



King's Research Portal

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Gray, J., & Lammerhirt, D. (2019). Making Data Public? The Open Data Index as Participatory Device. In A. Daly, K. Devitt, & M. Mann (Eds.), *Good Data* Institute of Network Cultures. http://networkcultures.org/wp-content/uploads/2019/01/Good_Data.pdf

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

GOOD DATA

EDITED BY
ANGELA DALY,
S. KATE DEVITT
& MONIQUE MANN

**THEORY
ON
DEMAND**

A SERIES OF READERS
PUBLISHED BY THE
INSTITUTE OF NETWORK CULTURES
ISSUE NO.:

29

GOOD DATA

Edited by Angela Daly,

S. Kate Devitt

and Monique Mann

Theory on Demand #29

Good Data

Editors: Angela Daly, S. Kate Devitt and Monique Mann

Copy editor: Harley Williams

Editorial assistant: Kayleigh Murphy

Funding: Queensland University of Technology Faculty of Law

Cover design: Katja van Stiphout

Design and EPUB development: Barbara Dubbeldam

Published by the Institute of Network Cultures, Amsterdam, 2019

ISBN 978-94-92302-27-4

Contact

Institute of Network Cultures

Phone: +3120 5951865

Email: info@networkcultures.org

Web: <http://www.networkcultures.org>

This publication is available through various print on demand services and freely downloadable from <http://networkcultures.org/publications>

This publication is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-SA 4.0).



CONTENTS

1: Introduction: What is (in) Good Data? **8**

Monique Mann, Kate Devitt and Angela Daly

THEME 1: GOOD DATA MANIFESTOS AND PRACTICES

2: Good Data Practices for Indigenous Data Sovereignty and Governance **26**

Raymond Lovett, Vanessa Lee, Tahu Kukutai, Donna Cormack, Stephanie Carroll Rainie and Jennifer Walker

3: The Good Data Manifesto **37**

Claire Trenham and Adam Steer

4: The Good, the Bad and the Beauty of ‘Good Enough Data’ **54**

Miren Gutiérrez

5: An Energy Data Manifesto **77**

Declan Kuch, Naomi Stringer, Luke Marshall, Sharon Young, Mike Roberts, Iain MacGill, Anna Bruce and Rob Passey

THEME 2: GOOD DATA AND JUSTICE

6: Trade-offs in Algorithmic Risk Assessment: An Australian Domestic Violence Case Study **96**

Daniel McNamara, Timothy Graham, Ellen Broad and Cheng Soon Ong

7: Unlawful Data Access and Abuse of Metadata for Mass Persecution of Dissidents in Turkey: the ByLock Case **117**

A. Sefa Ozalp

8: Not as Good as Gold? Genomics, Data and Dignity **135**

Bruce Baer Arnold and Wendy Elizabeth Bonython

9: Data Localization: Policymaking or Gambling? **156**

Nikita Melashchenko

THEME 3: GOOD DATA AS OPEN AND SHARED DATA

10: Making Data Public? The Open Data Index as Participatory Device **174**

Jonathan Gray and Danny Lämmerhirt

11: Data Journalism and the Ethics of Open Source **189**

Colin Porlezza

12: Governance of Communal Data Sharing **202**

Chih-Hsing Ho and Tyng-Ruey Chuang

THEME 4: GOOD DATA ACTIVISM AND RESEARCH

13: Provocations for Social Media Research: Toward Good Data Ethics **216**

Andrea Zeffiro

14: Data for the Social Good: Toward a Data-Activist Research Agenda **244**

Becky Kazansky, Guillén Torres, Lonneke van der Velden, Kersti Wissenbach and Stefania Milan

15: Good Data is Critical Data: An Appeal for Critical Digital Studies **260**

Chiara Poletti and Daniel Gray

16: The Fieldnotes Plugin: Making Network Visualisation in Gephi Accountable **277**

Maranke Wieringa, Daniela van Geenen, Karin van Es and Jelmer van Nuss

THEME 5: GOOD DATA AND SMART CITIES AND HOMES

17: Algorithmic Mapmaking in ‘Smart Cities’: Data Protection Impact Assessments as a means of Protection for Groups **298**

Gerard Jan Ritsema van Eck

18: Truly Smart Cities. Buen Conocer, Digital Activism and Urban Agroecology in Colombia **317**

Juan-Carlos Valencia and Paula Restrepo

19: Intelligent Warning Systems: ‘Technological Nudges’ to Enhance User Control of IoT Data Collection, Storage and Use **330**

Rachelle Bosua, Karin Clark, Megan Richardson and Jeb Webb

20: Domesticating Data: Socio-Legal Perspectives on Smart Homes and Good Data Design **344**

