intended to address third-party or secondary uses of users' personal data. It means that users would have to explicitly opt-in to having their data used."[53] The Bill is still in the early stages of the legislative process and yet to influence the US data governance model in a direct way.

Technology architecture: As with the policy, regulatory and institutional environment, the current technology architecture of US data governance is complex and distributed across state and local government as well as the private sector. There is no single, unifying national architecture for data management. Again, paralleling the policy environment, particular types of regulated data – such as in the health sector – have developed technology standards and implementation specifications to allow for interoperability of data across systems. The health sector aims to integrate with private health data, such as that produced by wearable technology, by 2020.[54] In the absence of nationwide data portability standards and infrastructure for most types of data, data aggregators often fill this gap in the US marketplace. This enables data to be transferred between entities via data aggregators whose business it is to collect and transfer (or sell) data.

Impact on user participation, agency and choice: Despite state and federal legal and regulatory constructs aimed at protecting individuals' data, the US model has lacked the citizen-centric approach to data rights that the EU has had since 1995 under the Data Protection Directive, reinforced by the new GDPR. With a few exceptions, such as the right to request a free copy of credit reports annually, US consumers have had few ways to manage their data – government data is distributed across multiple state and federal databases, regulated data is difficult to access, and private data (like that created on social media platforms) has been controlled and used by private companies.

---

[52] Pfeifle, S. (2018, 18 June 2018). California passess landmark privacy legislation [blog]. Retrieved from https://iapp.org/news/a/california-passes-landmark-privacy-legislation/

[53] Watson, Katie. (2018, 30 April 2018). CONSENT: Privavcy is Key to Reinforcing Trust [blog]. Retrived from www.internetsociety.org/blog/2018/04/consent-privacy-is-key-to-reinforcing-trust/

[54] HealthIT.gov. Shared Nationwide Interoperability Roadmap: The Journey to Better Health and Care. Retrieved from www.healthit.gov/infographic/shared-nationwide-interoperability-roadmap-journey-better-health-and-care

Furthermore, a recent Pew Research Center study found that 64% of all Americans have personally experienced a major data breach, and roughly half believe their personal information is less secure now than five years ago.[55]

This combination of forces highlights three particular challenges:

- **Consumer awareness challenges:** The complexity of the rules that govern health information in the US illustrate the consumer awareness challenge. HIPAA, as the primary law governing health information in the US, only applies to "covered entities" holding "protected health information". There is general acknowledgement by federal regulators that individuals do not understand which entities fall into this category, nor do they understand which health information is "protected" and which is not. Specific laws that apply to other aspects of the health system, such as the Family Educational Rights and Privacy Act (FERPA) which governs management of student health records, further compound the challenge of consumer awareness.[56]

- **Data oversight challenges:** To add to the confusing rules of government data management, third-party integration with major platforms like Facebook, Google, Amazon, or Salesforce through APIs contributes to further loss of control of personal data by individuals. Much of the internet economy runs on services offered through such integration, enabling many of the services that consumers value. However, this also creates systems where individuals' data are bought and sold by third parties for uses that go well beyond consumer awareness and control.

- **Consent process challenges:** Currently, most consumer products in the US market are designed in such a way that most individuals simply click through terms of service and long legal policies without comprehension of what it is they are agreeing to. Frequently this consumer behaviour is seen simply as a trade-off: a willingness to give up some privacy in order to take advantage of valuable products. US consumers who want to exert more privacy over their data often have to go through an extensive and complicated process to opt out of personal data usage practices. Frequently, consumers are unaware that such options exist, and this limits their ability to consent to how their personal data is handled.[57] Of course, in the fragmented, sector-specific US approach, some sectors do offer a pathway for broader opt-in policies. The Consumer Financial Protection Bureau, for instance, developed a set of Consumer Protection Principles for companies that are authorised by consumers to use their personal financial data to analyse aggregate consumer needs and develop new products.[58]

---

[55] Olmstead, K. and Smith, A. (2017, 26 January 2017). Americans and Cybersecurity. Pew Research Center | Internet and Technology. Retrieved from www.pewinternet.org/2017/01/26/americans-and-cybersecurity/

[56] O'Connor, N. (30 January, 2018). Reforming the U.S. Approach to Data Protection and Privacy. Council on Foreign Relations | Digital and Cyberspace Policy Program. Retrieved from www.cfr.org/report/reforming-us-approach-data-protection

[57] Watson, Katie. (2018). CONSENT: Privavcy is Key to Reinforcing Trust [blog].

