

Framing Policy Visions of Big Data in Emerging States

Laura C. Mahrenbach
Technical University of Munich

Katja Mayer
University of Vienna

ABSTRACT

Background *Emerging states, such as Brazil, India, and China (the BICs), have big plans for big data and digitalization. Research has identified distinct policy visions regarding how technological advances can facilitate economic development and improve governance.*

Analysis *This article examines how BIC governments frame data-driven ambitions across the diverse issue areas in which governments plan to use big data, as well as how they frame the role(s) of the government and citizens in the era of big data.*

Conclusion and implications *We find clear differences in discussions of big data across the BICs and across issue areas. Moreover, we show the societal changes that governments seek to effect using big data vary greatly in scope, with Brazil and India seeking more fundamental changes than China.*

Keywords *Data policy; Big data; Emerging powers; frames; Digital strategy; Global South*

RÉSUMÉ

Contexte *Les États émergents, tels que le Brésil, l'Inde et la Chine (les BIC), ont de grands projets pour Big Data et la numérisation. La recherche a identifié des visions politiques distinctes concernant la façon dont les progrès technologiques peuvent faciliter le développement économique et améliorer la gouvernance.*

Analyse *Cet article examine la manière dont les gouvernements BIC définissent les ambitions basées sur les données dans les divers domaines problématiques dans lesquels les gouvernements envisagent d'utiliser big data, ainsi que la façon dont ils définissent le ou les rôles du gouvernement et des citoyens à l'ère des big data.*

Conclusion et implications *Nous constatons des différences claires dans les discussions sur big data entre les BIC et entre les domaines problématiques. De plus, nous montrons que les changements sociétaux que les gouvernements cherchent à effectuer en utilisant big data*

Laura C. Mahrenbach works as a post-doctoral researcher at the Chair of Computational Social Sciences and Big Data, Hochschule für Politik München an der Technischen Universität München, Germany. Email: laura.mahrenbach@hfp.tum.de. **Katja Mayer** is a sociologist working at the interfaces of science, technology, and society, Department of Science and Technology Studies, University of Vienna, Austria. Email: katja.mayer@univie.ac.at.

varient considérablement, le Brésil et l'Inde recherchant des changements plus fondamentaux que la Chine.

Mots clés *Politique de données; Big data; Puissances émergentes; cadres; Stratégie numérique; Sud global*

Introduction

Over the last decade, “big data” have moved from buzzword to political tool. Governments in the Global North and the Global South are developing—and in some cases already implementing—plans to leverage technological advances for better economic, social, and developmental outcomes (United Nations Economic Commission for Africa, United Nations Development Programme, International Development Research Centre, and World Wide Web Foundation, 2017; European Commission, 2014). Crucial to these efforts are government policy visions that set goals for incorporating data into policymaking and contain plans for goal achievement. Research has shown that the big data policy visions of larger states in the Global South, including emerging market countries such as Brazil, India, and China (henceforth the BICs), link big data to improved governance and economic development (Mahrenbach, Mayer, & Pfeffer, 2018a; CIS, India 2018; Ruan, 2018). The extensive media campaigns, international conferences, and speechifying undertaken by the BIC governments to promote big data initiatives (see Srinivasan & Johri, 2013, or Zheng, 2013) indicate that these governments view communication as a critical tool to cultivate support for and implement such visions. In fact, studies have shown that what is said and to whom affects citizen evaluations of programmatic and organizational legitimacy (Deephouse & Suchman, 2008). Governments consequently target their messages according to the audience they are addressing and the goals they seek to achieve (Gronau & Schmidtke, 2016).

And yet, despite the proliferating uses of big data in public sector activities (Fredriksson, Mubarak, Tuohimaa, & Zhan, 2017) and the related dangers big data pose for citizen privacy protection and the autonomy of government decision-making from corporate actors (Taylor & Schroeder, 2015; Zeng, 2016), the issue of how big data policy visions are communicated to the actors who will be affected by them remains under-examined. While existing studies highlight the general challenges and opportunities linked to big data, they often fail to examine how and why these challenges and opportunities—and government aims—may vary. Performing a qualitative content analysis of BIC government strategy documents to map how BIC governments frame big data in relation to two factors adds nuance to current understandings of big data initiatives. First, the scholarship cited above suggests governments may communicate big data plans differently in relation to different issue areas. This is because policy goals vary across issue areas, implying that big data may be more, less, or differently useful depending on the issue area of focus. As a result, this article examines how the BICs frame big data in seven issue areas. Second, scholars argue that digital governance is a means of “changing the relationship between citizens and the government” (Mossberger, Yonghong, & Crawford, 2013, p. 351). Big data programs such as China’s social credit score, which derives an individual’s score from their collective online presence and has implications for that person’s access to healthcare, education,

and government services, underline how wide-ranging this impact can be. Given this, from a communication perspective, governments should seek to influence how that relationship develops. Consequently, this article examines how BIC governments frame the roles of governments and citizens in reference to big data.