Martin Flintham, Murray Goulden, Dominic Price and Lachlan Urquhart

Bibliographies **361**

10: MAKING DATA PUBLIC? THE OPEN DATA INDEX AS PARTICIPATORY DEVICE

JONATHAN GRAY AND DANNY LÄMMERHIRT

Introduction

The Open Data Index is a 'civil society audit' which strives to shape the availability and openness of public sector data from around the world. In this chapter we examine the social life of this project, including how it evolved, the changing visions and practices associated with it, and how it serves to assemble and involve different publics in the assessment of institutional practices and forms of datafication. Drawing on recent work on stactivism, data activism and the performative effects of numbers, rankings and indices, we look at how the index organises participation and data politics in specific ways, raising questions about not only making data public but also the making of public data. It plays two roles which are sometimes in tension: (i) conventionalising assessment to facilitate comparability, and (ii) reflecting the diversity of different interests, issues and settings involved in opening up public sector data. It also facilitates the creation of 'enumerated entities' as objects of concern in open data advocacy and policy. The Open Data Index may thus be viewed as a site where participation is both configured and contested, and where practices of valuation and enumeration are both conventionalised and brought into question.

How can various publics participate in shaping what data is created, collected and distributed by states? How might public participation around the availability and openness of public data lead to 'good data' (or at least 'better data')? In this chapter we reflect on the social life of the Open Data Index, a civil society project which aims to measure and influence how government data is made available. In particular we attend to how the index organises participation and data politics in specific ways, raising questions about not only making data public but also the making of public data.

There are many ways that one might appraise such an initiative. From the perspective of what Bruno Latour calls an 'anthropology of modernity',¹ there is a lot to be unpacked in these three little words: 'open', 'data' and 'index'. For example, one might consider imaginaries and practices of the 'index' in light of research on cultures of auditing and accountability.² One might tell a tale of the rise of the index as a style of political mobilisation alongside the rally, the petition or the hashtag. As well as public sector indices such as the United Nations 'Human Development Index' and 'Human Poverty Index', there are now a wide variety of

1 Bruno Latour, *An Inquiry into Modes of Existence: An Anthropology of the Moderns*, trans C. Porter, Cambridge MA: Harvard University Press, 2013.

2 Michael Power, *The Audit Society: Rituals of Verification*, Oxford: Oxford University Press, 1999; Marilyn Strathern, *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy*, London: Routledge, 2000.

non-governmental and civil society indices, such as the 'Corruption Perceptions Index', the 'Press Freedom Index', the 'Happy Planet Index', the 'Financial Secrecy Index', the 'Global Slavery Index' and the 'Global Food Index'.³

The rise of the index can also be understood in relation to recent research exploring the social life and performative effects of rankings, indices and indicators.⁴ Indices not only *represent* aspects of the world, they can also participate in *shaping* the world, including through reactive effects. They enable scores, rankings and comparisons of different issues across countries through processes of commensuration and quantification. Indices can thus be envisaged as devices for the production of social facts, which in turn enable different dynamics of competition and concern. The following passage from the press release accompanying a recent edition of the 'Corruption Perceptions Index', presented alongside coloured maps and rankings, provides an example of how indices can give rise to 'enumerated entities'⁵ as objects of attention and concern:

This year, New Zealand and Denmark rank highest with scores of 89 and 88 respectively. Syria, South Sudan and Somalia rank lowest with scores of 14, 12 and 9 respectively. The best performing region is Western Europe with an average score of 66. The worst performing regions are Sub-Saharan Africa (average score 32) and Eastern Europe and Central Asia (average score 34).⁶

Here the function of the index is render corruption across different countries commensurable, comparable and quantifiable. The scores enable normative claims about which countries are doing better and worse, and aspirations for future action.⁷

The Open Data Index may be considered a *database about data*, as it is concerned with the availability of governmental data across borders. Along with other similar initiatives such as the Open Data Barometer, the index raises questions about which data is to be made available and how it is to be made available. In doing so it also surfaces issues around the governance and politics of public data, such as who gets to decide what is measured and what kinds of reactive effects are imagined and observed. While previous literature discusses this index in

3 The corresponding organisation in order of the indices named: Transparency International, Reporters Without Border, New Economics Foundation, Tax Justice Network, Walk Free, Oxfam.

4 Wendy N. Espeland and Michael Sauder, 'Rankings and Reactivity: How Public Measures Recreate Social Worlds', *American Journal of Sociology*, 113 (2007): 1-40. Richard Rottenburg, Sally E. Merry, Sung-Joon Park, and Johanna Mugler (eds.), *The World of Indicators: The Making of Governmental Knowledge through Quantification*, Cambridge: Cambridge University Press, 2015.

