SITUATING **OPEN DATA**

GLOBAL TRENDS IN LOCAL CONTEXTS

Edited by Danny Lämmerhirt, Ana Brandusescu, Natalia Domagala & Patrick Enaholo

Situating Open Data

Global Trends in Local Contexts

Edited by
Danny Lämmerhirt, Ana Brandusescu,
Natalia Domagala & Patrick Enaholo

AFRICAN
MINDS

Published in 2020 by African Minds
4 Eccleston Place, Somerset West 7130, Cape Town, South Africa
info@africanminds.org.za
www.africanminds.org.za

This work is published under a Creative Commons Attribution 4.0 International License (CC-BY).

ISBN Paper 978-1-928502-12-8
ISBN eBook 978-1-928502-13-5
ISBN ePub 978-1-928502-14-2

Orders:
African Minds
4 Eccleston Place, Somerset West 7130, Cape Town, South Africa
info@africanminds.org.za
www.africanminds.org.za

For orders from outside South Africa:
African Books Collective
PO Box 721, Oxford OX1 9EN, UK
orders@africanbookscollective.com
www.africanbookscollective.com

Introduction

The Open Data Research Symposium (ODRS) is a bi-annual gathering designed to provide a space for researchers working on open data to reflect critically on their findings and to apply and advance theories that explain the dynamics of open data as a socially constructed phenomenon and practice. ODRS is intended to be a safe space for debate in the presence of demands for fast results and proofs of impact – although this is not to suggest that researchers should be immune to considerations of relevance and transfer. Therefore, ODRS is usually organised alongside the International Open Data Conference (IODC) as a way for researchers to present their latest work, learn about other projects in the open data space and brainstorm new ideas. The symposium inspired the emergence of the Open Data Research Network with an aim to keep the momentum going in the ODRS community by sharing ongoing research.

For the 2018 edition of the symposium, the organisers received a total of 30 extended abstracts of which 12 were accepted for presentation in Buenos Aires. Selection was based on single-blind review of each abstract by at least two independent experts. Authors were required to submit full papers of their abstracts one week ahead of the symposium, and these papers were shared with those who had registered for the symposium. After the symposium, accepted authors were invited to revise and submit full papers for consideration in this edited volume. Papers underwent double-blind review by at least two peers, and authors were required to revise their papers before being accepted for publication.

Situating open data in critical data studies

The ODRS 2018 itself explored how open government data research should position itself within broader debates on data. Some attendees were in favour of including other types of open data beyond government data. Attendees also reflected on the fact that much open data research focuses narrowly on data from public institutions and governments, excluding research on open access and open science for instance. Others suggested the need to revive engagement with topics such as data ownership, the politics of making data public, or the relationship between open data, data rights and data justice. The role of open data for corporate actors, automated systems, artificial intelligence, and other emerging trends was also highlighted.

This suggests shifting viewpoints on open data initiatives from what they were some years ago. Instead of deploying and implementing open data programmes in social contexts, or seeking to understand their benefits and impacts, the value

proposition behind these programmes should be interrogated, in order to zoom into the politics of the data. Data politics examines the influence data has on how humans relate to one another, who benefits from these relations, how value is created, and for whom (Ruppert et al. 2017). It offers an alternative to framing data as a resource to be unlocked, pooled, tamed, or unleashed, since it considers data as an essential element that is embedded in histories, social orders, situated ways of knowing, and a wider political economy.

This book provides empirical accounts to understand the situatedness of open data along the following themes: 1) open data practices; 2) the local implementation of global trends; and 3) open data ecosystems. Many chapters in this volume simultaneously address several of these themes. The thematic grouping of chapters is an attempt to foreground salient questions for open data research.

In addition, the book covers country-specific, localised applications of open data with a few chapters explicitly focusing on how open government data initiatives unfold within different socio-political contexts. The geographical scope of the contributions spans four continents, providing insights on open data practices in Europe (Kosovo, Belgium, United Kingdom), Africa (Nigeria, Tanzania), Asia (Indonesia, the Philippines), and Latin America (Paraguay, Brazil).

Open data practices

Several chapters discuss how actors practice open data. Each of them points – more or less explicitly – to the dual role data literacy plays. It may at once be a resource to upskill people and work with data and a form of critical practice that reflects on the historical functions of data within public institutions and the public sphere. Several chapters explicitly make a case for boosting data literacy as a matter of upskilling and point out resource and other constraints hindering the take up of data literacy.

Perhaps less explicitly, but not less importantly, several chapters critically examine how people practise data in different social settings – from health data reporting to data journalism – to show how data practice and the functioning of institutions are mutually related. Regarded in this way, data literacy becomes a form of critique that understands data not only as a product of institutions but as an active element shaping how institutions administer the world.

