COGITATIO

Politics and Governance (ISSN: 2183–2463)
2018, Volume 6, Issue 4, Pages 1–4
DOI: 10.17645/pag.v6i4.1810

Editorial

Big Data Applications in Governance and Policy

Sarah Giest [1,*] and Reuben Ng [2,3]

[1] Institute of Public Administration, Leiden University, 2511 DP The Hague, The Netherlands;
E-Mail: s.n.giest@fgga.leidenuniv.nl
[2] Lee Kuan Yew School of Public Policy, National University of Singapore, 259772 Singapore, Singapore;
E-Mail: spprng@nus.edu.sg
[3] Geriatric Education and Research Institute, 769027 Singapore, Singapore

Submitted: 5 November 2018 | Published: 21 November 2018

Abstract
The editorial sets the scene for this thematic issue on big data applications in governance and policy. It highlights the lack of engagement in the current literature with the application of big data at the cross-section of governance of data and its utilization in the policy process and draws out aspects related to its definition and future research agenda. The contributions highlight several aspects related to big data in different contexts, such as local and national government as well as a variety of policy areas. They converge on the idea that big data applications cannot overcome existing political and structural limitations that exist in government. This leads to a future research agenda that looks at the disconnect between data production and usage as well as identifying policy issues that are more or less suitable for data analytics.

Keywords
big data; governance; policy analytics; policymaking; politics

Issue
This editorial is part of the issue "Big Data Applications in Governance and Policy", edited by Sarah Giest (Leiden University, The Netherlands) and Reuben Ng (National University of Singapore, Singapore).

1. Introduction

Recent literature has been trying to grasp the extent as to which big data applications affect the governance and policymaking of countries and regions (Boyd & Crawford, 2012; Giest, 2017; Höchtl, Parycek, & Schöllhammer, 2015; Poel, Meyer, & Schroeder, 2018). The discussion includes the comparison to e-government and evidence-based policymaking developments that existed long before the idea of big data entered the policy realm. The theoretical extent of this discussion however lacks some of the more practical consequences that come with the active use of data-driven applications. In fact, much of the work focuses on the input-side of policymaking, looking at which data and technology enters the policy process, however very little is dedicated to the output side. In short, how has big data shaped data governance and policymaking? The contributions to this thematic issue shed light on this question by looking at a range of factors, such as campaigning in the US election (Trish, 2018) or local government data projects (Durrant, Barnett, & Rempel, 2018). The goal is to unpack the mixture of big data applications and existing policy processes in order to understand whether these new tools and applications enhance or hinder policymaking.

Existing research is split regarding the usefulness of big data in the policy realm. Some are convinced that there is nothing new in the way data is being used—even if it is *big* data. This argument is in reference to the large administrative datasets that government has handled prior to the big data idea and the technological shift that came with the introduction of computers and increasingly refined software to utilize data (Connelly, Playford, Gayle, & Dibben, 2016). Others however see a shift at the scale of the Industrial Revolution (Richards & King, 2014), due to the type and speed of information being available. Since there is a variety of big data applications and governance systems, it is difficult to find

one answer to the question whether big data will permanently alter the policymaking process. With this thematic issue we aim to contribute to this discussion by highlighting applications in a variety of contexts to show that they come to a common conclusion: there is benefit to using big data in the policy realm, however (1) a more nuanced look at ongoing applications reveals a complex picture of politics entering the process, and (2) contextual factors, such as the level of government, the policy field and the hierarchical structure affect data utilization. In other words, big data applications cannot overcome existing political and structural limitations that exist in government. This finding might be a less exciting one, but is a cautionary warning to those governments that portray big data as numbers-only, neutral information that can solve a variety of issues.

The following section gives an overview over the definition of big data in the governance and policymaking literature and is followed by a summary of the contributions to this thematic issue. The editorial concludes with ideas for future research.

2. Big Data in Governance and Policy

The concept of big data is vague and has not been clearly defined (Connelly et al., 2016). The articles in this issue converge on a definition that acknowledges the different forms in which big data can appear in the policy process. For example, Durrant et al. (2018) consider administrative data as a form of big data, because it is exhaustive, highly granular, large and found and repurposed, rather than made. Trish (2018) also focuses on the use of administrative and performance data as part of a long-standing evidence-based policy movement in the US government. Longo and Dobell (2018) acknowledge census data as big data, and focus in their paper on its velocity and variety as a foundation for policy analytics. Ng (2018) defines big data as unstructured data that a city produces such as video, audio, sensor data, citizens' conversation online and social media. This zooms in on the volume and veracity of the data available.

