Yale Law School
Paul Tsai China Center
耶鲁大学法学院蔡中曾中国中心

AI NOW

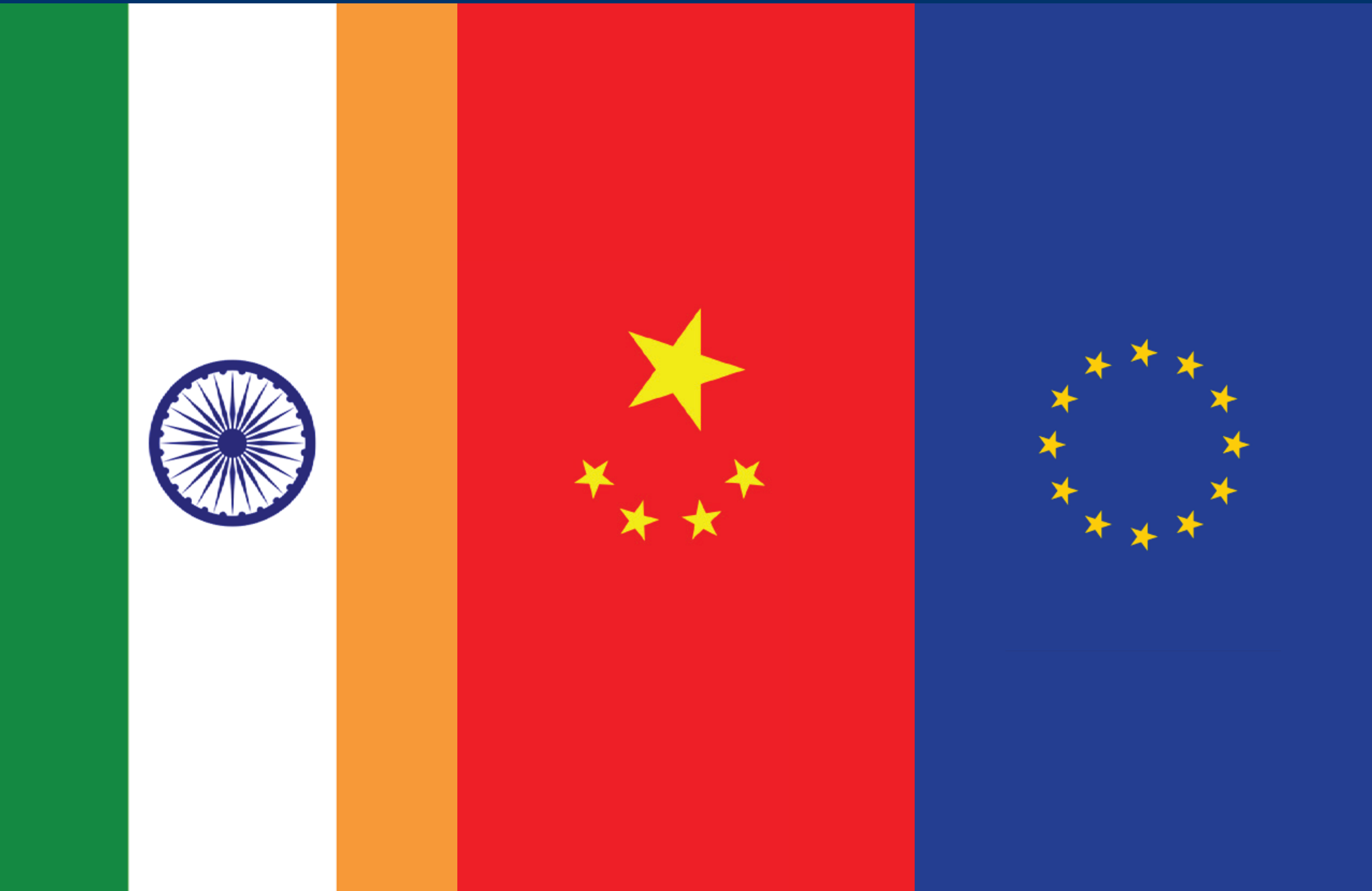
NEW AMERICA

Shifting Narratives and Emergent Trends in Data-Governance Policy

Developments in China, India, and the EU

Amba Kak and Samm Sacks

POLICY MEMO AUGUST 2021



About the Authors

AMBA KAK

Amba is currently Director of Global Policy & Programs at AI Now Institute at New York University where she leads the institute's global policy engagement and partnerships, and is also a fellow at the Engelberg Center at the NYU School of Law. Amba has over a decade of experience in the field of technology-related policy across multiple jurisdictions and has provided expertise to regulatory bodies, civil society organizations, and philanthropies. Before AI Now, Amba was a Global Policy Advisor at Mozilla, where she led the organization's work in India and contributed to global policy fora on issues such as data protection, online content regulation, and network neutrality. Amba has previously been the recipient of the Google Policy Fellowship and the Mozilla Technology Policy Fellowship which supported her research and advocacy in this field.

Trained as a lawyer, Amba received her BA LLB (Hons) from the National University of Juridical Sciences in India. She has a Masters in Law (BCL) and an MSc in the Social Science of the Internet at the University of Oxford, which she attended as a Rhodes Scholar.

SAMM SACKS

Samm Sacks is Cyber Policy Fellow at New America and a Senior Fellow at Yale Law School's Paul Tsai China Center. Her research examines China's information and communications technology (ICT) policies, with a focus on the U.S.-China technology relationship and the geopolitics of data privacy and cross-border data flows. She has worked on Chinese tech and cyber policy for over a decade, in both the national security community and the private sector. Previously, Samm launched the industrial cybersecurity business for Siemens in Japan, Korea, and China. She also led China technology sector analysis at the political risk consultancy Eurasia Group.

Samm holds an M.A. from Yale University in international relations and a B.A. from Brown University in Chinese literature. Her writing on China's data policies has appeared in *Foreign Affairs, MIT Tech Review, The Atlantic,* and *Slate,* among other outlets. She reads and speaks Mandarin and was a Fulbright Scholar in Beijing.

Contents

Introduction

This policy paper provides a synthesis of some key trends in data-governance policy in India, China, and the European Union (EU), alongside more detailed explanations of how the trends manifest themselves across these regions. The term *data governance* helps capture a rapidly expanding body of law and other (softer) policy frameworks that regulate access to and transfer of data between different entities in the digital economy. We focus, in particular, on an interconnected subset of policies relating to data privacy, transnational *data flows,* access to datasets held by governments or companies, and rules that promote the competitiveness of "national champion" tech platforms. In addition to analyzing the regulatory frameworks, we also attempt to capture the dominant tropes or rhetorical claims made by the stakeholders who endorse these policies, as well as policymakers and analysts looking into these regions from outside. In fact, while regulatory texts often read like mundane or bureaucratic stipulations or classifications, tracking these policy narratives in parallel reveals the broader political and economic vision they are part of. It also complicates a policy narrative that there is a simple data-governance "model" represented by each of these regions: many elements remain unresolved and fluid amid negotiation among various actors.

We chose to focus on these regions because they are at the forefront of articulating national-level data-governance approaches in ways that are already creating ripple effects and reference points around the world. One of our own revelations through this project has been the strikingly similar themes that cut across India, China, and the EU. We want to find ways to highlight that similarity without losing national or regional nuances. Our goal is certainly not to try to flatten contrasts or smooth over contradictions across these varied political economies; on the contrary, where relevant, we explain how a seemingly similar high-level trend manifests very differently when understood more granularly, and within its richer—and messier—national or regional context.

1. Data Governance As Economic Policy

Trend 1: National or regional economic development is emerging as an explicit justification for data-governance policy. This often translates into policy that incentivizes data sharing within national or regional borders to aid domestic businesses, and stricter controls on data flows outside of the region.

Trend 2: Recent data- and AI-governance frameworks include clear elements of industrial policy, such as the facilitation of domestic data markets, promotion of domestic cloud infrastructure, and public investment in skill building and research infrastructure.

Summary While the economic imperatives for data governance are not new by any stretch of the imagination, the national development rationale for policies is certainly becoming more explicit, alongside traditional justifications like privacy or national security.

The Indian government has stuck its neck out in national and international fora, arguing that low- and middle-income countries like India have the prerogative to shape data policy to meet their developmental needs, and to leverage the untapped data potential of the country's large population to create a more competitive national data economy. While much of this policy vision is still nascent, with proposals for facilitating access to data for domestic actors still in draft stages, data-localization restrictions have already been implemented in the financial sector. The developments around the draft Non-Personal Data Governance Framework, in particular, triggered lively public debate, with civil society pushing back against the conflation of domestic business interests in data access with the more ambiguous notion of "national interest" or "the public good."

In the EU, the data and AI strategies of 2020, in particular, foreground the economic rationale for data governance, including through more permissive data-sharing arrangements in pursuit of the "single market for data" in the EU region. While the single-market rationale is by no means a new motivation for European data-governance policy and has consistently factored in EU data-protection policy for decades, the economic justifications for regulation are now more explicit, especially in comparison to the discourse around the General Data Protection Regulation (GDPR) that emphasized the imperatives created by European human rights obligations. In addition to a more explicit economic rationale, certain traditional data-protection rights, such as the right to data portability, are now resurfacing in economic regulations like the 2020 Digital Markets Act proposal.

In China, there is increasing official acknowledgment of the need to balance the national-security rationale to strictly control data flows with economic imperatives to create more fluid data sharing and data markets within the country. A growing chorus of voices is also arguing for loosening restrictions on international data flows to benefit Chinese companies that have a global presence. Economic policy imperatives are not recent, but official recognition of a need to strike a balance between security and economic development in data policies marks an important shift.