In Chapter 1: ‘What technology and open data can do for women in Kosovo: A critical assessment of the potential of ICT skills programmes and open data to empower women in the ICT sector in Kosovo’, Natalia Domagala critically assesses how open data is used in ICT grassroots programmes created for young women to gain ICT skills and work with open data, and whether this contributes to their empowerment. Domagala answers pertinent questions such as: What are the sources of disempowerment of women in the ICT sector in Kosovo? What are the strengths of ICT skills programmes and open data for the empowerment of women? What are their limitations? Through key informant interviews,

INTRODUCTION

Domagala identifies that these programmes can counter the negative effect of socio-cultural norms, which end up discouraging women from undertaking employment in the ICT sector. These programmes do so by providing highly demanded skills, support networks and role models. Yet, ICT skills building programmes can exclude women who lack access to the internet, provide little practical employment assistance and fall short on strategic ways to strengthen political engagement even though they are often touted as championing civic engagement. The chapter provides practical insights for policymakers on how to improve the effectiveness of ICT skills programmes and their ability to create broader social benefits by using open data.

In Chapter 2: 'Journalists and the intermediation of open data: A Nigerian perspective', Patrick Enaholo and Doyinsola Dina examine the work of open data intermediation by journalists in Nigeria. The chapter discusses how journalists, as intermediaries of government data, engage with end-users in the open data ecosystem. The chapter builds on the competency framework proposed in Van Schalkwyk et al. (2016) to understand the role of journalists as intermediaries of open data. Using a mixed method approach, the authors determine that, within the specific context they operate, technical as well as creative competencies are key to ensuring the effectiveness of journalists' role in the open data ecosystem. Although the study also identifies the lack of technical skills and competencies as the primary challenge for journalists in the Nigerian context, they show how these are compensated for by other intermediaries in the open data value chain such as civil society and civic tech groups.

In Chapter 3: 'Using open data for public services', Miranda Marcus, Ed Parkes, Therese Karger-Lerchl, Jack Hardinges and Roza Vasileva explore the use of open data in public service delivery in a collaborative manner to solve problems and support innovation in the UK public sector. The chapter distinguishes three patterns of open data in public sector institutions and their effects on delivering public services: (1) open data for increasing access to public services; (2) open data for more efficient public service delivery; and (3) open data for policy development. The research has been conducted through a combination of interviews, desk research, and visualisation techniques. Due to the complexity and non-linear process of public service delivery, the authors used a rich picturing method to portray public services as ecosystems. The chapter concludes with comparisons across the three patterns of open data use in public services, looking at the following aspects: organisational collaboration, technology infrastructure, digital skills and literacy, foundational data infrastructure, open standards for data, senior-level championing, intermediaries, problem focus, open innovation and peer networks.

Local implementation of global trends

In Chapter 4: 'Localising global commitments: Open data in sub-national contexts in Indonesia and the Philippines', Michael Cañares analyses the

role and scope of local government in implementing the Open Government Partnership commitments. The chapter develops case studies of early open data adopters in Bohol province in the Philippines and the city of Banda Aceh in Indonesia. The cases describe how local open data initiatives are shaped by political leadership, civil society, and data literacy. It then outlines challenges and opportunities for implementing open data initiatives. The chapter concludes with recommendations for open data practitioners in local governments to: 1) engage in local consultations; 2) ensure a legal basis for open data initiatives; 3) strengthen citizen engagement; 4) and invest in local capacity for data suppliers and users.

In Chapter 5: ‘Closing the gaps in open data implementation at sub-national government level in Indonesia’, Ilham Cendekia Srimarga and Markus Christian trace the implementation of an open data project in a sub-national context. The authors discuss their involvement in the transition of Bojonegoro district government in Indonesia from one stage of aggregated data to one in which disaggregated data is employed in order to facilitate higher levels of public participation with open data. The challenges encountered during the first stage included the obsolescence of the data and the difficulty of processing due to its closed format which fell short of expectations and led to a predominantly low response from citizens. The second stage involved civil society organisations and other stakeholders to provide analyses and recommendations that eventually engendered a ‘data revolution’ concept. The authors examine how data provisioning technology (such as the format and representation of open data) and the context in which this happens (such as the low rate of adoption of open data by stakeholders) can shape the eventual outcomes of such implementations.

In Chapter 6: ‘The cost of late payments in public procurement’, Juan Pane, Camila Salazar, and Julio Paciello examine how granular open data on public procurement payment processes can be used to contribute to cost savings for public administration. Based on public institutional data from Paraguay, the authors assess the duration of public procurement payments to estimate the cost of payment process delays and how they affect suppliers. In turn, the authors determined the variables that contribute to these delays using descriptive analysis and financial cost estimation to calculate the cost of late payments on suppliers, and modelled the duration of payments using survival analysis to identify which variables have a role in delaying payments. When properly implemented and analysed, the study found that open data can positively impact the public procurement cost saving process. The study finds that the duration from the moment an invoice is issued to when the payment occurs can be of approximately 55 days on average for each payment as opposed to international practice of a 30-day average. Between 2011 and 2017, late payments accumulated to USD 81.07 million. For cost saving efforts to be successful, the authors recommend the Paraguay public administration to analyse the payment process and apply an appropriate corrective normative framework.