5 Helen Verran, 'Number as an inventive frontier in knowing and working Australia's water resources', *Anthropological Theory*, 10 (2010): 171-178; Helen Verran, 'Enumerated Entities in Public Policy and Governance', in Ernest Davis and Philip J. Davis (eds), *Mathematics, Substance and Surmise*, New York: Springer, 2015, pp. 365-379.

6 Transparency International, 'Corruption Perceptions Index', *Transparency International blog*, 21 February 2018, https://www.transparency.org/news/feature/corruption_perceptions_index_2017.

7 Jane Guyer, 'Percentages and perchance: archaic forms in the twenty-first century', *Distinktion, Journal of Social Theory* 15 (2014): 155-173.

the context of information policy and the practices of public institutions,⁸ it can also be considered in terms of recent work on data politics, including 'stactivism'⁹ and 'data activism'.¹⁰ The Open Data Index can be envisaged not only as a way to measure accessibility but also as a particular kind of intervention around official regimes of quantification and datafication - including around the horizons of intelligibility, the formation of collectives and the varieties of transnational coordination that they give rise to.¹¹

Year	Number of submissions	Number of countries	% of "open" datasets
2012	177	34	Not given
2013	597	77	16%
2014	970	97	12%
2015	1586	122	9%
2016/17	1410	94	11%

Table 1: Numbers of submissions, numbers of countries and % of open datasets from Open Data Index and Open Data Census 2012-2017. Numbers obtained from archived websites and materials.

In what follows we examine what the Open Data Index is, what it does and how it has developed from 2012 to 2017 (Table 1) with reference to its various websites (current and archived), blog posts, reports, events, videos, software repositories, mailing lists and other materials. We also draw on our own involvement with the project in various capacities. We are particularly interested in how it organises participation around both *making public data* (i.e. what counts, what is counted, the forms of quantification and datafication which are held to matter) and *making data public* (i.e. the specific social and technical arrangements for making information available). We consider the index as an online 'device',¹² which both

-
- 8 Sharon S. Dawes, Lyudmila Vidiysova and Olga Parkhimovich, 'Planning and designing open government data programs: An ecosystem approach', *Government Information Quarterly* 33 (2016): 15-27; Jeffrey Thorsby, Genie N.L. Stowers, Kristen Wolslegel, and Ellie Tumbuan, 'Understanding the content and features of open data portals in American cities', *Government Information Quarterly* 34 (2017): 53-61.
- 9 Isabelle Bruno, Emmanuel Didier and Tommaso Vitale 'Statactivism: Forms of Action Between Disclosure and Affirmation', *Partecipazione e Conflitto*, 7 (2014).
- 10 Stefania Milan and Lonneke van der Velden, 'The Alternative Epistemologies of Data Activism', *Digital Culture & Society* 2 (2016): 57-74.
- 11 Jonathan Gray, 'Three Aspects of Data Worlds', *Krisis: Journal for Contemporary Philosophy* (2018).
- 12 John Law and Evelyn Ruppert, 'The Social Life of Methods: Devices', *Journal of Cultural Economy* 6 (2013): 229-240.

shapes and is shaped by the assembly of publics around the openness and availability of data as a 'matter of concern'.¹³

As a participatory device, the index plays two roles which may sometimes be in tension: (i) to *conventionalise* the assessment of the openness of data (thus facilitating comparability, objectivity and coordination across settings); and (ii) to facilitate public involvement in a way which is *receptive and flexible* enough to align with diverse interests, issues and activities around opening up governmental data. With the Open Data Index it is notable that this tension between conventionalisation and receptivity plays out through an open-ended invitation from a non-profit organisation to involve various publics through open source software and replicable components which enable the adaptation and multiplication of projects, including through forking.¹⁴

Dataset *

Please pick from the list. This is dataset you are responding about (for different datasets just submit this form multiple times)

Data Availability *

Please answer regarding open data availability *right now*. If data will be released soon you can mention this in the details and comments section below.

	Yes	No	Unsure
Does the data exist?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Is it in digital form?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Is it machine readable? (E.g. spreadsheet not PDF)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Available in bulk? (Can you get the whole dataset easily)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Is it publicly available, free of charge?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Is it openly licensed? (as per the http://OpenDefinition.org/)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Is it up to date?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Location of Data Online

(Optional) e.g. URL, if available

Date it became available

(Optional) yyyy-mm-dd format if possible

Details and comments

e.g. is the data partially available, are there plans to make it available in the future?

Figure 1: Detail of Open Data Census submission form showing types of data availability.

13 Bruno Latour, 'Why Has Critique Run out of Steam? From Matters of Fact to Matters of Concern', *Critical Inquiry* 30 (2004): 225-248.

14 Nathaniel Tkacz, *Wikipedia and the Politics of Openness*, Chicago: University Of Chicago Press, 2014.

In July 2011 the idea of an index on 'Open Knowledge' or 'Open Data' arose on the mailing list of a working group on 'open economics', advocating for open access to economic data, publications, code and other materials. As one group member put it:

There are many indices out there - for tracking democracy, corruption, innovation and human development - so why not a measure to track progress in opening government?