CHINA

Beijing views data as a strategic resource not just for national security, but also for economic development.[1] According to a senior cybersecurity figure in China who also contributed to drafting China's Data Security Law (DSL),[2] data must be secured in order to make use of it for development. "The two go hand in hand," he said.[3] In fact, key parts of the law could be interpreted as more of an industrial strategy for data than a law meant to strengthen cybersecurity. Chapter 2 of the law emphasizes that national security objectives do not mean sacrificing the opportunity to use data to fuel innovation and the digital economy. In particular, the law introduces two concepts with more detail about how, concretely, the government aims to use data toward development:

- "The State firmly places equal emphasis on safeguarding data security and promoting data development and use." (数据开发利用) (Article 12)
- "The State establishes and completes data exchange management systems, standardizes data exchange activities, and cultivates a data trade market." (数据交易管理制度, 规范数据交易 行为,培育数据交易市场) (Article 17)

Article 12 makes explicit that security and development must be balanced in China's data-governance system.[4] These two competing tensions have shaped China's cyber bureaucracy for years. This long-standing internal source of friction has contributed, at least in part, to the Chinese government not necessarily enforcing (or unevenly enforcing) the strictest or most conservative security-oriented readings of Chinese cybersecurity laws and regulations. Devoting an entire early chapter of the draft law to the topic indicates recognition by Chinese authorities that state power hinges not only on security of data, but also on its commercial use, and that China must therefore find an effective way to leverage both at once.

Over the past year, the government has issued policy directives that elevate the concept of data as an economic asset. An April 2020 directive issued by the State Council and Central Committee of the Chinese Communist Party designated data as the fifth factor of production—after land, labor, capital, and technology.[5] Then, at the National People's Congress in March 2021, the outline of the 14th Five-Year plan called for "improving the market of data factors" (健全数据要素市场), sparking a wave of commentaries exploring unlocking the value of data to fuel the shared and digital economy.[6] One commentator in *Sina Finance,* for example, writes that companies should open up their search, e-commerce, and social data, and that "data ownership does not mean an exclusive right, but rather the right to access and use."[7]

[1] The idea of data as a strategic resource is not a recent development in China. Echoes of these principles are evident in a series of Big Data White Papers (2014, 2016, 2018) published by an influential think tank under the Ministry of Industry and Information Technology, as well as in the Big Data Strategy (2017). The 13th Five Year Plan (2016–2020) calls for "fully implementing the promotion of the big data development initiatives and accelerating the sharing of data resources and development of applications, to assist in industrial transformation and upgrading . . . "

[2] Emma Rafaelof et al., "Translation: China's 'Data Security Law (Draft)'," New America, July 2, 2020, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinas-data-security-law-draft/.

[3] Interview conducted by phone, October 23, 2020.

[4] Article 12: "The State firmly places equal emphasis on safeguarding data security and promoting data development and use."

[5] Ouyang Shijia, "New guideline to better allocate production factors," April 10 2020, China Daily, https://www.chinadaily.com.cn/a/202004/10/WS5e903fd7a3105d50a3d15620.html

[6] Sina Online, "What Is the Meaning of the '14th Five-Year Plan' Outline (Draft) to Improve the Market of Data Elements?," March 5, 2021, https://finance.sina.com.cn/china/2021-03-05/doc-ikftssaq1688850.shtml.

[7] Ibid.

Hong Yanqing, one of China's most influential scholars of data law, explores the relationship between data sovereignty and the emphasis on data as an economic asset. He writes that data sovereignty should be viewed from a development perspective, arguing that "to safeguard data sovereignty, China should also consider how to enable Chinese enterprises to control and use more data globally. After all, the United States can extend its 'arm' because its enterprises are all over the world."[8] Hong calls for adhering to data sovereignty while also creating data flows to Chinese internet companies operating overseas: "It is time to broaden the scope of data sovereignty. While adhering to the traditional concept of sovereignty . . . China needs to establish a good data flow order for Chinese information and communication technology (ICT) companies with global businesses. It is a critical component in the design of China's model of cross-border data access."[9] In doing so, Hong observes a tension between data sovereignty and data as a tool of economic development, one that, he argues, could create a disadvantage for Chinese companies in the future, especially if the US and EU are able to align on digital policies. He points to the disadvantages of creating split products for different markets (for example, Bytedance segmenting its global and Chinese versions of the apps TikTok and Douyin). He writes that this approach "prevents Chinese ICT companies from upgrading services by using a global data pool and limiting the gains from the economics of scale. Once the United States and the European Union reach an agreement, at least their enterprises can avoid data localization and segregating storage, which puts Chinese ICT enterprises at a disadvantage."

China's private sector, too, has picked up the concept of nationalist elements of data as an economic resource. An article by the Tencent Research Institute discusses what the idea of data as a factor of production may mean, emphasizing nationalist elements of facilitating more data flows to China's large tech platforms. Citing an International Data Corporation (IDC) estimate, the article states that "by 2025, the proportion of the world's data held by [China] will increase from 23.4% in 2018 to 27.8%, making China the first in the world. The open use of data resources will determine whether our country can seize the initiative in a new round of international competition and guarantee national data security through the development and growth of the digital industry."[10]

INDIA

In the past few years alone, India has developed an international reputation for pursuing "data nationalism" through a range of data-governance policies. While the rhetoric used by political leadership and Indian business tycoons—ending "data colonialism," establishing "data sovereignty"[11]—has been politically charged, it is economic imperatives that

[8] Hong Yanqing, "'Game of Laws': Cross-Border Data Access for Law Enforcement Purposes – The Models in the USA, Europe, and China," forthcoming trans. Yale Law School Paul Tsai China Center, originally published in *Global Law Review* in Chinese. This article is the result of a special 2018 project by the Ministry of Justice, "Big Data and Cybersecurity Legislation" (18SFB1005), in which the author participated.

[9] Hong, "Game of Laws."

[10] Chen Weixuan et al., "Data Production Factors in the Framework of Macroeconomic Growth: History, Theory and Prospects," Tencent Research Institute, June 12, 2020, https://tisi.org/14625.

[11] See Mahesh Langa, "Mukesh Ambani Urges Modi to Take Steps against Data Colonisation by Global Corporations, *Hindu*, January 18, 2019, https://www.thehindu.com/news/national/mukesh-ambani-urges-modi-to-take-steps-against-data-colonisation/article26025076.ece; Nandan Nilekani, "Why India Needs to Be a Data Democracy," *LiveMint*, July 27, 2017, https://www.livemint.com/Opinion/gm1MNTytiT3zRqxt1dXbhK/Why-India-needs-to-be-a-data-democracy.html; Aroon Deep, "#NAMAprivacy: BJP's Vinit Goenka on Data Localisation, Sovereignty, and 800 Years of Slavery," *MediaNama*, September 11, 2018, https://www.medianama.com/2018/09/223-license-patent-colonisation-bjp-vinit-goenka-namaprivacy/; and Pankaj Doval, "Apple Engaged in Data Colonization," *Times of India*, August 8,https://timesofindia.indiatimes.com/business/india-business/apple-engaged-in-data-colonisation/articleshow/59961875.cms.

have been key to this emergent policy vision aimed at the competitiveness of domestic enterprise. These policies, if implemented, would impact data-driven services broadly, and yet much of the emphasis has been on the consumer-technology sector currently dominated by US "big tech" companies like Amazon, Microsoft, Google, and Facebook.

The first clear and official articulation of this economically driven policy vision was the Draft National e-Commerce Policy of February 2019, published by the Ministry of Commerce.[12] Subtitled "India's Data for India's Development," the draft puts forth a laundry list of government proposals to favor the growth of domestic digital players. Items include, for example, access to data for smaller Indian firms as a potential policy lever to boost the domestic digital economy by countering the high barriers to entry created by larger market entities. The draft states that these smaller firms, granted "infant-industry" status, should be given preferential access to data about Indians, which is currently disproportionately controlled by large foreign companies. The data-localization mandate in the Personal Data Protection Bill of 2019 was also viewed as a penalty against US companies, which lobbied publicly against it.[13]

During the same period, the 2019 Economic Survey of India, the government's flagship economic planning document, included its first section on data markets.[14] Comparing data to a natural resource, the section emphasizes the need for India to harness the economic value of data, stating that once anonymized, data should be understood as a public good to be used for public benefit. The 2020 National Strategy for Artificial Intelligence echoes the need to overcome the hurdle of limited data access that acts as a barrier to AI innovation in India.[15]

But it is the developments around the draft Non-Personal Data Governance Framework that demonstrate how this economic justification for data access can exist in tension with the notion that data should be used in the "national interest" or for "the public good." For example, the first version of the draft framework, published in July 2020 by a government-appointed expert committee,[16] recommended that large global companies be required to share data with smaller Indian businesses to break down the barriers to market entry for smaller firms. This received wide-ranging critique from civil-society advocates and business actors alike. While businesses (both domestic and foreign)

[12] Government of India, Ministry of Commerce and Industry, Department for Promotion of Industry and Internal Trade, "Draft National e-Commerce Policy: India's Data for India's Development," February 23, 2019, https://dipp.gov.in/whats-new/draft-national-e-commerce-policy-stakeholder-comments.

[13] See, e.g., Aditya Kalra and Aditi Shah, "Exclusive - U.S. Tech Giants Plan to Fight India's Data Localisation Plans," Reuters, August 18, 2018, https://www.reuters.com/article/uk-india-data-localisation-exclusive-idUKKBN1L30CN; "Exclusive: Data Localisation Will Affect Growth of Indian IT Companies, Says Facebook," *India Today,* September 13, 2019, https://www.indiatoday.in/technology/story/data-localisation-will-affect-growth-of-indian-it-companies-says-facebook-1598923-2019-09-13; Nikhil Pahwa, "US Trade Secretary Wilbur Ross Highlights Data Localisation, High Tariffs on Electronics, Telecom Products in India as Trade Issues," *MediaNama,* May 9, 2019, https://www.medianama.com/2019/05/223-us-trade-secretary-wilbur-ross-highlights-data-localisation-high-tariffs-on-electronics-telecom-products-in-india-as-trade-issues/; and Nimish Sawant, "Sundar Pichai's Letter to RS Prasad Hints Why Data Localisation Isn't Feasible," *Firstpost,* September 10, 2018, https://www.firstpost.com/tech/news-analysis/sundar-pichais-letter-to-rs-prasad-hints-why-data-localisation-isnt-feasible-5151061.html.