The findings indicate not only that BIC governments discuss big data differently within and across issue areas as well as in relation to different actors. They also demonstrate that the societal changes governments seek to effect using big data vary greatly in scope. Both recognitions are crucial if scholars and policymakers are to accurately assess the extent to which big data disrupt and preserve existing political processes and relationships.

Big data and the BICs

Big data are generally defined in reference to the 3Vs: volume, which refers to the exponential growth in the size of data; variety, which refers to the ever-more diverse types of data; and velocity, or the speed at which data can be analyzed (Laney, 2001). However, subsequent studies have underlined the importance of moving beyond these basic characteristics to understand how collecting, curating, and utilizing such data are transforming social practices and relations (Kitchin, 2014). After all, different concepts can lead to different “decisions, judgments and notions of value” (Beer, 2016, p. 5) derived from big data. Given this article’s focus on the communication of BIC policy visions to citizens, as well as the expectation that governments tailor messages to the preferences of their target audience, the definition of big data is left open here. Specifically, it is assumed that, just as governments use different policy visions to advance different ambitions, so too will governments draw on different conceptualizations of “big data” in different situations.

Why focus on the BICs? China, India, and Brazil are among the top four countries in terms of internet populations (calculated using data from World Bank, 2018). With 733 million internet users, China alone ranks higher than all the European Union countries together. Similarly, India, with 391 million internet users, ranks higher than all of North America. This implies that how BICs communicate their big data visions is important on a macro level. These governments possess the resources as well as the datasets necessary to innovate in relation to big data; consequently, understanding how they do so is relevant for other countries interested in doing the same. Furthermore, the BICs have already begun to incorporate big data into government activities. For instance, the Indian government’s Digitize India Platform provides the technological tools necessary “for government agencies to transform themselves into digital enterprises” (Government of India, 2018b). Similarly, the Chinese government’s most recent five-year economic plan provides implementation details for the National Big Data Strategy (Central Committee, 2015). As such, BIC government communications appear well-suited to examining how governments discuss big data across issue areas and in relation to citizens.

Before proceeding to the analytical framework, it is important to examine the context within which BIC governments communicate about big data. In the early 2010s, big data emerged as a business paradigm (Manyika, Chui, Brown, Bughin, Dobbs, Roxburgh, & Byers, 2011), soon becoming an economic imperative (Rieder, 2018).

Policymakers accordingly emphasized novelty and excitement in political discussions of big data. Big data were painted as the “new oil,” a “gold mine,” a “game changer” (2018, p. 92), that is, as a strategic resource and advantage. Also, media discussions have amplified such discussions by linking data with power competition at the international level, see, for example, “China and the US Compete to Dominate Big Data” (*Financial Times*, 2018). In fact, BIC policy visions and regulations regarding big data are not developed in a geopolitical vacuum. The European General Data Protection Regulation, which came into effect in May 2018, externalizes European Union (EU) law for companies wishing to do (digital) business in the EU. This particularly affects China, the EU’s second largest importer and exporter of commercial services, but also Brazil and India, whose combined trade in services with the EU totalled U.S.\$50.3 billion in 2016 (European Commission, 2019; World Trade Organization, 2019). It is thus perhaps not surprising that China introduced complementary rules governing data use and storage the same month (Lucas, 2018), nor that India and Brazil signed data protection laws shortly thereafter.¹ At the same time, media platforms have given voice to fears linked to big data, including worries about privacy violations and exploitation. *The Times of India* has argued that data concentration in the hands of governments and tech companies could constitute data colonisation (Doval, 2017), while the newspaper *O Estado de S. Paulo* detailed the personal costs of people blindly following the big data trend (Rocha, 2013). These fears also appear to have demonstrable political effects. For example, in 2016, the Indian government rejected Facebook’s offer of free but limited internet access for Indian citizens. This reflected an extensive public consultation process as well as new regulations forbidding “discriminatory prices to consumers based on the content, applications, services or any other data being used” (Telecom Regulatory Authority of India, 2017, p. 4).