INTRODUCTION

In Chapter 10: 'From theory to practice: Open government data, accountability and service delivery', Michael Jelenic interrogates the extent to which the use of open government data is associated with higher levels of accountability and improved service delivery in developing countries. Constructing a unique dataset that operationalises open government data, government accountability, service delivery, as well as other intervening and control variables, the research adopts secondary data from 25 countries in sub-Saharan Africa to find a number of significant associations between open government data, accountability, and service delivery. However, the findings suggest differentiated effects of open government data across the health and education sectors, as well as with respect to service provision and service delivery outcomes. Although this early research has limitations and does not attempt to establish a purely causal relationship between the variables, it provides initial empirical support for claims about the efficacy of open government data for improving accountability and service delivery.

Open data ecosystems

Several chapters attend to the constituent parts of open data ecosystems to understand how these can be infrastructured for broader socio-economic benefits. The chapters reflect a recent academic debate among platform studies and infrastructure studies scholars who argue that there is a hybridisation of concepts usually considered apart from another. Platforms that are usually considered as programmable, modular systems become infrastructural, meaning engrained in daily lives and serving critical functions. Infrastructures which are usually considered as learned practice, critical for daily lives, and visible only during breakdown become increasingly 'platformised', meaning splintered across service providers (Plantin et al. 2016).

François van Schalkwyk's contribution to this book adds to this discussion, reflecting on how an open data platform could shape data flows in existing information infrastructures. In Chapter 7: 'Connecting flows and places: Flows of (open) data to, from and within hyperlocal communities in Tanzania', he discusses how health-related data travels in a data ecosystem between public institutions and civic actors. Building on theories of data ecosystems, the chapter maps out how infrastructural elements shape 'blockages, breaks, switchers and connectors of data flows'. These infrastructural elements include databases and how they connect to data exchange, reporting procedures to digitise and share data, as well as trained social conventions and (dis-)incentives to report and share data across local, regional, and national scales. Using the method of tracing data flows, the chapter draws attention to how data flows can have centralising tendencies, with many reporting routines steered towards higher-level administrations, instead of enabling data flows towards civic communities. The chapter provides fresh thinking not only on how open data platforms can integrate in information systems, but also how data flows are shaped by

organisational interests, and how they create inherent conflicting interests, exclusion, as well as power dynamics.

Other contributions approach open data ecosystems from the viewpoint of platform architectures to understand how the design of platform elements (such as APIs) interact with the performance of application ecosystems and user satisfaction. In Chapter 8, Julián Rojas, Bert Marcelis, Eveline Vlassenroot, Pieter Colpaert, Matthias van Compernelle and Ruben Verborgh discuss ‘Decentralised open data publishing for the public transport route planning ecosystem’. The chapter studies the implementation of a decentralised open data publishing strategy in the public transport route planning ecosystem. Comparing two data architectures – REST APIs and decentralised linked data – the chapter discusses how each enables users to withdraw, compile, and calculate different types of data. The chapter provides insights on how platform elements influence the performance of network components, including apps and mobile phones. In addition, the chapter provokes thinking on how the platform architecture relates to privacy-by-design, technical performance, and user-perceived satisfaction with traffic routing apps.

In Chapter 9: ‘Building a framework for analysis of factors to creation and growth of an open data ecosystem’, Edson Carlos Germano, Nicolau Reinhard and Violeta Sun propose a framework for analysing and exploring the factors that can influence the economic and social sustainability of Open Government Data (OGD) ecosystems comprising producers, intermediaries, and users who maintain relationships with each other by creating a flow of information, services, products and financial resources. The authors set out to examine the information flow throughout the OGD value chain with a focus on data reuse in emerging countries. Based on an in-depth literature review, they propose five characteristics for the analysis and understanding of OGD ecosystems. The proposed framework considers rules and governance aspects defining the relations between ecosystems’ data producers and final users of the products and services, focusing on informal membership rules in the ecosystem. The authors argue that with such a framework – which is currently being adopted for the Brazilian federal government OGD ecosystem – it will be possible to analyse and understand the governance necessary for the construction and sustainability of OGD ecosystems. This, to the authors, can help society obtain greater benefit from using OGD analysis to understand how government resources are used more efficiently.

We hope that this book sparks a critical reflection on the way open data is practised, implemented locally, and its strategic role in ecosystems. The volume departs from framings of data as a resource, and on impact as an end-goal. Rather, the chapters focus on the process and how open data influences social relations, shapes histories and power structures affecting information systems.

The editors
October 2019

INTRODUCTION

REFERENCES

- Plantin J-C, Lagoze C, Edwards PN & Sandvig C (2016) Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society* 20(1): 293–310. <https://doi.org/10.1177/1461444816661553>
- Ruppert E, Isin E & Bigo D (2017) Data politics. *Big Data & Society* 4(2): 1–7. <https://doi.org/10.1177/2053951717717749>
- Van Schalkwyk F, Cañares M, Chattapadhyay S & Andreason A (2016) Open data intermediaries in developing countries. *Journal of Community Informatics* 12(2): 9–25. <http://ci-journal.net/index.php/ciej/article/view/1146>