The immediate predecessor to the project was the 'Open Data Census', conceived in 2011 and described as 'an attempt to monitor the current status of open data across the globe'. Information about these government datasets was initially gathered with a Google Form embedded on a dedicated website, which recognised and recorded four types of availability: 'a) available in a digital form; b) machine-readable; c) publicly available, free of charge; d) openly licensed'. This later developed into seven questions about data availability (Figure 1), which were used to evaluate the openness of ten areas of data 'which most governments could reasonably be expected to collect':

1. Election Results (national)
2. Company Register
3. National Map (Low resolution: 1:250,000 or better)
4. Government Budget (high level - spending by sector)
5. Government Budget (detailed - transactional level data)
6. Legislation (laws and statutes)
7. National Statistical Office Data (economic and demographic information)
8. National Postcode/ZIP database
9. Public Transport Timetables
10. Environmental Data on major sources of pollutants (e.g. location, emissions)

The census was promoted via email lists, events, blog posts and social media platforms, including through two URLs: one for submissions (with the embedded Google Form) and one with visualisations and numbers summarising the results (Figure 2).

Open Government Data Census

Number of countries **34** Number of responses **177** Number of datasets **171** Number of open key datasets **33**

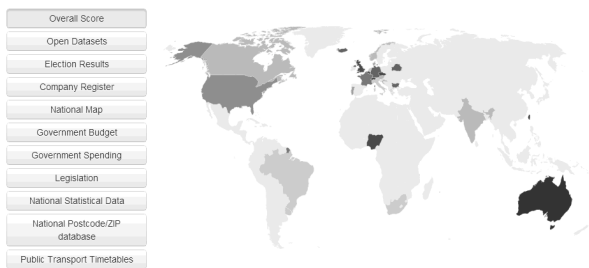


Figure 2: Map showing preliminary results of the 2012 Open Government Data Census.

The Open Data Census was originally envisaged as a means to 'gather responses from every country in the world' around the openness of different types of government data through pre-structured options ('yes', 'no', 'unsure') to questions through which this could be evaluated. This basic format served as the basis for the Open Data Index and associated projects. As the project served to assemble publics to monitor the openness of government data, we can consider it not only as in terms of its *analytical* capacities, but also in terms of its *interactivity* as a form of 'infrastructuring' around public sector data.¹⁵ In the case of the 2012 census, the input was structured through the form, and feedback was invited through an email alias and a public mailing list.

What kind of participation does the index facilitate? One approach would be to consider the dynamics between the 'formal social enterprise' of Open Knowledge International, the non-governmental organisation which coordinates the project;¹⁶ and its 'organised publics', i.e. the various contributors to the census.¹⁷ The politics of the project play out between the views of contributors, how project coordinators manage input and resolve issues, and how the project was situated in relation to the strategic and organisational prerogatives of its host NGO, its collaborators, funders, states, IGOs, and other actors. Some of these issues were raised in a blog post by open data advocate David Eaves highlighting the potential risk that 'British biases - with its highly centralized national government - have strongly shaped the census':

Thus, while the census evaluates countries some of the data sets being counted are not controlled by national governments. For example - will national governments Canada or the United States get counted for public transport data if any of their cities release transit data? Indeed, transit data - while demonstrably useful - strikes me as an odd duck choice since it is almost always not managed by national governments. The same can be said for company/corporate registers, in which the most important data sets are managed by sub-national entities.¹⁸

Inquiring about the 'details about the selection process' regarding the datasets included in the 2012 census, he further suggested others that he wished to see added to the list, including:

- Access to Information (ATIP or FOIA) made, completed, rejected and average response time, broken down by government entity
- Government procurements and contracts, broken down by government entity
- Electoral Boundary Data
- Voting Booth Locations

15 Christopher A Le Dantec, and Carl DiSalvo, 'Infrastructuring and the formation of publics in participatory design', *Social Studies of Science* 43 (2013).

16 The NGO underwent several name changes over the course of the project. It started as Open Knowledge Foundation (OKF), then became Open Knowledge (OK) and finally Open Knowledge International (OKI). For clarity we use the current name throughout.

17 Adam Fish, Christopher M. Kelty, Luis F.R. Murillo, Lily Nguyen, and Aaron Panofsky, 'Birds of the Internet', *Journal of Cultural Economy* 4 (2011): 157-187.

18 David Eaves, 'How To Evaluate The State Of Open Data', *Techpresident*, 8 May 2012, <http://techpresident.com/news/wegov/22161/how-evaluate-state-open-data>.