Although measuring the impact of geopolitical, corporate, and civil society pressure falls beyond the scope of this case study, it is clear that the communicative context within which the BICs discuss big data is characterized by great scepticism, great promise, and significant agency. As such, one can conceive the context within which BIC governments conceptualize big data as one populated by both challenges and opportunities. Examining how policy documents define those challenges and opportunities can shed light on both the scope of change that officials expect big data to stimulate in these countries, as well as nuances regarding how officials will/should effect change.

Analytical framework

Reviewing the literature on big data in the public sector, Cecilia Fredriksson, Farooq Mubarak, Marja Tuohimaa, and Ming Zhan (2017) concluded that contemporary studies often focus on general observations, such as the potential for big data to increase trust in government, and ignore the nuances relevant to government visions of big data in different sectors (i.e., in political versus economic contexts). In fact, a report by the Organisation for Economic Co-operation and Development (OECD, 2014) highlights that some sectors, including public administration, health, and education, are more likely to benefit substantially and rapidly from “data-driven innovation” than others. This article takes a more detailed view of BIC government communication in relation to big data.

Building on calls to engage in more qualitative research about big data, especially in emerging states (Arora, 2016; Kshetri, 2014), this study presents the results of a content analysis of 13 BIC policy documents.² The corpus includes strategy papers, decrees, white papers, and speeches, as well as the descriptive sections of government regulations. An overview of the corpus, including descriptions of each document, appears in Appendix 1. Relevant documents were identified through targeted internet searches as well as in discussion with experts living and working in BIC countries. The coding scheme combined deductive elements (e.g., regarding which issue areas governments might link with big data) with inductive ones (e.g., the specific ways in which these issue areas are linked with big data) to ensure a complete picture of the textual data (Hsieh & Shannon, 2005). Building on existing literature, it was expected that governments would link big data to the following issue areas: politics (e.g., e-government), the economy (e.g., manufacturing), environment (e.g., climate change), health (e.g., improved medical treatment), security (e.g., threats to citizens), education (e.g., training programs), and science/technology (e.g., artificial intelligence).

In addition, the presence of normative frames, such as accountability or transparency was sought. Theoretically this made sense as framing theorists argue that politicians discuss political issues in a certain way (the frame), to create situations in which they are more likely to achieve their goals (Daniels & Martin, 1998; Gamson & Modigliani, 1989). This is true in both democratic and non-democratic settings (Fearon, 1994; Zeng, Stevens, & Chen, 2017). Empirically, frames are useful because big data programs face particular legitimation difficulties: it is hard to anticipate how the costs and benefits of such programs will be shared across the population and their achievements, for instance, the creation of technological infrastructure, are generally not visible to the general population (Srinivasan & Johri, 2013). Thus, *how* actors discuss big data should be especially pertinent to understanding their use. Frames provide insight by enabling us to systematically evaluate similarities and differences in how the BICs discuss big data vis-à-vis diverse issues and actors.

Findings

In line with previous analyses (Fredriksson et al., 2017), rhetoric highlighting the efficiency and effectiveness of big data was present in all corpus documents. To move beyond this “marketing campaign” approach to big data (Rieder, 2018, p. 92), this section takes a more nuanced look, examining first how governments frame big data in diverse issue areas and, subsequently, citizens’ and governments’ role(s) in these visions.

Big data across issue areas

There were few differences among the BICs in some of the issues examined. For instance, regarding health, all three BICs discussed big data collection/analytics as a means of improving the quality of medical care and facilitating positive health outcomes (e.g., Central Committee, 2015; Ministro da Ciência, Tecnologia, Inovações e Comunicações, 2018), thereby adopting a productivity and progress frame. Similarly, all BICs viewed big data as a tool for monitoring environmental conditions, improving government responses to natural disasters, and facilitating economic development (see Government of India, 2012b; Ministro da Ciência, Tecnologia, Inovações e

Comunicação, 2018; State Council Information Office, 2010), combining productivity and progress and transparency frames.