[14] Gulveen Aulakh, "Econ Survey: Anonymised Data Should Be Treated as 'Public good,'" *Economic Times,* July 5, 2019, https://telecom.economictimes.indiatimes.com/news/econ-survey-anonymised-data-should-be-treated-as-pubic-good/70082749.

[15] "National Strategy on Artificial Intelligence," Government of India, NITI Aayog, accessed April 7, 2021, https://niti.gov.in/national-strategy-artificial-intelligence.

[16] Ministry of Electronics and Information Technology (MeitY), Government of India, "Report by the Committee of Experts on Non-Personal Data Governance Framework," 2020, https://static.mygov.in/rest/s3fs-public/mygov_159453381955063671.pdf.

argued that coercive data sharing would impinge on intellectual property rights and commercial interests in datasets,[17] civil-society advocates pushed back, citing data privacy and security concerns and the primacy to business interests over the privacy rights of Indian citizens.[18]

Likely in response to this backlash, the revised and final version of the framework released in December 2020[19] is significantly narrower in scope, removing data sharing between businesses and instead focusing exclusively on a public-good rationale to increase the use of data for development purposes. The proposal is that "high-value datasets" that can be used for developmental purposes can be subject to mandatory data sharing based on a public-good rationale. Examples of high-value datasets include data collected by ride-hailing platforms about city traffic, or by public utilities about energy data, or by telecom companies about users—and the framework contends that such data should be accessible and managed by "data trusts" in the interest of "communities" that are directly impacted.[20] The fact that the purely commercial rationale for mandatory data access is more subdued in the revised version of the framework indicates that the government's conflation of domestic economic interests in data with that of public interest still comes up against a range of barriers in India.

EU

In contrast to the discourse around the GDPR, recent policy developments in the EU reveal a more prominent emphasis on the economic drivers of data governance. The EU has been widely credited as a global leader in setting data-governance norms; the Data Protection Directive and, later, the GDPR have played an outsize role in shaping data-governance norms around the world. These regulatory efforts have foregrounded a political commitment to fundamental rights, and to Europe being accorded a certain kind of moral leadership as the frontrunner in legislating data-protection law.[21] Despite a consistent chorus of voices alleging that European data-privacy laws are a form of economic protectionism and a barrier to the free flow of trade,[22] the EU has successfully stood its ground, arguing that any restrictions on data flows outside the EU are based first and foremost on protecting Europeans' fundamental right to privacy.

[17] PTI, "Tech Trade Group Urges India Not to Accept Draft Report on Non-Personal Data Governance," *Business Today*, September 9, 2020, https://www.businesstoday.in/technology/news/tech-trade-group-urges-india-not-to-accept-draft-report-on-non-personal-data-governance/story/415580.html.

[18] Apar Gupta, Internet Freedom Foundation, submission on the draft Non-Personal Data Governance Framework, https://drive.google.com/file/d/1Nb3UeyDbpUSvW3DmDmdra8Eh_vsMR4SM/view. "The present expropriation of personal or non-personal data from silicon valley [sic] firms will not result in any clear user benefits but will lead to greater data collection and generation of sets."

[19] MeitY, "Report by the Committee of Experts on Non-Personal Data Governance Framework," December 2020, https://static.mygov.in/rest/s3fs-public/mygov_160922880751553221.pdf.

[20] See page 29 of the December 2020 MeitY report.

[21] See generally Lee Andrew Bygrave, *Data Privacy Law: An International Perspective* (Oxford: Oxford University Press, 2014); and Abraham L. Newman, *Protectors of Privacy: Regulating Personal Data in the Global Economy* (Ithaca: Cornell University Press, 2008), 121.

[22] See Susan Ariel Aaronson, "Digital Protectionism? Or Label the U.S. Government Uses to Criticize Policy It Doesn't Like?," Council on Foreign Relations, March 3, 2016, https://www.cfr.org/blog/digital-protectionism-or-label-us-government-uses-criticize-policy-it-doesnt, accessed March 5, 2020; see also Peter P. Swire and Robert E. Litan, "None of Your Business: World Data Flows, Electronic Commerce, and the European Privacy Directive," *Harvard Journal of Law & Technology* 12, no. 3 (Summer 1999), http://jolt.law.harvard.edu/articles/pdf/v12/12HarvJLTech683.pdf; and Svetlana Yakovleva and Kristina Irion, "Pitching Trade against Privacy: Reconciling EU Governance of Personal Data Flows with External Trade," *International Privacy Law* 10, no. 3 (August 2020), https://academic.oup.com/idpl/article/10/3/201/5813832/.

The current wave of policymaking in the EU is more explicitly focused on enhancing the European Union region's competitiveness in the data economy. The 2018 Framework for the Free Flow of Non-Personal Data is aimed at preventing restrictions on data flows between entities in different EU member states on the grounds that data localization "constrains the data economy's development."[23] More recently, the 2020 European Strategy for Data also foregrounds an enabling industrial strategy aimed at making "the EU a global leader in the data economy" and creating "a genuine single market for data" alongside acknowledging the need for restrictive guardrails around AI development in alignment with fundamental rights.[24] As the 2020 European Commission White Paper on Artificial Intelligence notes, the objective is to increase Europe's "technological sovereignty" in the global data economy, especially given that it lags behind US and Chinese counterparts in the consumer technology and online platforms market.[25]

The goal of a "genuine single market for data" cited in the 2020 Data and AI strategy is also defended primarily in terms of giving EU businesses the opportunity to build "on the scale of the single market."[26] The single-market rationale is by no means a new motivation for European data-governance policy and has in fact consistently factored in EU data-governance policy at least since the 1980s.[27] However, in the context of recent policy developments, there appears to be a more explicit foregrounding of the economic rationale motivating these data-governance policies, and especially so in comparison with the discourse around the GDPR.

In addition to a more explicit economic rationale, certain traditional data-protection rights, such as the right to data portability, are now resurfacing in economic regulation. The December 2020 Digital Markets Act proposal, aimed at tackling market dominance in the platform space, creates the category of "gatekeeper firms" that are subject to a range of requirements, including ensuring data portability for both end users and business users using these platforms.[28]

[23] "Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a Framework for the Free Flow of Non-Personal Data in the European Union," *Official Journal of the European Union*, November 28, 2018, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807&rid=2.

[24] European Commission, "Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A European Strategy for Data," February 19, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN.

[25] European Commission, "White Paper on Artificial Intelligence: A European Approach to Excellence and Trust," Brussels, February 19, 2020, https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

[26] "White Paper on Artificial Intelligence," 25.

[27] See European Community, "The Politics of the European Community," press release, June 16, 1988, https://ec.europa.eu/commission/presscorner/detail/en/DOC_88_5. "[A] harmonious development of economic activities within the common market calls for the creation of a genuine common market in data-processing in which the free movement of goods and freedom to provide services are assured and competition is not distorted."

[28] European Commission, "Proposal for a Regulation of the European Parliament and of the Council on Contestable and Fair Markets in the Digital Sector (Digital Markets Act)," December 15, 2020, https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A842%3AFIN:2020:842:FIN.

2. Data Classification

Trend 1: Regulation is increasingly targeted at specific subcategories of data in order to justify tailored interventions that can meet economically oriented policy goals, beyond those of data privacy or security. There is a particular uptick in regulatory activity around the (unstable) category of "non-personal data."

Summary Creating tailored regimes for specific kinds of data, like non-personal or anonymized data, is an increasingly prominent regulatory strategy to diversify the objectives of data-governance regulation beyond data protection and security. In China, for example, the draft Data Security Law calls for establishing a top-down data-classification rule to differentiate between categories of data and level of risk. Recent regulations for auto-sector data also suggest that authorities are undertaking a sweeping effort to categorize data using a more granular, subsectoral approach that details definitions for "important" and "personal" data in discrete contexts. The aim appears to be to create more flexibility in the use of certain kinds of data as a way to support the growth of the digital economy.

In India and the EU, too, the distinction between personal and non-personal data regimes has become more prominent in recent policy moves. This bifurcation appears motivated by the desire to create exceptional rules for non-personal data that allow for relatively unrestricted data sharing and reuse (compared to the data covered by personal data-protection laws). In the EU, the focus on non-personal datasets in recent policy frameworks is justified as part of a broader strategy to strengthen the EU's potential competitive advantage for "industrial AI" across sectors like agriculture, energy efficiency, and healthcare, in contrast to consumer technology, where foreign (particularly US) companies are dominant. In India, on the other hand, the bifurcation could also be explained as a strategy to mitigate anticipated pushback that any mandatory data-access provisions would bring in terms of clashing with the upcoming data-protection law, violating constitutional privacy protections, as well as India's obligations under trade treaties. However, it remains uncertain (and unlikely) that merely scoping this to non-personal data will relieve these concerns. The practical distinction between these categories of data remains on shaky ground. Moreover, if the definition of non-personal data is strictly construed, it will apply to a very narrow band of datasets with limited applicability in the consumer-technology sector. Given that the underlying justification for this regulatory activity has been the need for a nationally competitive data economy to rival Silicon Valley, the link between the means (more access to and sharing of non-personal data) and the ends (a more competitive consumer-technology sector) remains tenuous.

CHINA

The Chinese government has introduced a national or top-down data-classification system that will classify data by category and then grade it by level of risk and importance. Article 19, among the most significant developments in a draft Data Security Law released in July 2020, states:

> The State shall implement data protection for data at different grades and classifications, according to the degree of importance to economic and social

> development; and according to the impact on national security, the public interest, or the lawful rights and interests of citizens or organizations if it is falsified, destroyed, leaked or illegally acquired, or illegally used.
>
> Each region and department, according to relevant national provisions, shall determine a regional, departmental, and industrial important data protection catalog, and undertake special protections for that which is listed in the catalog.[29]

The term "data classification" refers to categorizing data into different types (分类), while "data grades" refers to different levels (分级) of risk, importance, and impact. Although previous laws, regulations, and standards going back to at least 2016 refer to "data classification," the DSL marks the first time that the words "data grades and classifications" appear together at the national level.