- Land Registry Data
- Payments to Government for Extractive Industry Resources
- Foreign Aid Data
- Campaign Finance Data
- Registered Lobbyists List

In 2013 saw the release of the Open Data Index as a 'community-based effort' which is 'compiled using contributions from civil society members and open data practitioners around the world' and 'peer-reviewed and checked by expert open data editors'.¹⁹ Described as 'a Scoreboard for Open Government Data', it was explicitly positioned in relation to other civil society index projects:

Inspired by work such as Open Budget Index from the International Budget Partnership, the Aid Transparency Index from Publish What You Fund, the Corruption Perception Index from Transparency International and many more, we felt a key aspect is to distill the results into a single overall ranking and present this clearly.²⁰

The transition from 'census' to 'index' brought a number of changes. The census remained the mechanism for collecting data about the openness of government data, and the index became the mechanism for displaying rankings. A numerical score was introduced for each country by weighting different aspects of data availability (Figure 3). These weightings rewarded specific legal and technical conventions associated with open data.²¹ For example, while 5 points (out of a total of 100) would be added if the dataset was available in digital form, 15 points would be accrued if it was 'machine readable' and 30 points if a dataset was 'openly licensed'. These dataset scores would then be used as the basis for an overall country score. Any number of submissions could come in for a given country through the website. These would be reviewed by a 'country editor' (which could be one person or a shared role), who would check and approve incoming submissions (Figure 4). A community manager at OKI worked to coordinate the project and communicate with contributors and editors.

19 Open Knowledge International, 'Government data still not open enough - new survey on eve of London summit', *Open Knowledge International blog*, 28 October 2013, <https://blog.okfn.org/2013/10/28/government-data-still-not-open-enough/>.

20 Rufus Pollock, 'The Open Data Census - Tracking the State of Open Data Around the World', *Open Knowledge International blog*, 20 February 2013, <https://blog.okfn.org/2013/02/20/open-data-census-tracking-the-state-of-open-data-around-the-world/>.

21 Jonathan Gray, 'Three Aspects of Data Worlds', *Krisis: Journal for Contemporary Philosophy* (2018).

Question	Details	Weighting
Does the data exist?	Does the data exist at all? The data can be in any form (paper or digital, offline or online etc). If the answer is "no", then please submit the form (below); there is no need to answer any further questions.	5
Is data in digital form?	Is it in digital form (stored on computers or digital storage) or is it only in e.g. paper form.	5
Publicly available?	Is the data "public" - this does NOT require freely available but does require that "someone" outside of the government can access in some form (e.g. if the data is available for purchase it is public, if the timetables exist as PDFs on a website that you can access it is public, if you can get it in paper form it is public). If a freedom of information request or similar is needed to access the data, answer "no".	5
Is the data available for free?	Is the data available for free or is there a charge? If there is a charge please note this in the further information below.	15
Is the data available online?	Is the data available online from an official source. If yes please put the URL in the Source field below.	5
Is the data machine readable?	Data is machine readable if it is in a format that can be easily processed by a computer. Data can be digital but not machine readable. For example, consider a PDF document containing tables of data. These are definitely digital but are not machine-readable because a computer would struggle to access the tabular information (even though they are very human readable). The equivalent tables in a format such as a spreadsheet would be machine readable. Note: The appropriate machine readable format may vary by type of data - so, for example, machine readable formats for geographic data may be different than for tabular data. In general, HTML and PDF are "not" machine-readable.	15
Available in bulk?	Data is available in bulk if you download or access the "whole" dataset easily. Conversely it is non-bulk if you are you limited to just getting parts of the dataset (for example, are you restricted to querying a web form and retrieving a few results at a time from a very large database).	10
Openly licensed?	As per http://opendefinition.org/ - does the terms or use or license allow anyone to freely use, reuse or redistribute the data (subject at most to attribution or sharealike requirements). It is vital that a licence is available (if there's no licence, the data is not openly licensed). Open Licences which meet the requirements of the Open Definition are listed at http://opendefinition.org/licenses/ .	30
Is the data provided on a timely and up to date basis?	Is the data up to date and timely or long delayed. For example, for election data is it made available immediately or soon after the election or is it only available many years later. Please add comments below if you have uncertainty here.	10

Figure 3: Table showing weightings for Open Data Index 2013.