There are two ways of understanding the use of big data in government. One is to look at the governance of big data, which includes the handling and regulation of data. The other perspective is to focus on the utilization of big data for specific policy problems. In this issue we collapse both into the idea of big data in governance and policy-making based on the assumption that they are intrinsically linked. This linkage is visible when data regulations prevent the collaboration of government units for addressing cross-cutting issues (Durrant et al., 2018). Another intersection of big data for policymaking and the governance of it are the challenges highlighted by Trish (2018) around public scrutiny of the information being used by government. Here, questions are raised around how the data is governed in terms of its transparency and values and as well as how this information is used to make decisions around public policy.

3. Contributions to This Thematic Issue

Longo and Dobell (2018) begin the thematic issue with an overview of theoretical and applied work in policy analytics. They define policy analytics as a modification to the traditional policy analysis approach and position this idea in a wide variety of literature while giving practical examples of its application. By looking at the emergence of policy analytics within the policy sciences, they find that new ways to analyze policies is much more than just data analysis. Based on a review of recent literature, they show that the promises that data-driven applications make is met with the complexity of policy decisions. This intersection is where less researched, but interesting questions are raised in terms of whether the policy environment is too complex for even advanced policy analytics to contribute or whether the effects of one policy decision is so diffused in a variety of sectors and governmental levels that the effect of policy analytics is hard to grasp. Longo and Dobell (2018) conclude with a matrix for the applicability of policy analytics across scale (from local to global) and complexity (from uncertain to certain). This illustration shows that policy analytics can best support local problems that have a degree of certainty, such as monitoring, implementation and enforcement.

In their contribution, Durrant et al. (2018) pick up on the idea of local big data applications and use participatory action research to observe activities of identifying, integrating and analyzing multiple and diverse forms of data. Based on this, they theorize about the social contexts of both data production and policymaking to better understand the boundaries and barriers to big data for policy. In their work, Durrant et al. (2018) find that the context for data applications is deeply value-laden and political, which leads them to draw the following conclusions. First, there is an absence of data sufficiently relevant for addressing specific policy questions. In other words, the questions being raised in, for example health policy, could not be answered with the data being routinely collected and made available to the responsible agency. Second, the data being collected is largely used for service administration and audit rather than tackling underlying issues, such as reasons for low service take-up. Finally, the cost of providing data is greater than the perceived benefit. This has to do with having to establish the validity of data access requests by different authorities, data holders and project teams being involved. Taken together, Durrant et al. (2018) conclude that the insights from available data are not always actionable in the local context due to the factors mentioned above and that caution should be exercised when it comes to which questions can be asked of big data.

Trish (2018) focuses on the data use at national level in the US. She analyses three cases from the Obama Administration: microtargeting in electoral campaigns, performance management in government and signature drone strikes. Whereas these applications are highly technical in nature, the paper shows that, similar to

the previous two analyses, underlying assumptions and power relationships impact the usefulness of data. In fact, decisions are often made based on incomplete information and Trish (2018) cautions the uncritical use of data by having efficiency as a foundation for such decisions. She finds that there is limited public scrutiny in combination with an undercurrent of market-based influence. Looking ahead, Trish (2018) concludes that using big data in this way reinforces existing biases of society and gives decisions an appearance of objectivity, which is not warranted based on the type of data that is being used. With this assessment, she underpins the previous findings that big data has a role to play, but the information drawn from the data has to be used with caution in terms of their completeness, applicability and the type of question they are supposedly answering.

Finally, Ng (2018) provides a case study of Singapore's big data applications in governance and policy that are enabled by cloud computing adoption. He distills five key factors that drive the use of big data in public management and policy: (1) public demand for big data applications; (2) focus on whole-of-government policies and practices; (3) restructuring of technology agencies to integrate strategy and implementation; (4) creating the Smart Nation Platform; (5) purpose-driven big data applications especially in healthcare. Taking lessons from Singapore, he concludes that other countries can promote regulatory sandboxes to experiment with policies that proactively manage novel technologies and business models that may radically change society, and establish more public-private partnerships to co-innovate on challenges.