However, there were notable differences in the fields of politics, economy, education, and security. Starting with politics, all three BICs see big data as an opportunity to transform political structures and processes. China and India focus their attention on government institutions. For example, India underlines the necessity of digitizing public services, and envisions big data as a means to improve cooperation among government departments, thus adopting a collaboration frame (Government of India, 2012b). Similar expectations are apparent in Chinese documents (China State Council, 2015b) and in a few Brazilian documents (e.g., Ministro da Ciência, Tecnologia, Inovações e Comunicação, 2018). However, Brazilian documents primarily focus on big data's potential for improving political communication between the government and citizens, for instance, via the creation of dialogue platforms (Ministry of Planning and Budget, 2016) or by ensuring the government's digital services are compatible with diverse operating systems and device types (Congresso Nacional, 2014). As such, a participation frame dominates Brazilian political discussions of big data.

Turning to the economy, China and Brazil depict big data as an opportunity to boost productivity and enhance the international competitiveness of various industries (China State Council, 2015b; Ministro da Ciência, Tecnologia, Inovações e Comunicação, 2018). For instance, Brazil plans to expand its supercomputing infrastructure and create new data centres incorporating technological advances in robotics, the internet of things, et cetera (Ministro da Ciência, Tecnologia, Inovações e Comunicação, 2018). Both countries encourage private actors to take the lead in pursuing these opportunities, limiting the government's role to ensuring access to data (Ministro da Ciência, Tecnologia, Inovações e Comunicação, 2018) or funding research and development (China State Council, 2015b). Indian documents, in contrast, simply advocate for more cooperation among industry and universities as a means of maximizing big data's impact on the Indian economy (Government of India, 2012b). As such, while all three BICs demonstrate a normative frame of productivity and progress in relation to the economic impact of big data, the Indian frame appears less actionable than that of its fellow BICs.

Education was the third major issue discussed in which the BICs' emphases differed. Drawing on transparency and inclusion frames, Brazilian documents argued that making big data accessible to the public will increase the quality and accessibility of education (e.g., Ministro da Ciência, Tecnologia, Inovações e Comunicação, 2018). China, in contrast, centres its vision on professionals rather than citizens, depicting the training of data scientists and engineers as a crucial step in restructuring China's industrial base (China State Council, 2015a, 2015b). Finally, India's focus is on the educational impact of big data for government officials. Documents argue that, in light of programs such as *Aadhaar*, a demographic and biometric identification program, officials must not only be able to analyze big data but must also be trained regarding privacy rights, data security, and data quality (Government of India, 2012b). Both India and China adopt a productivity and progress frame, albeit targeting different types of progress (political versus economic) and evaluating progress in reference to different actors (officials versus professionals).

Finally, while “cybersecurity” is a buzzword in Chinese and Brazilian documents, depicted as both the cause of and solution to contemporary security problems (e.g., Ministry of Planning and Budget, 2016), it rarely appears in Indian documents. However, all three countries underline citizens’ rights to such security, as well as governments’—and in China businesses’—responsibilities in securing it (China State Council, 2015b; Congresso Nacional, 2014). In addition to this frame of individual rights, India and China adopt a law and order frame regarding government responsibilities for protecting data privacy, with documents proposing government mechanisms to protect privacy and specifying penalties for failing to follow regulations (Central Committee, 2015; Government of India, 2015). Perhaps the most striking difference in the security issue area addresses the link between big data and national security that features prominently in the Chinese documents (China State Council, 2015b), but is mentioned only in passing in Brazilian and Indian documents (Congresso Nacional, 2011; Srikrishna, 2017).

Government-citizen relationship

All documents expect changes in political culture as a result of big data. Accountability, transparency, and participation are the main normative frames linked to governments’ role(s) in this new relationship. Chinese documents frame governments, both national and local, as potentially corrupt entities (Central Committee, 2015). Big data should make it easier for citizens to organize to identify and address this type of behaviour. Furthermore, big data enables new channels of political communication for Chinese citizens (Central Committee, 2015). Consequently, Chinese documents depict big data as a tool for increasing government accountability and promoting citizen participation in government. In contrast, the transparency and participation frames are more prevalent in Brazilian documents. The Brazilian government enables and encourages citizen participation in policymaking by making big data accessible and, as in China, by creating communication platforms (Ministry of Planning and Budget, 2016). Finally, Indian documents share the participation frame with fellow BICs but additionally (similar to China) assume an accountability frame. The National Data Sharing and Accessibility Policy, for example, argues that the government must collect, classify, and make available “enough” data for citizens to hold the government accountable (Government of India, 2012a).