Data classification in China appears to be working toward three main objectives: (1) an attempt to centralize state control over data management, (2) an effort to carve out space for data-fueled economic development (as discussed in Section 1); and (3) the beginning of an inventory process to better understand and sort data so that businesses know what data they hold and can make the best use of their data resources.

The data-classification system remains in early stages, with questions about how new and existing parts of the system are meant to interact, creating areas of uncertainty and confusion as the various pieces take shape.

The first point of confusion is that "important data" is at once a fixed category and also a way to evaluate risk levels across all kinds of data. The idea of "important data" as a fixed category dates back to the 2016 Cybersecurity Law,[30] which identifies two kinds of data: "important data" and "personal data," laying the foundation for a bifurcated legal regime to treat these two kinds of data separately. The 2020 draft Data Security Law is one of a pair meant to formally divide requirements for these two kinds of data. It ostensibly aims to focus on important data, while its counterpart, a draft Personal Information Protection Law (2020), focuses on personal data. As scholar Lu Chuanying writes, "the relationship between the two laws must be handled well. The Personal Information Protection law treats data security issues more from the perspective of protecting citizen privacy, while the Data Security Law sets out from the perspective of national security and public security."[31]

The meaning of the term "important data" is itself the subject of intense debate among scholars, practitioners, and government officials. According to one practitioner, casting a wide net by defining the term broadly creates an impossible situation where the government lacks resources to implement requirements. Yet defining the term more narrowly means that a breach or disclosure could cause damage to state security that may have been overlooked; government officials may not want to expose themselves to either risk,

[29] "Translation: China's 'Data Security Law (Draft)'," New America, July 2, 2020, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinas-data-security-law-draft/.

[30] Rogier Creemers, Paul Triolo, and Graham Webster, "Translation: Cybersecurity Law of the People's Republic of China (Effective June 1, 2017)," New America, June 29, 2018, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-cybersecurity-law-peoples-republic-china/.

[31] Lu Chuanying, "A Chinese Scholar Outlines Stakes for New 'Personal Information' and 'Data Security' Laws (Translation)," trans. Graham Webster and Rogier Creemers, New America, May 28, 2020, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/chinese-scholar-outlines-stakes-new-personal-information-and-data-security-laws-translation.

so the term remains undefined beyond a list of 27 broad-sector categories in the "Data Outbound Transfer Security Assessment Guideline" associated with "data that when leaked can endanger national security, public interest, life or property interest, national key fundamental infrastructure, market order, national secrets, etc."[32] A separate forthcoming "important data" standard led by Zuo Xiaodong (an influential cybersecurity expert and vice president of the China Information Security Research Institute) will aim to define what constitutes important data. This promises to become a reference point for the kinds of data that will be subject to certain enhanced security measures like encryption, anonymization, backups, and audits.

Alongside the law calling for a top-down classification system, there are also regulations and voluntary standards broken down according to different sectors. The Cyberspace Administration of China (CAC) recently published draft provisions on data security in automobiles that signal that authorities are seeking to categorize data using a more granular, subsectoral approach, defining with greater specificity what the two categories of "important" and "personal data" mean in discrete contexts. If implemented, the provisions would represent the most detailed binding definition of what constitutes "important data" in any sector since the government to date.

Kendra Schaefer, who leads China tech policy at the Beijing consulting firm Trivium, speculates that in the future, data classification in China could consist of overlapping schemes made up of both laws and sector-level standards. Layers sorting data into different groups could coexist across various schemes, not just dividing out personal from national security data, but also breaking out data into subsets within industry type, and distinguishing data created by people from machines and metadata.[33]

Tensions also exist in the relationship between "personal" and "important" data – specifically the collision between a need to protect the privacy of citizen data while also treating data as a national security tool (see Section 3). At the same time, the rise of data as an economic asset (as discussed in Section 1) creates new fault lines when it comes to the demands of both national security and privacy concerns.

Even as the interactions among the different data frameworks remain unresolved, China is at the beginning stages of putting in place the building blocks of a system that aspires to enable companies to capitalize on the troves of data they hold, in furtherance of economic development goals (as discussed in Section 2.2). In theory, defining what kind of data is sensitive – needing further protection and regulation – could create more space to make better use of other data for industrial upgrading. According to the state news outlet *Xinhua,* the role of laws and regulations is to clarify data ownership among platforms, the state, and consumers in order to facilitate data markets.[34] And Chinese scholar Zhang Jihong writes:

> A data grades and classifications system is necessary because data protection needs tailoring. Resources must be devoted to the most vulnerable and important data. A flat regulatory regime that does not differentiate between

[32] "Data Outbound Transfer Security Assessment Guideline," National Information Security Standardization Technical Committee, August 30, 2017, https://www.tc260.org.cn/front/bzzqyjDetail.html?id=20170830211755&norm_id=20170221113131&recode_id=23883.

[33] Interview with Kendra Schaefer via Wechat, April 27, 2021.

[34] Qiao Ruiqing, "培育数据要素市场，助力数字经济高质量发展, *Xinhua*, April 14, 2020, http://www.xinhuanet.com/fortune/2020-04/14/c_1125854357.htm.

> data is bound to be ineffective and overly restrictive, hampering the healthy development of the digital economy.[35]

Chinese scholars and cyber officials also recognize and are grappling with the gaps in China's data-governance regime that may harm privacy because of more lenient requirements for use and sharing of anonymized data.[36] A new draft standard by the national cybersecurity standards body TC260 lays out a framework for evaluating both the effectiveness of personal information deidentification and risks of reidentification.[37] The document is likely to serve as a reference for assessing what counts as a suitable level of identifiability when it comes to data use in different contexts. More broadly, with this standard, China appears to be breaking new ground, since existing similar standards (like ISO/IEC 20889) provide less detail on identifiability.[38]

INDIA

Although the expert committee's draft Non-Personal Data framework has not yet been formally accepted, the Indian government has already made clear its intention to carve out a separate regulatory regime for non-personal data in a number of other recent policies.[39] This is pitched as an entirely distinct policy domain from that covered by the Personal Data Protection Bill (nearing its final vote in Parliament at time of writing[40]), which applies to personal data and defines anonymization of the data as the boundary condition for determining scope. In July 2020, a committee of experts appointed by India's Ministry of Electronics and Information Technology (MeitY)[41] proposed a draft governance framework for non-personal data. They released their final report in August 2020 after a hurried and largely confidential public consultation,[42] and set the foundation for a legislative proposal for the category of non-personal data, including the creation of a new non-personal data regulator. Barely three months later, in December, the government issued a slightly modified version of the proposal, claiming that it had responded to the feedback in the public consultation.[43] Creating a parallel policy process

[35] Zhang Jihong, "Data Security Law (Draft Seeking Comment): Data Classification System" 数据安全法（草案）》（二）：数据分级分类制度 （张继红）."

[36] "Data Anonymization or It Is Difficult to Protect Personal Privacy," Cyberspace Administration of China, July 24, 2019, http://www.cac.gov.cn/2019-07/24/c_1124790603.htm.

[37] "Notice on Soliciting Opinions on the Draft of the National Standard 'Information Security Technology Personal Information De-identification Effect Grading Evaluation Specification' (关于国家标准《信息安全技术 个人信息去标识化效果分级评估规范》征求意见稿征求意见的通知)," April 12, 2021, https://www.tc260.org.cn/front/bzzqyjDetail.html?id=20210412183118392628&norm_id=20201104200026&recode_id=41659.

[38] Email exchange with Alexa Lee, April 20, 2021.

[39] The Draft National eCommerce Policy released in 2019 indicated that non-personal data could be used for the benefit of Indian companies and the Indian government for governance purposes. The Personal Data Protection Bill also makes reference to non-personal data, providing for mandatory sharing of non-personal or anonymized data "to enable better targeting of delivery of services or formulation of evidence-based policies by the Central Government." The National Strategy on Artificial Intelligence, for instance, contemplates making some types of government data available for the "public good" and requiring corporations to share aggregated data as a means of overcoming the hurdle of limited data access within India's AI ecosystem. Elsewhere, the 2018–2019 Economic Survey of India likened data to a natural resource and stated that personal data, once anonymized, becomes a "public good" that should be utilized for public benefit.

[40] The bill is currently in the Lower House of Parliament and is expected to be voted on (and passed, given the government's clear majority) after review by the Parliamentary Standing Committee.

[41] MeitY, "Office Memorandum: Constitution of a Committee of Experts to Deliberate on Data Governance Framework," September 13, 2019, https://www.meity.gov.in/writereaddata/files/constitution_of_committee_of_experts_to_deliberate_on_data_governance_framework.pdf.

[42] The public consultation—likely prompted by a media leak—was heavily critiqued by civil-society advocates because no submissions were made public. Interview with Udbhav Tiwari, policy expert based in New Delhi, November 2020. Transcript on file with the authors.

[43] MeitY, "Report by the Committee of Experts on Non-Personal Data Governance Framework," December 16, 2020, https://static.mygov.in/rest/s3fs-public/mygov_160922880751553221.pdf.

for non-personal data before the comprehensive data-protection legislation is enacted and implemented has created a great deal of uncertainty and confusion around how these bodies of law (and potentially separate regulatory bodies) will interact.

This interaction is, of course, inevitable. The distinction between personal and non-personal data remains slippery, particularly in the context of most datasets being "mixed" and dynamic, or having elements of both at different periods of time. The Committee Report has clarified that mixed datasets that have inextricably linked personal and non-personal data will be governed by the Personal Data Protection Bill, which means that effectively there will be a very narrow band of datasets that falls within scope. While anonymization is referred to in both the Personal Data Protection Bill and the draft Non-Personal Data Framework as the process that can convert personal to non-personal data, decades of evidence point to the possibility of reidentification of anonymized data, which has been repeatedly demonstrated in studies across sectors. While there are certainly pure non-human datasets (e.g., industrial or supply-chain data) that might be considered strictly non-personal, this raises the question of whether this very limited subset of datasets would serve the grand objectives of the policy to decentralize market power or incentivize innovation.