4. Concluding Remarks and Ideas for Future Research

This thematic issue raises a non-exhaustive list of issues linked to big data in governance and policy. The contributions shed light on a range of factors that have been partially overlooked in current research on the topic. In particular, all papers converge on the idea that policymaking is a complex process in which data analytics is one factor that might have positive, negative or no effect at all. In fact, the papers highlight that the positive effect is over-valued, which leads to decisions being made based on incomplete evidence (Trish, 2018) or irrelevant information regarding the problem at hand (Durrant et al., 2018). The contributions further give the sense that the production and use of data remain two separate processes, which means that the data are not answering the questions linked to specific policy issues. This disconnect leads to data-based evidence that is incomplete, not actionable or even confusing from the perspective of policymakers looking for answers. Hence, a future research question would be how this disconnect comes about and how policy issues can inspire data collection rather than existing data informing solutions for policy problems.

Another issue raised in the contributions for this thematic issue, is the complexity of policymaking, which cannot be simplified by more data. In fact, data has not shown to be as disruptive to existing processes due to long-standing political, hierarchical and procedural structures. As Longo and Dobell (2018) point out, the context in which big data tools are applied matters in terms of its complexity and scale. Looking ahead it raises the question for big data and policy research, whether data use is particularly applicable to activities, such as monitoring and unfit for more complex issues, such as community health services (Durrant et al., 2018). Essentially, data have to achieve a purpose. As Ng (2018) concludes, a data project without clear policy goals careens into disillusionment, and negatively impacts the perception of data in the policy process.

Finally, the contributions agree that data use in policymaking is not a linear process where data is analyzed and then information fed into the policy cycle. In fact, barriers to data use occur in unlikely situations, such as the sharing of data with private companies who then deny access to it for integration (Durrant et al., 2018) or the need for qualitative statements next to predictive models due to unexpected outcomes (Longo & Dobell, 2018). This points towards questions of trust in the process of sharing and using certain type of data. Looking ahead, research institutions, such as universities, could play a unique role by bringing together public and private organizations to achieve mutually beneficial outcomes. Ng (2018) describes such a formalized approach where a Memorandum of Understanding (MOU) is signed among multiple private and public entities to co-create solutions for complex societal challenges.

Acknowledgements

The editors would like to acknowledge and thank the contributing authors and the anonymous reviewers of this thematic issue for their work.

Conflict of Interests

The authors declare no conflict of interest.

References

Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, *15*(5), 662–679.

Connelly, R., Playford, C. J., Gayle, V., & Dibben, C. (2016). The role of administrative data in the big data revolution in social science research. *Social Science Research*, *59*(C), 1–12.

Durrant, H., Barnett, J., & Rempel, E. S. (2018). Realising the benefits of integrated data for local policymaking: Rhetoric versus reality. *Politics and Governance*, *6*(4), 18–28.

Giest, S. (2017). Big data for policymaking: Fad or fasttrack? *Policy Sciences*, *50*(3), 367–382.

Höchtl, J., Parycek, P., & Schöllhammer, R. (2015). Big

data in the policy cycle: Policy decision making in the digital era. *Journal of Organizational Computing and Electronic Commerce*, *26*(1/2), 147–578.

Longo, J., & Dobell, R. (2018). The limits of policy analytics: Early examples and the emerging boundary of possibilities. *Politics and Governance*, *6*(4), 5–17.

Ng, R. (2018). Cloud computing in Singapore: Key drivers and recommendations for a smart nation. *Politics and Governance*, *6*(4), 39–47.

Poel, M., Meyer, E. T., & Schroeder, R. (2018). Big data for policymaking: Great expectations, but with limited progress? *Policy and Internet*, *10*(3), 347–367.

Richards, N.bM., & King, J. H. (2014). Big data ethics. *Wake Forest Law Review*, *49*, 393–432.

Trish, B. (2018). Big data under Obama and Trump: The data-fueled U.S. presidency. *Politics and Governance*, *6*(4), 29–38.

About the Authors

Sarah Giest is an Assistant Professor at the Institute of Public Administration, Leiden University. She is conducting research on urban policymaking, specialising in networking among various governmental and non-governmental stakeholders and the policy design connected to that. Sarah is currently working on several papers and projects connected to big data use by policymakers together with researchers from Computer Science with the goal of unpacking the way data information enters the policy process in the city context.

Reuben Ng spent 16 years in government, consulting, and research. In government, he was in the Prime Minister's Office of Singapore driving evidence-based policy-making through data analytics, and Smart Nation strategies. In consulting, he co-built the Advanced Analytics practice at a top firm and implemented complex analytics projects across industries and functions. In research, he is an expert in quantitative social sciences, social gerontology, and credited with creating innovative techniques to measure societal perceptions/stereotypes that are applied to policy, and program evaluation. Reuben trained as a behavioural scientist at NUS, Oxford and Yale where he was Singapore's first Fulbright Science and Technology Scholar.