Dataset	Score	Breakdown	Location (URL)	Information
Transport Timetables	10%			Propose Revisions
← Awaiting review	9%		https://www.obfahrplan...	Review Now
Government Budget	60%		https://www.govdata.d...	Propose Revisions
← Awaiting review	9%		https://www.govdata.d...	Review Now
Government Spending	10%			Propose Revisions
Election Results	100%		https://www.govdata.d...	Propose Revisions
Company Register	25%		https://www.unterneh...	Propose Revisions
← Awaiting review	9%		https://www.unterneh...	Review Now
National Map	75%		https://www.govdata.d...	Propose Revisions
← Awaiting review	9%		https://www.govdata.d...	Review Now
National Statistics	60%		https://www.destatis.d...	Propose Revisions
← Awaiting review	9%		https://www.destatis.d...	Review Now
Legislation	60%		http://www.gesetze-im...	Propose Revisions
← Awaiting review	9%		http://www.gesetze-im...	Review Now
Postcodes / Zipcodes	15%			Propose Revisions
← Awaiting review	9%			Review Now
Emissions of pollutants	40%		http://www.portalu.de	Propose Revisions
← Awaiting review	9%		http://www.thru.de	Review Now

Figure 4: Review process for contributors and editors, Open Data Index 2013.

In terms of its material organisation, the index was a *bricolage* of different elements. It continued to use a Google Form embedded on a website to gather submissions accompanied by a review process conducted with Google Spreadsheets. Submissions were gathered through the census, then normalised, reviewed and published as part of the index. Results were displayed on a NodeJS application, deployed on Heroku, and the code was made available on GitHub. A dedicated '[open-data-census]' public mailing list was set up for contributors to discuss and support the review process. On Twitter the hashtags #OpenDataCensus and #OpenDataIndex were used to promote activity and discussion around the project.

Activity around the census was organised to align with relevant international events. An 'Open Data Census Challenge' was hosted on International Open Data Day in February 2013 to encourage public involvement in the assessment of not only country-level data, but also city-level datasets. Events and activities took place in Amsterdam, Berlin, Prague, London, Shanghai, Montevideo and other cities, organised in association with local groups and partners such as the newspaper Zeit Online in Germany and the Fond Otakara Motejla in the Czech Republic, resulting in a dedicated city-focused section of the census (census.okfn.org/city).²² Another push of activity took place in the run up to the 39th G8 Summit in Lough Erne, UK, which included a focus on 'tax and transparency' and the release of a G8 'Open Data Charter'.²³ Another mini-site focusing on G8 members was released ahead of the summit.²⁴ This was editorialised with a press release contending that 'G8 countries must work harder to open up essential data' and that 'progress is lagging behind promise', which was translated into several languages by volunteers.²⁵ These two examples suggest how the index was adapted and aligned with subnational and transnational events and advocacy opportunities, in order to intervene at different scales in accordance with different event-based and policy-oriented rhythms.

The mailing lists and communications activities around the index also surfaced frictions in the process of creating a single score and ranking for the openness of data in countries and cities around the world. Submissions included comments and queries about pricing, licensing and the availability of data - such as concerns that the norms of the index were in tension with administrative and governance arrangements of countries being scored. A researcher contributing to the Canadian assessment suggested 'the methodology does not work well for federations with divisions of powers and thus different data responsibilities at different levels of jurisdiction'.²⁶ In a similar vein, a submission for Germany stated:

22 Christian Villum, 'The Open Data Census Challenge on Open Data Day 2013', *Open Knowledge International blog*, 5 March 2013, <https://blog.okfn.org/2013/03/05/the-open-data-census-challenge-on-open-data-day-2013/>.

23 Prime Minister's Office, '2013 Lough Erne G8 Leaders' Communiqué', 18 June 2013, <https://www.gov.uk/government/publications/2013-lough-erne-g8-leaders-communique>.

24 census.okfn.org/g8.

25 Rufus Pollock, 'G8 countries must work harder to open up essential data', *Open Knowledge International blog*, 14 June 2013, <https://blog.okfn.org/2013/06/14/g8-countries-must-work-harder-to-open-up-essential-data/>.

26 <http://lists-archive.okfn.org/pipermail/open-data-census/2013-July/000082.html>.

... the census it very tricky for a country like Germany with a federal system. Some of the datasets simply are not available at the national level as they belong to the federal states. To draw a [realistic] picture for Germany, we would need to do a census 16+1 (for the 16 federal states + 1 for the federal government).²⁷

Over the coming years the index explored different ways of responding to the tensions between conventionalising the assessment of open data for a range of data types, and the wide range of practices, policies and advocacy initiatives around government data. One way was to provide more support for common assessment practices. As well as having editors and review processes, 2014 saw the introduction of 'mentors', video tutorials and online drop-in sessions catering to different time zones. Continuing to evaluate government data from around the world according to a common set of metrics enabled the production of rankings, as well as claims such as 'only 11% of the world's published datasets are open according to the open definition' (Figure 5). We might consider the value and practical utility accorded to such rankings and claims in relation to their reactive effects in the context of information policy. As with the examples of indices above, activists and public servants were thus provided with material and 'enumerated entities' to make the case for changes.²⁸

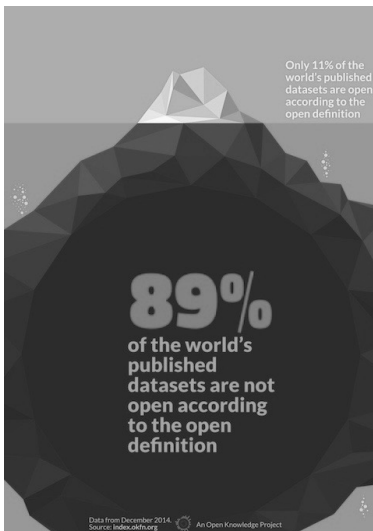


Figure 5: Graphic to illustrate results of Open Data Index 2014.