It also begs the question of what else might have motivated this bifurcation of policy regimes. One possible explanation could be that non-personal data as a category might avert some of the anticipated legal and political pushback on the government's aggressive proposals for more control over data. This includes (1) restrictions imposed by the Personal Data Protection Bill or constitutional jurisprudence on the right to privacy, (2) concerns raised about compatibility with India's global free-trade commitments,[44] and (3) heated public discourse in India around the need for stronger protections against state control over citizen personal data in the absence of surveillance laws.[45] Promoting non-personal data as a "new" data category offers the potential for giving the government and other proponents additional leeway to mold a policy vision that posits data as a national asset in pursuit of the industrial policy goals described in Section 1.

Eventually, the Personal Data Protection Bill will certainly be implemented before the Non-Personal Data Framework, which means this apparent bifurcation might have subdued practical implications. Udbhav Tiwari, a policy expert based in New Delhi who has been actively involved in providing feedback on these developments, posits: "Over the next twelve to twenty-four months, I'd say many of the key debates we're having in the context of non-personal data will play out in the context of personal data protection first. Non-personal data is the flavor of the month."[46]

EU

In the EU, as in India, non-personal data has been the subject of increasingly lively policy development. These policy moves focus on creating an enabling framework to incentivize increased sharing of this category of data within the EU. The 2018 Framework

[44] Arindrajit Basu and Justin Sherman, "Key Global Takeaways from India's Revised Personal Data Protection Bill," *Lawfare,* January 23, 2020, https://www.lawfareblog.com/key-global-takeaways-indias-revised-personal-data-protection-bill.

[45] Vrinda Bhandari and Renuka Sane, "Protecting Citizens from the State post Puttaswamy: Analysing the Privacy Implications of the Justice Srikrishna Committee Report and the Data Protection Bill, 2018," *Socio-Legal Review* 14, no. 2 (n.d.), http://docs.manupatra.in/newsline/articles/Upload/7B08CF55-E27D-4A44-A292-3882F08E9053.pdf.

[46] Interview with Udbhav Tiwari, November 2020. Transcript on file with the authors.

for the Free Flow of Non-Personal data was in fact the culmination of a series of policy documents that surface the economic imperatives for incentivizing the sharing of non-personal data and creating an effective "single European data market." These economic incentives include creating more access to data between companies in different member states that could lead to more competitive data-driven products and services. This has since been reinforced in the 2020 European Strategy for Data and the 2020 Data Governance Act proposals put forth by the Commission.

The motivation for special emphasis on "non-personal data" and "industrial data" as an essential input for innovation appears pragmatic: it offers a pathway for innovation policy that runs parallel to (rather than counter to) the GDPR's requirements. Common illustrative examples of high-value non-personal datasets in these policy documents include data on precision farming, city planning, energy efficiency systems, or data on maintenance needs for industrial machines, indicating that the primary targets of this policy will be industrial rather than consumer-technology firms.[47] The European Commission defends this focus on industrial data by arguing that it can build a competitive advantage in these domains since it already has "a strong position in digitised industry and business-to-business applications, but a relatively weak position in consumer platforms."[48] Along with this commercial justification, the Commission emphasizes the societal benefits that will flow from strengthening European innovation based on transport, agricultural, and energy data in the form of "improved healthcare," "fewer breakdowns of household machinery," "safer and cleaner transport systems," "better public services," and so on.[49]

Despite these strategic justifications, a policy regime predicated on the slippery distinction between personal and non-personal data is on shaky footing across most of these strategic datasets, whether in the realm of healthcare, transport, or mobility data. This will inevitably lead to ambiguity around the scope of these laws, and could cultivate strategic behavior by companies seeking to exploit this uncertainty. The European Data Protection Supervisor (EDPS) has acknowledged this by stating that the non-personal data framework "carries significant risks of overlap or conflict with the GDPR, thus undermining legal certainty and causing difficulties of practical application."[50]

[47] See "Common European Data Spaces in Strategic Sectors and Domains of Public Interest" on page 26 of the 2020 "European Strategy for Data," https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN; and see page 3 of the European Commission's "White Paper on Artificial Intelligence," https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

[48] "White Paper on Artificial Intelligence," 1–2.

[49] "White Paper on Artificial Intelligence," 2.

[50] European Data Protection Supervisor, "Comments of the EDPS on a Proposal for a Regulation of the European Parliament and of the Council on a Framework for the Free-Flow of Non-Personal Data in the European Union," June 8, 2018, https://edps.europa.eu/sites/edp/files/publication/18-06-08_formal_comments_freeflow_non_personal_data_en.pdf.

3. Access to Data

Trend 1: Creating increased access to data is a central motif in recent data-governance proposals across jurisdictions, which view data as the foundational raw material for enabling domestic data businesses as well as data-driven governance by state actors. In the EU and China, there is a heightened focus on access to government datasets for private data businesses (alongside government access to private data in China), while in India, draft policies gesture toward access to data held by foreign private firms.

Summary Recent data-governance policies prioritize greater availability and access to data for national and regional businesses. This is based on an intuitive understanding that data is a foundational raw material for a competitive data economy. In India, policy documents sometimes refer to this as moving beyond data protection to "data sharing" and "data empowerment" as the driver of data-governance policy. In the EU, it is framed in terms of the broader policy goal of a "genuine single market for data"; the 2020 data and AI strategies, as well as the Framework for the Free Flow of Non-Personal Data, push for more frictionless data transfers within companies (and countries) in the region.

That said, there are crucial differences in the direction (state access to private data versus private access to state data) this conversation is taking in different regions. In China, for example, recent regulatory efforts appear to put more guardrails on the state's ability to access data held by companies. Yet many argue that creating a more institutionalized process by which the government should make data requests could eventually serve to promote *more* data sharing between companies and the government rather than less.

Alongside this, however, there are ongoing efforts to create increased access to government datasets that can be used by Chinese companies to develop apps. In the EU, too, the most recent 2020 Data Governance Act (DGA) proposal is entirely focused on creating greater access to government datasets, including those that contain personal data.[51] These proposals do not undercut the GDPR, but instead leave open the possibility of techno-legal arrangements that will allow this data sharing while still ensuring compliance with data protection norms. In India, the emphasis in recent policy moves has been on creating increased access to data held by foreign companies in favor of domestic enterprise as well as public agencies alongside a (still vague) notion of creating access to private datasets for "the public good." India still lacks a comprehensive data-protection law, and the latest version of the Personal Data Protection Bill has been criticized for being particularly weak on restricting state access to data. In this context, a prominent civil-society demand has been that these developments must follow rather than precede data-protection frameworks and restrictions on government surveillance.

CHINA

Recent developments in China's data-governance system aim to create a lawful process for government access to data held by companies and across government agencies. These efforts seek to break down data silos that stem from a combination of factors, including

[51] European Commission, "Proposal for a Regulation on European Data Governance (Data Governance Act)," November 25, 2020, https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-data-governance-data-governance-act.

reluctance to share data as a form of political or commercial power and bureaucratic inefficiency.

Data silos or data “islands” have long plagued China’s government interagency and created barriers within and among private companies. These companies are reluctant to share their data as valuable commercial intellectual property,[52] while government agencies often push back against one another’s access requests, guarding their data as a form of political power. Government data use is not monolithic, with different actors seeking data not just for security and surveillance, but also for digital-economy and other administrative needs. Within the government, failure to share data also sometimes reflects mundane factors like bureaucratic inertia apart from political drivers. This may be changing, though, as the government looks to set up processes to make data sharing more efficient and facilitate lawful access to both state and private-sector data to spur innovation, while also centralizing control over information flows tied to governance. Data-exchange markets (discussed in Section 4) appear to work in tandem with changes to the legal system to facilitate data access among different actors while also establishing guardrails.

The draft DSL and Personal Information Protection Law (PIPL) both contain provisions on government access to data held by companies. According to an interview with one leading legal scholar in China, the intent behind these requirements is to rein in the excessive data-extraction power of the state.[53]As Jamie Horsley observes, the draft PIPL “subjects state organs to its general limiting principles of legality, legitimacy, necessity, and minimum scope for data handling, and specifies they must handle personal information according to their legal authority and not exceed the scope and limits necessary to carry out their statutory duties (Article 34).” [54] She also notes that the PIPL contains requirements relating to automated decision-making (Article 25) and use of facial recognition and surveillance for public safety purposes (Article 27).

Under the DSL, the relevant article is 32, which states:

> Where public security departments and national departments need to consult data in order to lawfully safeguard national security or investigate a crime, they shall, according to relevant State regulations, undergo strict approval procedures and proceed according to the law; relevant organizations and individuals shall grant cooperation.

The lack of specifics in these key passages has led to debate among Chinese scholars and policy experts about whether the vague language is sufficient to protect privacy from government actors. Wang Xixin, a law professor at Peking University Law School, told the media outlet *Caixin* that “the section on data collection by the state should be expanded into a stand-alone chapter to include more-detailed requirements. Government agencies

[52] China now protects both trade-secret and confidential business information, which is being identified as a subset of traditional trade secrets as IP. See the draft “Notice of the Ministry of Justice on the ‘Guiding Opinions on Strengthening the Protection of Commercial Secrets and Confidential Business Information in the Process of Administrative Licensing (Consultation Draft)’ (司法部关于《关于强化行政许可过程中商业秘密和保密商务信息保护的指导意见（征求意见稿）》公开征求意见的通知),” published for comment in August 2020, http://www.moj.gov.cn/government_public/content/2020-08/14/657_3254208.html.

[53] Phone interview with source, February 2021.