Turning to citizens’ role(s), there are two primary normative emphases in the documents. First, all three countries extensively linked big data with the frame of privacy protection.³ The BIC governments envision citizens’ role in relation to privacy as passive: citizens should be protected, not protect themselves. Along these lines, Brazil and India propose government training programs to raise officials’ awareness of privacy and security issues and advocate for international best practices of data protection (Rousseff, 2013 ; Srikrishna, 2017). In other words, privacy protection is primarily a political issue and, in India in particular, one in which the government should play the dominant role (Srikrishna, 2017). In contrast, Chinese documents discuss privacy in the context of economic activity: privacy protection is the responsibility of internet service providers (ISPs) and frequently linked to information security (e.g., protecting trade secrets; see State Council Information Office, 2010).

The other big data frame discussing citizens' role(s), the participation frame, grants citizens a more active role. Digital India envisions digital channels via which citizens can engage government officials: at the time of writing, the homepage featured 30 active discussion threads, covering issues as diverse as waste management, health, infrastructure, and child protection (Government of India, 2018a). Brazilian documents highlight similar participatory initiatives, for instance, Participa.br, which resembles Digital India. They additionally commit to facilitating communication via social media platforms (Ministry of Planning and Budget, 2016). In contrast, China encourages participation through incentive structures, including financial incentives for businesses with research that helps the government reach social goals (Li, 2016) and for citizens who participate in state-sponsored crowdsourcing efforts to innovate state-owned enterprises (Central Committee, 2015). As such, the Chinese envision a more indirect role for the government in protecting citizen privacy (e.g., via the regulation of ISPs) and promoting citizen participation (e.g., via incentives) than is true for Brazil and India, where the governments explicitly play the guiding role in framing privacy regulations and creating mechanisms to directly encourage participation.

Discussion

This article seeks to contribute to the greater scholarly endeavour of exploring “the concept of big data in historical, political and sociological terms” (Beer, 2016, p. 1). Specifically, it moves beyond the mythmaking of big data as a force for general good (efficiency, trust, etc.) by examining how the BIC governments discuss big data in reference to a variety of issue areas, as well as the impact big data is expected to have on the government-citizen relationship. In so doing, it sheds light both on how the BICs seek to promote change via the use of big data as well as the scope of that change.

Starting with the former, study findings confirm the BICs view of big data as both a challenge and an opportunity. On the one hand, statements depict big data as a challenge in the fields of education (India/China) and security (BICs). Regarding education, big data demands new training for both government officials and professionals. The productivity and progress frame implies that mastering the educational challenge of big data is crucial for achieving diverse economic and political goals. On the other hand, big data is viewed as an opportunity in the realm of politics (BICs), the economy (BICs), health (BICs), and education (Brazil). What kind of opportunity big data represent differs across the BICs. For example, while all three BICs viewed big data and digital communication as opportunities to improve political communication among actors, they disagreed regarding which actors. India and China, for instance, focused on intra-government communication, while Brazil emphasized citizen-official communication. This implies that India and China see big data as a functional opportunity, whereas Brazil may be more politically motivated. Additionally, the BICs disagree regarding how to make the most of the opportunities big data offer. For instance, all three countries employed the productivity and progress frame in relation to the economy. However, while Brazil and China focused on using big data to enhance existing industrial processes and boost competitiveness, that is, as an opportunity to optimize existing production, India describes developing new auditing techniques and technological platforms, that is, as an opportunity for economic innovation and novelty.

Succinctly, even though the BICs collectively view big data as a challenge and opportunity, they differ in how they discuss these challenges and opportunities within a given issue area, across issue areas, and across the BICs.

Turning to the scope of change sought, this study shows Brazil and India seek more fundamental changes within their society, and particularly government-citizen relations, arising from big data than does China. For example, all three BIC governments saw big data as a means of collecting more information from and providing more information to citizens. The Brazilian and Indian governments hoped doing so would enhance citizen participation in policymaking, thereby opening a door for citizens to play a greater role in setting government priorities and changing government policies. As such, Brazil and India raise the possibility of fundamental changes to political processes arising from big data. In contrast, Chinese documents depict information gains from big data as a means for citizens to identify political corruption and hold governments accountable. Here big data enable only modest change, enabling citizens to police officials but not empowering them to set political tone or alter government policy. A similar relationship is evident in how the documents depict citizens' participation in government-created digital discussion forums. In Brazil and India, the empirical breadth of platforms such as Participa.br or Digital India, as well as the potential to initiate discussion on topics, at least theoretically enables citizens to provide input across the entire spectrum of government policymaking. In contrast, digital discussion forums discussed in Chinese documents appear more restricted in their content, implying more substantial government control over the areas in which big data may affect the citizen-government relationship.