Another way was to enable more flexibility and customisation for different practices for assessing the openness of data at different scales and in different settings. In addition to the coun-

²⁷ Submission via Google Form for Open Data Census 2013.

²⁸ Helen Verran, 'Number as an inventive frontier in knowing and working Australia's water resources'. *Anthropological Theory* 10 (2010): 171-178. Helen Verran, 'Enumerated Entities in Public Policy and Governance', in Ernest Davis and Philip J. Davis (eds), *Mathematics, Substance and Surmise*, New York: Springer, 2015, pp. 365-379.

try-level censuses that were the basis for the Open Data Index, there were a proliferation of more local censuses and topical censuses, including through collaborations and alliances with other civil society organisations and projects such as the Sunlight Foundation and Code for America.²⁹ This enabled the inclusion of datasets which were deemed most relevant for different situations. For example, several groups in Belgium ran a series of workshops, consultations and other activities to solicit input on what should be included in their census to create 'a crowdsourced product that is specifically designed for cities in Belgium'.³⁰

The 2015 Global Open Data Index (GODI) saw further efforts to both consolidate the transnational assessment process as well as to broaden it to include other interests and concerns. A list of 'dataset definitions' sought to clarify assumptions around how the openness of government data should be assessed, including the role of states in coordinating the production of data. For example, it was claimed:

Our assumption the national government has a role as a regulator to create and enforce publishing [of] such data. Therefore, even if the data is not produced by the government, we see it as responsible to ensure the open publication of the data.

The question of which forms of data public institutions do create and should create raised further issues and concerns around different histories and contexts of governance and administration, such as the following comment from a Canadian open data advocate:

This doesn't work for Canada. There is a clear division of powers in our constitution between the federal government and the provinces. The federal government isn't going to walk into a constitutional battle over e.g. education data just to satisfy this index... Different provinces collect health and education data in different ways. It may not be possible to have one consolidated dataset. Please, just respect countries' differences instead of harmonizing everything to a central government structure like the UK.

UK government officials from Government Digital Service also commented that 'some criteria are an awkward fit in different national contexts', contending that changes in the assessment criteria around election data had 'cost [them] the top spot' and that '[they'd] need to change the laws governing our electoral system to make progress'.³¹

There was also a consultation around which datasets should be included in the 2015 edition of the Global Open Data Index, which encouraged participants to vote on different datasets as well as to propose other suggestions (Figure 6). This was part of a repositioning from 'a

29 Rufus Pollock, 'Announcing the Local Open Data Census', *Open Knowledge International blog*, 4 February 2014, <https://blog.okfn.org/2014/02/04/announcing-the-local-open-data-census/>.

30 Pieter-Jan Pauwels, 'Results of Wiki Survey and final steps', 12 September 2014, <http://openbelgium.be/2014/09/results-of-wiki-survey-and-final-steps/>.

31 Oliver Buckley, 'Open Data - the race to the top', *Gov.Uk blog*, 15 December 2015, <https://data.blog.gov.uk/2015/12/15/open-data-the-race-to-the-top/>.

simple measurement tool' to 'civil society audit of the open data revolution' exploring 'which datasets are of high social and democratic value', resulting in the addition of five new datasets: government procurement data, water quality, weather forecasts, land ownership and health performance data.³² Public input and involvement was also encouraged in a forum, which saw more extended discussions about questions and issues around the index.

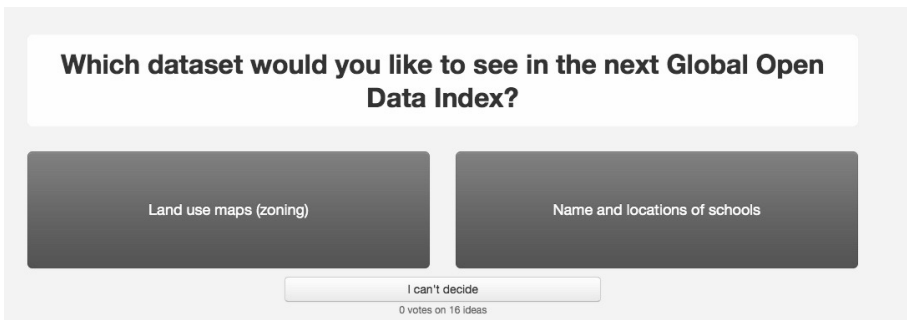


Figure 6: "Crowdsourced survey" to inform choice of datasets included in 2015 Global Open Data Index.