[58] US Consumer Financial Protection Bureau. Consumer-authorized financial data sharing and aggregation. (18 October 2017). Retrieved from https://files.consumerfinance.gov/f/documents/cfpb_consumer-protection-principles_data-aggregation_stakeholder-insights.pdf

Against this backdrop, the current US model has mixed results for the guiding principles of a 'data democracy':

1. Participation: With nearly 80% of the population using the internet, and one of the largest smartphone markets in the world, the barriers to participation in the data economy are low in the US. Likewise, the ability of individuals and small businesses to engage online is generally unrestricted. However, in terms of data management, the lack of consumer awareness, data oversight, and consent challenges are limiting factors in individuals' ability to actively manage and benefit from their data 'endowment'.

2. Agency: In 2017 alone, there were three major incidents: a security breach at Equifax exposed detailed personal information – including social security and passport numbers – of nearly 150 million people; Deep Root Analytics accidentally leaked the personal information of nearly 200 million US votes; and it was revealed that Uber had attempted to conceal a data breach of 57 million accounts. These security breaches at private companies, combined with the recent revelations surrounding the Facebook–Cambridge Analytica scandal, have significantly eroded US consumer trust in public and private institutions to manage their data.[59] With respect to the other key element of agency – user-centric policies and technologies – the U. model is mixed. There are examples of regulated data, particularly in the financial sector, where individuals are empowered to manage their data. However, that is not consistent across regulated data, nor true for fragmented government data that resides with many different governmental agencies at both state and federal levels, or for private data that resides with private companies.

3. Choice: The preference for the multi-stakeholder model of internet governance in the US, and the associated involvement of industry in the operations of the internet, have fostered a strong culture of innovation in the ICT sector. Likewise, the free market approach to the digital economy has created competition in the US market. However, the strength of a few data dominators – Google, Facebook, and others – limits the leverage that consumers have in choice surrounding how to manage their data. Only recently have private companies like Digi.me started to introduce innovative products for personal data management and increase consumer choice.

[59] Olmstead and Smith. (2017) Americans and Cybersecurity.

Section IV: Emerging insights

Policymakers are increasingly understanding the governance of personal data as a strategic imperative that requires urgent attention. This is being driven by a variety of forces, including more frequent hacks of technology systems and databases. This can result in stolen personal data being used for identity theft and other illicit purposes. There is a new awareness that data can be used for discriminatory and other negative purposes that put individuals and society at risk. There is also a recognition of the information and power asymmetries that exist between individuals and the companies that manage their data. Many policymakers now believe that they must shift from a reactive to a proactive approach when it comes to personal data, especially if they view such data as a potential mechanism for accelerating human and societal development.

This paper has explored the approaches being taken by four countries – China, Estonia, India and the US – by looking at their potential impact on individual participation, agency, and choice with regard to their data. In other words, we evaluate how policymakers can make personal data a force for good, even a pathway to prosperity, in the lives of ordinary people, especially people who will likely be "data rich" before they are financially so.

Based on the lessons from China, Estonia, India, and the US, policymakers looking to take action in this space should keep several considerations in mind:

1. Left untouched, the power dynamics between individuals, corporations and government as it relates to personal data tend to skew away from individuals. This is because, on their own, individuals have very little bargaining power to determine how their personal data is treated. The information and power asymmetries that exist between them and those who manage their data are simply too large for individuals to exert meaningful influence. This could be considered a market failure that cannot be resolved without deliberate government action.

2. While most individuals in the world may not currently have the ability to actively manage their personal data for their own benefit, they are nonetheless generating detailed personal data histories that, with the right opportunity, could be leveraged to improve their lives. While the extent to which a 'data endowment' can lead to real livelihood improvement is still unknown, specific cases demonstrate its potential value. However, unlocking this benefit for individuals may not be possible through policy and regulation alone. While rights over their personal data may give individuals the legal basis for exerting greater control over their data, it does not necessarily provide them with the means of doing so. Easily accessible and usable data management tools may also be necessary.

3. Governments have an important role to play beyond just establishing data rights. For example, an essential precondition to the governance of personal data is the ability for individuals to first participate in the digital realm. This requires a variety of public goods digital infrastructures such as internet connectivity. The example of India demonstrates

that such infrastructure, when designed as platforms, can serve to benefit individuals and the marketplace alike by creating opportunities for individual participation and providing a foundation for private sector innovation and competition. This broader view on the government's role with regard to personal data likely requires deliberate national visions that extend beyond individual rights and the prevention of harm to proactive individual empowerment through data. National visions are also needed to avoid the fragmentation that can naturally occur between parts of government or subnational jurisdictions.

4. Governments may find it helpful to think about policy and technology development as two parallel tasks that can be mutually reinforcing and necessary to achieve desired outcomes. Establishing technology standards and frameworks in parallel with policy rules and rights may be more effective than just one of these approaches on its own.


BLAVATNIK
SCHOOL OF
GOVERNMENT
UNIVERSITY OF
OXFORD


Pathways
for Prosperity
Commission
Technology &
Inclusive Development