Clearly, these findings must be interpreted within the diverse cultural and socio-economic contexts in which the document corpus was created. This has been only partially achieved, given the reliance on translated documents as well as due to space restrictions. Nonetheless, this study identified several avenues for future research that appear especially promising for unpacking the agency and politics of framing big data in both the BICs and the Global South more broadly. First, scholars should systematically analyze contextual factors, including traditional and social media discussions, civil society movements, external economic shocks, or even sporting mega events, to better understand how communicative context affects national strategies vis-à-vis big data and their implementation. Second, a closer look at how global corporations, such as Alphabet (Google), Alibaba, and international organizations such as the United Nations, intervene in big data strategies and programs in the Global South could help disentangle global and local visions and highlight restrictions on local actors' agency arising from dependencies, including the need for foreign aid or foreign direct investment. Finally, it is crucial to examine the temporal aspects of governments' engagement with big data, in particular how learning processes contribute to the narrative (and constant) reinvention of big data and the challenges and opportunities they provide within the Global South. Such research, will add additional nuance to global conversations about cross-cultural understandings of big data and provide a multinational, critical perspective on the promises and pitfalls of digital politics and innovation.

Notes

1. These documents were not analyzed, as they were finalized subsequent to the submission of this article.
2. Official versions of all Indian documents and all but one Chinese document were available in English. For the remaining documents, automatic translation tools (e.g., Google Translate), reputable crowd-source translation sites (e.g., China Law Translate, 2020), and deliberate and spot checks by native speakers were triangulated to ensure the accuracy of translations. All documents as well as the codebook are available at Mahrenbach, Mayer and Pfeffer, 2018b.
3. All three countries have only recently implemented privacy and data protection regulations. Future studies should evaluate the connections highlighted here anew once these regulations have become fully integrated within the BICs.

Websites

China Law Translate, <https://www.chinalawtranslate.com/en/>
 Digital India, <https://www.digitalindia.gov.in/>
 Google Translate, <https://translate.google.com>
 Participa.br, <http://participa.br/>

References

- Arora, Payal. (2016). The bottom of the data pyramid: Big data and the Global South. *International Journal of Communication*, 10(19) 1681–1699.
- Beer, David. (2016). How should we do the history of big data? *Big Data & Society*, 3(1) 1–10. doi:10.1177/2053951716646135
- Central Committee of the Communist Part of China. (2015). *The 13th five-year plan for economic and social development of the People's Republic of China*. Beijing, CN: Central Compilation & Translation Press.
- China Law Translate. (2020). [Website]. URL: <https://www.chinalawtranslate.com/en/> [January 8, 2020].
- China State Council. (2015a). *Notice of the state council on printing and distributing the action plan for promoting big data development*. Guo Fa [2015], No. 50. Beijing, CN: People's Daily Online.
- China State Council. (2015b). *Made in China 2025*. Beijing, CN: IoTOne, 1-38. Congresso Nacional. (2011). *Access to Information Law. C. Nacional*. Brasília, BR: Congresso Nacional.
- CIS. (2018). *Big data in governance in India: Case studies*. The Centre for Internet and Society, India.
- Congresso Nacional. (2014). *Internet Bill of Rights. C. Nacional*. Brasília, BR: Congresso Nacional.
- Daniels, Stephen, & Martin, Joanne. (1998). Punitive damages, change, and the politics of ideas: Defining public policy problems. *Wisconsin Law Review*, 1998(1), 71–100.
- Deephouse, David L., & Suchman, Mark. (2008). Legitimacy in organizational institutionalism. In R. Greenwood, C. Oliver, R. Suddaby, & K. Sahlin (Eds.), *The Sage handbook of organizational institutionalism* (pp. 49–77). Thousand Oaks, CA: Sage Publications. doi:10.4135/9781849200387.n2
- Doval, Pankaj. (2017, July 23). Need a law that ends data colonisation, says Nandan Nilekani. *Times of India*. URL: <https://timesofindia.indiatimes.com/india/citizens-should-have-ownership-of-digital-data-that-they-generate-nilekani/articleshow/59715877.cms> [August 31, 2018].
- European Commission. (2014). *Towards a thriving data-driven economy*. Brussels, BE: European Commission.
- European Commission. (2019). *Countries and regions*. Directorate-General for Trade. URL: <http://ec.europa.eu/trade/policy/countries-and-regions/> [January 31, 2019].
- Fearon, James D. (1994). Domestic political audiences and the escalation of international disputes. *The American Political Science Review*, 88(3), 577–592. doi:10.2307/2944796
- Financial Times. (2018, April 30). China and US compete to dominate big data. *Financial Times*. URL: <https://www.ft.com/content/e33a6994-447e-11e8-93cf-67ac3a6482fd> [July 27, 2018].
- Fredriksson, Cecilia, Mubarak, Farooq, Tuohimaa, Marja, & Zhan, Ming. (2017). Big data in the public sector: A systematic literature review. *Scandinavian Journal of Public Administration*, 21(3), 39–61.