[54] Jamie Horsley, “How Will China’s Privacy Law Apply to the Chinese State?,” New America, January 26, 2021, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/how-will-chinas-privacy-law-apply-to-the-chinese-state. As Horsley notes, state organs include China’s legislatures, courts, procuratorates, supervision commissions, and military commissions, in addition to administrative departments under the central government—the State Council—and all levels of government throughout the country.

need to seek a balance between data management efficiency for the public interest and protection of individual's privacy."[55] In an earlier scholar's version of the draft PIPL, there was in fact a longer chapter devoted to fleshing out this issue, but it was subsequently removed—reflecting the very real possibility that only vague and basic principles may be feasible within China's current political system. According to an interview with another scholar in China, further detail might have led to opposition to the law from the very "state organs" called out in the law whose data-extraction powers need to be constrained. The law may have gone as far as possible in acknowledging that government data access needs an institutional process and set of guardrails.[56] Horsley notes that the law also provides carveouts for national security and law enforcement access. Article 35 waives notice and consent requirements in matters requiring confidentiality, or where such requirements would impede performance of government duties.[57]

It is also possible that creating a more institutionalized process by which the government should make data requests could promote more data sharing between companies and the government. The new laws provide justification for security agencies to cite that could make it more difficult for companies to refuse. Multinational and domestic companies in China have used lack of legal justification as a reason to refuse to comply with data requests. To date, the only explicit data-access requirements in force appear to relate to the ride-hailing industry, as evidenced by the notice issued jointly by the Ministry of Transportation and Ministry of Public Security in 2018 after the ride-hailing company Didi refused to turn over data in the investigation of passenger murders. The notice states that ride-hailing platforms must "provide public security organs with technical interfaces to provide real-time data such as platform drivers, vehicle registration data and vehicle location, driving routes, and passenger information."[58]

The DSL, then, could create the first national law that explicitly outlines parameters for data access (beyond the vague requirements in other laws such as the National Intelligence Law requiring organizations and citizens to assist with national intelligence work.[59])

Yet the sheer vagueness of the requirements for state data access will contribute to uneven enforcement, or even to a complete lack of enforcement. The tension within the laws about government access is reinforced by developments on the ground, as companies and individuals are pushing back against data requests and surveillance technology—although much of this remains anecdotal. The *Financial Times* reports that Ant Group, the financial affiliate of Alibaba, has defied "intense government pressure" and shared just a "fraction of its customer data with China's central bank," even as the bank sought to create a pool of data for state-owned banks to assess consumer loans.[60] A recent govern-

[55] Qin Jianhang, Qian Tong, and Han Wei, "Cover Story: The Fight Over China's Law to Protect Personal Data," Caixin, November 20, 2020, https://www.caixinglobal.com/2020-11-30/cover-story-the-fight-over-chinas-law-to-protect-personal-data-101633699.html.

[56] Horsley, "How Will China's Privacy Law Apply to the Chinese State?"

[57] Horsley.

[58] Ministry of Transport of the People's Republic of China, "Urgent Notice of the General Office of the Ministry of Transport and the General Office of the Ministry of Public Security on Further Strengthening the Administration of Safety in Taxis Ordered via Online Booking and Private Passenger Car Sharing (交通运输部办公厅、公安部办公厅关于进一步加强网络预约出租汽车和私人小客车合乘安全管理的紧急通知)," September 10, 2018, https://xxgk.mot.gov.cn/jigou/ysfws/201809/t20180911_3084087.html.

[59] Chinese National People's Congress Network, "National Intelligence Law of the People's Republic of China," June 27, 2017, https://cs.brown.edu/courses/csci1800/sources/2017_PRC_NationalIntelligenceLaw.pdf.

[60] Sun Yu, "Jack Ma's Ant Defies Pressure from Beijing to Share More Customer Data," *Financial Times,* March 2, 2021, https://www.ft.com/content/1651bc67-4112-4ce5-bf7a-d4ad7039e7c7.

ment proposal led by the People's Bank of China (PBoC) would create an entity to pool credit data together from the e-commerce and payments platforms, giving the government more control over the way credit data is distributed, but this remains theoretical.[61]

Regulators in China increasingly view internet platforms' control over data as a competition issue and are seeking different ways to compel companies to share their data. The PBoC has also flagged concerns about the size of private companies such as Ant, striking out at the "inappropriate collection and control of data" by "leading internet platforms that have abused their market monopoly."

Data exchanges (discussed in more detail in Section 4) create another channel within the government and across the private sector to promote more sharing of data resources to combat data hoarding. Greater access to datasets through the exchanges strengthens the government's control and visibility, while also helping serve a variety of other purposes, from making e-government more efficient to fueling innovation in the digital economy.

INDIA

Creating increased access to data is the central motif in recent data-governance proposals, which view data as the foundational raw material for enabling domestic data businesses as well as data-driven governance by state actors. The NITI Aayog's proposals around Data Empowerment and Protection Architecture (DEPA), discussed in detail in Section 4, and the draft Non Personal Data Framework are aimed at increasing access to user data for domestic private actors and public agencies. The August 2020 DEPA explicitly identifies the need to move beyond data protection to *data sharing* and posits techno-legal arrangements to enable it. The concern, broadly stated, is that a small number of (mostly foreign) private companies, given their market dominance, currently control access to data about Indians. Therefore, breaking data silos is promoted in terms of widening the net of companies that have access to Indians' data so that "fintech or healthtech companies compete on product design, analytics, and value creation, rather than data access."[62] Access to data is also linked to the need for greater *financial inclusion* or bringing more Indians within the net of digital banking and payments. The argument is that India's 1.2 billion-strong population has vast untapped potential to generate data that can be used to create wealth. A prominent claim is that India's poor "will become data rich before they become economically rich"[63] (in contrast to those in high- or middle-income countries) and that this data trail or digital footprint will empower them to "build trust with institutions" and gain access to credit.[64]

These policy frameworks also propose a range of intermediaries—legal entities that are incentivized to enable these data transfers and are subject to regulation. They are also positioned to play a key role in ensuring that these data-access regimes do not conflict with existing or upcoming data-protection and security norms.

It's still unclear how these policies for increased and frictionless access to data interplay with data-privacy requirements. While data-protection norms that put restrictions on the

[61] Lulu Yilun Chen, "China Considers Creating State-Backed Company to Oversee Tech Data," Bloomberg, March 24, 2021, https://www.bloomberg.com/news/articles/2021-03-24/china-is-said-to-mull-state-backed-company-to-oversee-tech-data.

[62] Government of India, Niti Aayog, "Data Empowerment and Protection Architecture," August 2020, https://niti.gov.in/sites/default/files/2020-09/DEPA-Book_0.pdf, 5.

[63] "Data Empowerment and Protection Architecture," 3.

[64] Ibid.

collection and use of data have evolved rapidly over the past few years in response to both domestic and global developments,[65] India still lacks a comprehensive data-protection law and the latest version of the draft Non-Personal Data Framework has been criticized for being particularly weak on state access to data. In this context, a prominent civil-society demand has been that these developments must follow rather than precede data-protection frameworks and restrictions on government surveillance.

EU

In addition to creating more interoperable data spaces throughout the Union, the EU Data Strategy emphasizes taking greater advantage of the datasets held by government agencies.[66] This vast reservoir of datasets is identified as holding untapped value for the EU economy and society (that is to say, "accelerate the development of value-increasing information products across the EU"[67]), especially in strategic sectors. The proposal for the DGA of December 2020, announced pursuant to this strategy, takes steps in this direction.[68]

While frameworks to incentivize public sharing of government datasets exist, they have been limited to non-personal data or data that is not otherwise restricted by commercial, intellectual-property-rights, or data-protection restrictions. The DGA opens up the possibility for data sharing of personal datasets in ways that do not mitigate or undercut the GDPR, although the modalities of how that will work are not spelled out. It does not mandatorily force the sharing or reuse of government datasets, but does put in place a range of mechanisms (like a Data Innovation Board) that incentivize data sharing and create conditions for increased access to a number of actors.

There is a notable focus in the EU on preventing public-sector bodies from giving preferential treatment or exclusive access to particular private interests, and the DGA proposal includes a range of transparency requirements for agreements involving public-sector information between public and private parties, thereby avoiding exclusive deals.

In addition to government agencies, the European Commission also proposes mechanisms for the donation of data by entities for altruistic purposes like scientific research. The label "data altruism organizations" has been created and a certification process put in place for such entities that wish to share data for non-commercial purposes, for the common good—and only specific organizations with an EU certification for data altruism are to be permitted to process and store this data.

[65] This is the result both of domestic factors like the controversies around the Aadhaar biometric ID project and of global business imperatives like the possibility of adequacy under the GDPR.

[66] European Commission, "Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A European Strategy for Data," February 19, 2020, https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf.

[67] "The Data Governance Act & The Open Data Directive," data.europa.eu, March 2, 2021, https://data.europa.eu/en/highlights/data-governance-act-open-data-directive.

[68] European Commission, "Proposal for a Regulation on European Data Governance (Data Governance Act)," November 24, 2020, https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-data-governance-data-governance-act.

4. Data Exchanges

Trend 1: Policies for enabling data exchanges or similar mechanisms are increasingly common. Broadly, these are described as technical and legal mechanisms that can facilitate easier transfer of ownership and control over use of data between different entities. The practical forms and implementations of these nascent proposals remain less clear.

Summary Across China, India, and the EU, recent policy proposals recognize and promote the establishment of data exchanges or other data-sharing architecture to facilitate sharing of data between different entities. In these proposals, we see the conceptualization of datasets as commodities that can be traded seamlessly with the support of technical and legal mechanisms. The similarities, however, might end there, given that each region is promoting these techno-legal arrangements in pursuit of varied ends.

In China, data exchanges that exist in early stages—both government and private-sector-backed platforms—serve as testing grounds to experiment with policies and their economic impact, allowing policymakers to observe how rules related to data use, sharing, and transfer work and interact with one another in a mini data economy. Data exchanges fulfill a range of objectives: they facilitate greater access among private-sector companies and developments to datasets, give governments access to private data, and allow private companies to access government data.

In the EU, the DGA proposal puts forth a more narrowly tailored vision of technical and legal arrangements ("within a secure processing environment that is provided and controlled by the public sector") that enable greater sharing and reuse of data held by government agencies while still adhering to data-protection commitments under the GDPR. This will be no straightforward feat, and the details of how these arrangements will function are not provided. The very existence of the legislative proposal, however, signals an official recognition and encouragement of such future mechanisms.