- Gamson, William A., & Modigliani, Andre. (1989). Media discourse and public opinion on nuclear power: A constructionist approach. *American Journal of Sociology*, 95(1), 1–37. doi:10.1086/229213
- Government of India. (2012a). *National data sharing and accessibility policy*. New Delhi, IN: Ministry of Science & Technology.
- Government of India. (2012b). *National policy on information technology, 2012*. New Delhi, IN: Ministry of Electronics & Information Technology.
- Government of India (2015). *Principles of e-Kranti* (NeGP 2.0). New Delhi, IN: Ministry of Electronics & Information Technology.
- Government of India. (2018a). *All discussions under digital India*. URL: <https://www.mygov.in/home/13/discuss/> [September 20, 2018].
- Government of India. (2018b, September 19). *Digitize India platform*. URL: <https://digitizeindia.gov.in/> [September 20, 2018].
- Gronau, Jennifer, & Schmidtke, Henning. (2016). The quest for legitimacy in world politics – international institutions' legitimation strategies. *Review of International Studies*, 42(3), 535–557. doi:10.1017/S0260210515000492.
- Hsieh, Hsiu-Fang, & Shannon, Sarah E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288. doi:10.1177/1049732305276687
- Kitchin, Rob. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 1–12. doi:10.1177/2053951714528481
- Kshetri, Nir. (2014). The emerging role of big data in key development issues: Opportunities, challenges, and concerns. *Big Data & Society*, 1(2), 1–20. doi:10.1177/2053951714564227
- Laney, Doug. (2001, February 6). *3D data management: Controlling data volume, velocity and variety*. Stamford, application delivery strategies. File 949. Meta Group.
- Li, Keqiang. (2016). *Report on the work of the government*. Beijing, CN: T. S. C. o. t. P. s. R. o.
- Lucas, Louise. (2018, May 30). China emerges as Asia's surprise leader on data protection. *Financial Times*. URL: <https://www.ft.com> [August 31, 2018].
- Mahrenbach, Laura C., Mayer, Katja, & Pfeffer, Jürgen. (2018a). Policy visions of big data: Views from the Global South. *Third World Quarterly*, 39(10), 1861–1882. doi: 10.1080/01436597.2018.1509700
- Mahrenbach, Laura C., Mayer, Katja, & Pfeffer, Jürgen. (2018b). DATASET: Policy visions of big data: Views from the Global South. *Harvard Dataverse*, V1. doi: 10.7910/DVN/ZSTPBR.
- Manyika, James, Chui, Michael, Brown, Brad, Bughin, Jacques, Dobbs, Richard., Roxburgh, Charles, & Byers, Angela Hung. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute: McKinsey & Company.
- Ministro da Ciência, Tecnologia, Inovações e Comunicações. (2018). *Estratégia Brasileira para a transformação digital*. Departamento de Políticas e Programas Setoriais em TICs. Brasília, BR: Ministro da Ciência, Tecnologia, Inovações e Comunicações.
- Ministry of Planning and Budget. (2016). *Estratégia de Governança Digital da Administração Pública Federal*. Brasília, Brazil: Ministry of Planning and Budget, 1–34.
- Mossberger, Karen, Wu, Yonghong, & Crawford, Jared. (2013). Connecting citizens and local governments? Social media and interactivity in major U.S. cities. *Government Information Quarterly*, 30(4), 351–358. doi:10.1016/j.giq.2013.05.016
- Organisation for Economic Co-operation and Development. (2014). Data-driven innovation for growth and well-being. *New Sources of Growth: Knowledge-Based Capital*. Paris, FR : Organisation for Economic Co-operation and Development.
- Rieder, G. (2018). Tracing Big Data imaginaries through public policy: The case of the European Commission. In Ann Rudinow Saetnan (Ed.), *The politics of big data* (p. 89). Routledge.
- Rocha, Camilo. (2013, September 2). *Big data precisa de legislação e bom senso*. São Paulo, BR: O Estado de S.Paulo. URL: www.nexis.com [August 30, 2018].
- Rousseff, Dilma. (2013, September 24). *Statement by H.E. Dilma Rousseff*. United Nations General Assembly. New York, NY: United Nations.
- Ruan, Lotus. (2018). *When the winner takes it all: Big data in China and the battle for privacy*. Canberra, AU: Australian Strategic Policy Institute.

- Srikrishna, Bellur Narayanaswamy. (2017). *White paper of the committee of experts on a data protection framework for India*. New Delhi, IN: C. o. Experts.
- Srinivasan, Janaki, & Johri, Aditya. (2013). *Creating machine readable men: Legitimizing the 'Aadhaar' mega e-infrastructure project in India*. Sixth International Conference on Information and Communication Technologies and Development. Cape Town, South Africa.
- State Council Information Office. (2010). *The internet in China*. Beijing, CN: I. O. o. t. S. C. o. t. P. s. R. o.
- Taylor, Linnet, & Schroeder, Ralph. (2015). Is bigger better? The emergence of big data as a tool for international development policy. *GeoJournal*, 80(4), 503–518. doi:10.1007/s10708-014-9603-5
- Telecom Regulatory Authority of India. (2017). *Recommendations on net neutrality*. New Delhi, IN: T. R. A. o. I.
- United Nations Economic Commission for Africa, United Nations Development Programme, International Development Research Centre, and World Wide Web Foundation. (2017). *Africa data revolution report 2016*. Addis Ababa, UN Economic Commission for Africa, UN Development Programme, International Development Research Centre of Canada. Geneva, CH: World Wide Web Foundation.
- World Bank. (2018). *World Bank open data*. URL: <https://data.worldbank.org/> [August 30, 2018].
- World Trade Organization. (2019). *Statistics database*. Geneva, CH: World Trade Organization.
- Zeng, Jinghan. (2016). China's date with big data: Will it strengthen or threaten authoritarian rule? *International Affairs*, 92(6), 1443–1462. doi:10.1111/1468-2346.12750
- Zeng, Jinghan, Stevens, Tim, & Chen, Yaru. (2017). China's solution to global cyber governance: Unpacking the domestic discourse of "internet sovereignty." *Politics & Policy*, 45(3), 432–464. doi:10.1111/polp.12202
- Zheng, Lei. (2013). Social media in Chinese government: Drivers, challenges and capabilities. *Government Information Quarterly*, 30(4), 369–376. doi:10.1016/j.giq.2013.05.017

Appendix 1: Overview of document corpus

Country	Document	Year	Description
Brazil	Statement by H. E. Dilma Rousseff	2013	Speech by the Brazilian president at the opening of the 68th Session of the United Nations General Assembly
Brazil	Estratégia de Governança Digital da Administração Pública Federal (Digital Governance Strategy)	2016	Government strategy/roadmap with recommendations, terms and goals for e-government services and solutions
Brazil	Marco Civil da Internet, Law N° 12.965	2014	Legal framework governing the use of the Internet
Brazil	Decreto N° 8.771 (Internet Decree)	2016	Legal framework governing the use of the Internet
Brazil	Estratégia Brasileira para a Transformação Digital (Strategy for Digital Transformation)	2018	Government strategy for digitalization
China	Made in China 2025	2015	Government strategy detailing a ten-year action plan to support high-tech leadership and modernise production
China	Action Outline for the Promotion and Development of Big Data	2015	Government strategy to promote big data across all sectors via the "Platform for the Promotion of Big Data Development (No. 50 [2015])"
China	Report on the Work of the Government	2016	Report of the National People's Congress (NPC) on the government's work in the last year
China	The 13th Five-Year Plan for Economic and Social Development of the People's Republic of China (2016–2020)	2016	Government policy strategy for the coming five years
India	EKranti – National E-Governance Plan	2015	Government strategy for e-governance and its implementation
India	National Data Sharing and Accessibility Policy	2012	Policy regulating access to government data
India	Big Data Management Policy	2016	Policy for use of big data by the Indian Audit & Accounts Department
India	Big Data Initiative	2016	Homepage of the Indian Department of Science & Technology describing the government's Big Data Initiative