In India, the officially endorsed DEPA and the Non-Personal Data Framework both put forth a detailed vision for the creation of new data transaction markets. The DEPA, for example, is based on the notion of "electronic consent tokens" for every granular piece of data, which can then be managed through a technical portal (like a dashboard) that is optimized for ease of transfer from one entity to another and based on the data subject's consent. Notably, a range of non-government stakeholders have been influential in developing these frameworks, which are unconventionally rich in terms of technical and operational detail compared to typical policy or regulatory documents.

CHINA

China's draft Data Security Law is the first national law that recognizes and promotes the establishment of *data-exchange markets*. Although the draft offers no details, official recognition of the concept of a data-exchange market grants legitimacy to the idea that data resources should be shared and bought and sold as a way to promote economic development.

As noted earlier, Chinese scholars view data-classification systems as a basis for data markets because they delineate what kind of data can be part of such an exchange and create a process that seeks to clamp down on the unregulated or illegal data brokers that have plagued China's digital economy with data security problems and rampant scams—among the factors that led the CCP to declare infringement of personal information a threat to social stability.[69] By officially backing the idea of data exchanges, the government may be looking to gain more visibility and oversight in this gray zone while creating increased availability and access to data to fuel the digital economy.

The data exchanges that exist in early stages—both government and private-sector-backed platforms—serve as testing grounds to experiment with policies and their economic impact, allowing policymakers to observe how rules and new technologies related to data use, sharing, and transfer work and interact with one another in a mini data economy.[70] China is also incorporating technology such as blockchain to mitigate privacy and security concerns.

The exchanges serve a number of distinct yet mutually reinforcing purposes. First, they are meant to help break down data silos within the government bureaucracy to help make data sharing more efficient. For years, a series of policies have called for better coordination of government data resources to combat data hoarding, and now the exchanges are one way to compel government agencies to make their datasets available in a standardized format.[71] Premier Li Keqiang described the challenge created by these data barriers in a 2018 speech in which he said that "eighty percent of China's data resources are in the hands of the government; it will be a waste not to develop and exploit this data." Since at least 2007, when the State Council Regulation on Open Government Information was adopted, government agencies have shown an uneven record when it comes to sharing their data, despite efforts to mandate disclosure of government records, including the right to request and retain government-held information.

A significant shift occurred after the State Council declared data as a factor of production in April 2020, when the government began approaching data as an economic asset. The data exchanges have become a way for the government to at once build up the digital economy and control the flow of data within the country. Kendra Schaefer observes that in this way "the government now thinks of itself as a data administrator," in the sense that it maintains control over data by making it available to society to use while keeping a finger firmly on the source.[72] At the same time, the exchanges serve the purpose of allowing the private sector to make money from data transactions by charging fees for data providers, API access, or data services. Recent statements and pilot projects offer some indication of what the markets aim to accomplish and the policies and processes in place that could expand to a national-level system in the future.

For instance, within a month of the law's release, a state media article written by the head of the Pudong Leadership Academy (a Shanghai-based institution for training top Chinese Communist Party officials) provided more details about the government's vision. In the piece, He Lisheng advocates for "the open sharing of data resources" to "unlock the

[69] As Jamie Horsley has written: "To be sure, government transparency remains uneven and often unsatisfactory. Government officials concede much of the information they proactively release is 'garbage.'" See Horsley, "Will Engaging China Promote Good Governance?," Brookings Institution, January 2017, https://www.brookings.edu/research/will-engaging-china-promote-good-governance; and "How Will China's Privacy Law Apply to the Chinese State?"

[70] Interview with Kendra Schaefer by telephone, Beijing and New York, February 2021.

[71] The 13th Five Year Plan for Informatization (2016–2020) calls for building a national data resource. See the State Council, the People's Republic of China, "State Council Releases Five-Year Plan on Informatization," December 27, 2016, http://english.www.gov.cn/policies/latest_releases/2016/12/27/content_281475526646686.htm.

[72] Schaefer interview, February 2021.

spillover effect of the data economy" and to bring about "industrial and economic transformation in areas from automated factories and smart manufacturing to smart cities and financial technology."[73] He also explains that a data-classification system will be a fundamental part of creating the "orderly exchange of data" for "value creation and promoting the efficient development of the digital economy."

In practice, data is transferred through a combination of forces: the sale of datasets; the presence of free, downloadable datasets online; API service providers; and data services like data visualization, analysis, and cleaning. According to Schaefer, "[the data exchanges are] basically a shopping mall for data products and services."[74] Services include allowing access to datasets for querying without transferring the dataset itself. One example is a service provided by the Shanghai Data Exchange Corp called "China Audience Profile." The Exchange claims to aggregate data from dozens of data suppliers (many are registered exchange members) from different industries, providing integrated analysis for data buyers who want to understand more about what kinds of products users are looking for, as well as variables like users' buying power, preferences, habits, and geographical features, to project market demand.

Provincial government pilot zones offer a lens onto how the exchanges serve as testing grounds for creating a process and framework to share data that could be expanded to a national system in the future. Tianjin, for example, published a set of interim rules (for public comment) that includes requirements for the scope of data that can be traded: data must be "legally obtained, processed such that the original data generator can no longer be identified or recovered"; also, trading certain kinds of data, such as "data related to national security, public security, and personal privacy," is prohibited.[75] In theory, the type of data that is traded is regulated under the laws and standards of China's data-protection regime. Thus, it should be anonymized, or obtained through consent, although the inevitable privacy risks of re-identification remain an area to scrutinize as these markets develop.

Two of the largest "data-exchange institutions" (数据交易所) are the Guiyang Global Big Data Exchange (贵阳大数据交易所)[76] and Shanghai Data Exchange Corp (上海数据交易中心).[77] Their models are similar: any company with anonymized data can apply to those exchange institutions to be what are called "exchange members," which allows them to sell data to other registered members (or sometimes verified and certified third parties). Companies involved are often either businesses selling their consumer data directly to data buyers, or big data companies compiling data sold to them by consumer-facing businesses. The buyers, registered exchange members, could be app developers or people using the data for business development, marketing, or product development.

INDIA

Recent data-governance policies stand apart for the novel and detailed (if convoluted) techno-legal architectures they propose. These go beyond stating legal principles to

[73] He Lisheng, "推动数据由资源向要素转化 (Promote the Transformation of Data from Resources to Elements)," Xinhua, August 25, 2020, http://www.xinhuanet.com/tech/2020-08/25/c_1126408564.htm.

[74] Schaefer interview.

[75] 《天津市数据交易管理暂行方法（征求意见稿）》公开征求意见, http://tj.sina.com.cn/news/zhzx/2020-07-31/detail-iivhuipn6009892.shtml.

[76] Guiyang Global Big Data Exchange website, http://www.gbdex.com/website/view/aboutGbdex.jsp.

[77] Shanghai Data Exchange Corp website, https://www.chinadep.com.

present a combination of regulatory and technical systems that will together meet policy objectives. The NITI Aayog report notes that the DEPA could "do for India's data ecosystem what the TCP/IP Internet protocol or GPS—both powerful examples of American public digital infrastructure—did for communication and navigation respectively: introduce a new possibility that creates a Cambrian explosion of novel products and services that empower people."[78]

Both the DEPA and the draft Non-Personal Data Framework put forth techno-legal frameworks for the creation of new data-transaction markets. The DEPA, for example, is based on the notion of "electronic consent tokens" for every granular piece of data, which can then be managed through a technical framework that is optimized for ease of transfer from one entity to another and based on the data subject's consent. Given anticipated difficulties with individuals' managing this scale and depth of consent permissions, the framework proposes the creation of business entities known as *consent managers* that will act as intermediaries and manage permissions on behalf of users in exchange for monetary incentives. While this framework is posited as having value across data-driven businesses, it can be hard to conceptualize in the abstract. Sectoral regulators have committed to implementing versions of these new data-transaction markets. The financial regulator, the Reserve Bank of India, has already operationalized a system of "account aggregators" modeled on the DEPA in December 2020, but still at a very limited scale. In August, India's Prime Minister also endorsed the creation of a "National Health Stack" with a federated electronic health data system that is said to be modeled on the DEPA.

Similar to the DEPA, the draft Non-Personal Data Framework also proposes a range of new legal and market players (e.g., data trustees), new legal requirements for existing businesses ("data businesses" as a legal category), and an interoperable technical grid for seamless data sharing. Parminder Jeet Singh, a civil society member of the committee that drafted the framework on non-personal data acknowledged that organizations with a significant degree of "techno-organizational capacity" would be well suited to play the role of data trustees as envisioned in the framework. [79] This raises questions about the feasibility of these models in view of both the current levels of connectivity and the maturity of India's data markets, as well as unclear incentives for the creation of technically sophisticated market intermediaries to develop.

Another recurrent critique has been that the framework vests most rights and interests in the nebulous and undefined notion of "community". As policy expert Malavika Raghavan noted in a public workshop, the idea that data-driven communities would organically develop and organize around their common kind of interest in data is hard to imagine in the present context.[80]

A range of non-government stakeholders have been influential in developing the frameworks that are now being endorsed and implemented by government agencies. The government's NITI Aayog paper on DEPA, for example, acknowledges the proposal as a joint effort with several members of iSPIRT, a nonprofit think tank with membership from prominent members of Indian companies like Infosys, Paytm, and PhonePe. iSPIRT describes itself as "converting ideas into policy proposals to take to government

[78] "Data Empowerment and Protection Architecture," 5.

[79] Medianama, " Discussion on the Governance of Non Personal Data," video, 3:55:20, January 15, 2021, https://www.youtube.com/watch?v=9ynaYd1_A3A.

[80] See Soumyendra Barik, "#NAMA: Issues with definition of communities, public good, and unabated sovereign access to non-personal data," Medianama, January 22, 2020.

stakeholders"[81] and has courted controversy for its unusually deep influence and access within various sectors of government.[82] Similarly, the committee appointed by the government to prepare the draft Non-Personal Data Framework includes members of the technical and business community in India, as well as Parminder Jeet Singh, a civil-society advocate who has been vocal in support of community rights over data[83] as well as restrictions on cross-border data transfers in the national interest.[84]

EU

The proposal for the Data Governance Act of 2020 incentivizes the creation of techno-legal arrangements for increased availability, sharing, and reuse of data. It also creates a series of new intermediaries for the data market (such as data-sharing services and data-altruism organizations) that are supposed to act as "as a tool to facilitate the aggregation and exchange of substantial amounts of . . . data." [85] These data intermediaries are required to maintain independence and comply with strict requirements, to prevent "misaligned incentives that encourage individuals to make more data available for processing than what is in the individuals' own interest." The proposal also includes a certification or labeling framework, including subsequent monitoring of compliance with these requirements. The idea has been compared to the concept of "data trusts" that is being explored in several jurisdictions.[86]

As noted earlier, the Data Governance Act includes sharing and reuse of personal data held by the government as long as it is in compliance with existing GDPR requirements. The proposal does not prescribe how this balance will be achieved, but does indicate or gesture toward the use of modern technological tools that could allow this. For example, it states that public-sector bodies can impose obligations for the data only to be accessed "within a secure processing environment that is provided and controlled by the public sector." While the draft law doesn't provide any further details on how this would work, this could endorse nascent research proposals for mechanisms that retain the locus of computation with the dataset owner and prevent the sharing of raw data with potentially untrustworthy third parties.[87]

[81] iSPIRT website, https://ispirt.in.

[82] See Rohan Venkataramakrishnan, "Co-Founder of UIDAI-Associated Outfit Admits to Anonymously Trolling Aadhaar Critics on Twitter," Scroll.in, May 23, 2017, https://scroll.in/article/838468/co-founder-of-uidai-associated-outfit-admits-to-anonymously-trolling-aadhaar-critics-on-twitter.

[83] Parminder Jeet Singh, "Community Data in the Draft e-Commerce Policy", March 2019, https://www.medianama.com/2019/03/223-community-data-in-the-draft-e-commerce-policy.

[84] Parminder Jeet Singh, "Taking National Data Seriously," Hindu, October 17, 2019 https://www.thehindu.com/opinion/lead/taking-national-data-seriously/article29716990.ece.

[85] European Commission, "Proposal for a Regulation of the European Parliament and of the Council on Contestable and Fair Markets in the Digital Sector (Digital Markets Act)." The DGA press release notes that the act promises to "create new EU rules on neutrality to allow novel data intermediaries to function as trustworthy organisers of data sharing." See the official website of the European Union, "Commission Proposes Measures to Boost Data Sharing and Support European Data Spaces," November 25, 2020, https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2102.

[86] See Aline Blankertz and Louisa Specht, What regulation for data trusts should look like, SNV, July 2021, https://www.stiftung-nv.de/sites/default/files/regulation_for_data_trusts_0.pdf; Chris Martin, "'Data Trusts' Can Support Competing Interests, Studies Find," Pinsent Masons, Out-Law News, April 17, 2019, https://www.pinsentmasons.com/out-law/news/data-trusts-competing-interests-study.

[87] See Lisa M. Austin & David Lie, Safe Sharing Sites, New York University Law Review, 94 (4) October 2019 https://www.nyulawreview.org/wp-content/uploads/2019/10/NYULAWREVIEW-94-4-AustinLie.pdf

Coda: Beyond the Binary

Openness, and in particular the moniker of the "free and open internet" has been a canonical part of US government policy, associated with policy regimes promoting unrestricted global data flows and preventing data localization, as well as pushing back against restrictions on freedom of expression online, including abroad. Experts have pointed to the Chinese government's aggressive internet censorship regime as a failure of the "US open internet project".[88] In this section, however, we focus on how recent moves in US and Chinese data governance policy might point to the waning relevance and influence of "open" (and its corollary, "closed") as a binary lens with which to classify and evaluate data governance policy in the first place.

In September 2020, the Chinese Foreign Ministry released a set of principles to assert China's leadership in global data governance called the Data Security Initiative (DSI).[89] Through the principles outlined in the DSI, Beijing put a stake in the ground of advocating for an open global internet in reaction to recent measures by the United States.[90] The initiative represented China's response to former Secretary of State Mike Pompeo's Clean Network initiative, which called for purging Chinese hardware and software from US ICT infrastructure—a seismic shift in the State Department's long-held position around an open internet. Paradoxically, the language of China's DSI turns the canonical idea of the "open internet," championed by US government and tech companies for decades, on its head. At the very moment that US policy is closing off to Chinese technology under broad national security claims, Beijing is conducting global diplomacy calling for more openness and transparent criteria in assessing security risk. The first principle reads: "First, treat data security objectively and rationally, and work to maintain open, secure, and stable global supply chains."

In contrast to the approach under the Clean Network initiative, the Biden administration explicitly stated it is "committed to promoting an open, interoperable, reliable and secure Internet."[91] The statement appeared in a fact sheet for a new Executive Order (EO) revoking and replacing prohibitions on transactions with Wechat, TikTok, and other software applications. In their place, the new EO "directs the use of criteria-based decision framework and rigorous, evidence-based analysis" to evaluate risk in an effort to maintain openness and security at once. The way in which the Biden administration implements this EO will be an important litmus test for the prominence of "openness" as a guiding value for future US policy. Indeed, the EO of May 2019 that the Biden administration left in place (EO 13873) and the Commerce Department rule that implements it, grants the Commerce Secretary broad authority to ban any "transaction" that risks catastrophic effects on the security of the United States or the digital economy.[92] This latitude gives

[88] Jack Goldsmith, "The Failure of Internet Freedom", Knight First Amendment Institute, June 18th 2018, https://knightcolumbia.org/content/failure-internet-freedom.

[89] Ministry of Foreign Affairs of the People's Republic of China, "China proposes the 'Global Data Security Initiative'," September 8, 2020, https://www.fmprc.gov.cn/web/wjbzhd/t1812947.shtml.

[90] The main content of the initiative has been provided in a translation by the DigiChina Project. See Graham Webster and Paul Triolo, "Translation: China Proposes 'Global Data Security Initiative'," New America, September 7, 2020, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinese-proposes-global-data-security-initiative/.

[91] FACT SHEET: Executive Order Protecting Americans' Sensitive Data from Foreign Adversaries, June 9 2021, https://www.whitehouse.gov/briefing-room/statements-releases/2021/06/09/fact-sheet-executive-order-protecting-americans-sensitive-data-from-foreign-adversaries/.

[92] Securing the Information and Communications Technology and Services Supply Chain, January 19 2021, https://www.federalregister.gov/documents/2021/01/19/2021-01234/securing-the-information-and-communications-technology-and-services-supply-chain.

the government broad discretion in ways that mirror the opaque nature of China's own cybersecurity regime. Indeed, the idea of reciprocity in turning Chinese digital controls and market restrictions back on China is also gaining momentum in US policy discussions.[93] Meanwhile the Clean Network initiative remains as the unstated de facto policy, while there is a growing bipartisan consensus around the need to wall off American data and infrastructure from Chinese companies deemed untrustworthy.

Meanwhile, as US policymakers consider new tools to restrict access to American citizens' data, the Chinese government has signaled that it may be amenable to allowing more flexible data flows out of the country. The draft Data Security Law mentions the free flow of data twice.[94] Additionally, a number of provincial governments, including Hainan, announced pilots to allow the free flow of data to drive economic development. According to analyst Xiaomeng Lu, the Chinese government may be more amenable, given the economic slowdown occasioned by COVID-19, to allowing more cross-border data flows as part of a broader effort to attract much-needed foreign investment.[95]

Broad data-localization requirements remain in place under the Cybersecurity Law regime—and anecdotal conversations with company executives in China suggest the government will continue to require that significant swaths of data be stored on local servers. Nevertheless, it is worth highlighting the paradox that at the very moment when US policymakers may be shifting more toward an acceptance of a form of US data sovereignty, in China, at least at the margins, some voices may be pulling in the opposite direction, primarily driven by a growing recognition of the economic utility of data.

These developments prompt us to re-evaluate binary frames of analysis (such as open versus closed) which, over time, produce and sustain their own blind spots. The analysis in the body of this report demonstrates that flattening data policy into the "China model" or the "US model" (or even the European so-called "third way"[96]) obscures both the contradictions within these national policies, and overlooks their inter-dependencies.

[93] Matt Perault and Samm Sacks, "A Sharper, Shrewder U.S. Policy for Chinese Tech Firms," Foreign Affairs, February 19, 2021, https://www.foreignaffairs.com/articles/united-states/2021-02-19/sharper-shrewder-us-policy-chinese-tech-firms.

[94] Article 5 declares that "the state ensures . . . the lawful and orderly free flow of data." And Article 11 notes that "the state is to actively . . . promote the safe and free flow of data across borders.." Translation by Jeremy Daum, "Data Security Law of the People's Republic of China (Draft) (Second Deliberation Draft)," China Law Translate, https://www.chinalawtranslate.com/en/data-security-law-draft-2/.

[95] Xiaomeng Lu, "Is China Changing Its Thinking on Data Localization?", Diplomat, June 4, 2020, https://thediplomat.com/2020/06/is-china-changing-its-thinking-on-data-localization/.

[96] See Frederike Kaltheuner, "A New Tech Cold War? Not for Europe", AI Now Institute, July 29 2021, https://medium.com/@AINowInstitute/a-new-tech-cold-war-not-for-europe-4d4f2f8079b6

Acronyms

CCP	Communist Party of China
DEPA	Data Empowerment and Protection Architecture
DGA	Data Governance Act
DSI	Data Security Initiative
DSL	Data Security Law
EDPS	European Data Protection Supervisor
EU	European Union
GDPR	General Data Protection Regulation
ICT	Information and Communication Technology
IDC	International Data Corporation
MeitY	Ministry of Electronics and Information Technology
PBoC	People's Bank of China
PIPL	Personal Information Protection Law

Acknowledgements

The authors would like to thank Michael Rose, Kendra Schaefer, Udbhav Tiwari, Graham Webster, and colleagues at Yale Law School's Paul Tsai China Center—especially Robert Williams, Jamie Horsley, and Karman Lucero—for reviewing and providing detailed feedback on earlier drafts of this report.

Yale Law School
Paul Tsai China Center
耶鲁大学法学院蔡中曾中国中心

AI NOW

NEW AMERICA