The code for the project was used to create a prototype of a 'Global Goals Data Census', in order to track data relevant to the United Nations Sustainable Development Goals (SDGs), raising the question of the relationship between the transnational coordination of societal progress and the openness of governmental data.³³

The 2016 edition saw a shift from 'countries' to 'places'. This was said to be because national government is not always the most important unit of analysis (especially with regards to sub-national governments with 'administrative and legislative autonomy'). It also aimed to accommodate 'submissions for places that are not officially recognised as independent countries'.³⁴ A 'dialogue phase' in the assessment process was also introduced in order to enable 'civil society and government [to] talk to one another'.

More recently there has been more reflection about what the Open Data Index does. Blog posts in 2017 have warned against 'volunteer fatigue',³⁵ guarded against inappropriate com-

32 Mor Rubinstein, 'What should we include in the Global Open Data Index? From reference data to civil society audit', *Open Knowledge International blog*, 18 June 2015, <https://blog.okfn.org/2015/06/18/what-should-we-include-in-the-global-open-data-index-from-reference-data-to-civil-society-audit/>.

33 Mikel Maron, 'Let's Build the Global Goals Data Census', *Medium*, 5 October 2015, <https://medium.com/@mikelmaron/let-s-build-the-global-goals-data-census-c38b0458c9a>.

34 <https://index.okfn.org/methodology/>.

35 Oscar Montiel, 'Our Country Sample and What It Tells Us About Our Contributors', *Open Knowledge International blog*, 3 May 2017, <https://blog.okfn.org/2017/05/03/our-country-sample-and-what-it-tells-about-our-contributors/>.

parisons between years after changes in assessment methodology,³⁶ and advocated for a 'shift in focus from a mere measurement tool to a stronger conversational device'.³⁷

In summary, the Open Data Census and the Open Data Index aim to intervene around what is considered 'good data' by assessing the extent to which the publication of government data conforms with specific legal and technical norms and conventions, such as open licensing and machine-readability. It can thus be read in relation to the history of ideals and practices of open data, as well as free/open source software.³⁸ In contrast to other civil society indices which aim to advance their issue work by means of specialised research and policy teams, the Open Data Index has been designed to organise public involvement around the assessment of public data. It might thus be considered in terms of its capacities to enable not only *analysis*, but also *interactivity* using digital technologies - as well as structuring participation in particular ways which are not without tension.³⁹ Through this discussion of the development and social life of the Open Data Index, we suggest how it may be viewed as a participatory device whose function may vary from 'crowd-sourcing' assessment, to facilitating more substantive deliberation and public involvement around the politics of public data. Such indices may thus be viewed as sites where participation is both configured and contested, and where practices of valuation and enumeration are both conventionalised and brought into question.

References

- Bruno, Isabelle, Emmanuel Didier, Tommaso Vitale. 'Statactivism: Forms of Action. Between Disclosure and Affirmation', *Partecipazione e Conflitto* 7 (2014).
- Buckley, Oliver. 'Open Data - the race to the top', *Gov.Uk blog*, 15 December 2015, <https://data.blog.gov.uk/2015/12/15/open-data-the-race-to-the-top/>.
- Dawes, Sharon S., Lyudmila Vidasova and Olga Parkhimovich. 'Planning and designing open government data programs: An ecosystem approach', *Government Information Quarterly* 33 (2016): 15-27.
- Eaves, David. 'How To Evaluate The State Of Open Data', *Techpresident*, 8 May 2012, <http://techpresident.com/news/wegov/22161/how-evaluate-state-open-data>.
- Espeland, Wendy N., and Mitchell L. Stevens. 'Commensuration as a Social Process', *Annual Review of Sociology* 24 (1998): 313-343.
- Espeland, Wendy N., and Michael Sauder. 'Rankings and Reactivity: How Public Measures Recreate Social Worlds', *American Journal of Sociology* 113 (2007): 1-40.

36 Oscar Montiel, 'Our Country Sample and What It Tells Us About Our Contributors', *Open Knowledge International blog*, 3 May 2017, <https://blog.okfn.org/2017/05/03/our-country-sample-and-what-it-tells-us-about-our-contributors/>.

37 Open Knowledge International, 'The Future of the Global Open Data Index: Assessing The Possibilities', *Open Knowledge International blog*, 2 November 2017, <https://blog.okfn.org/2017/11/01/the-future-of-the-global-open-data-index-assessing-the-possibilities/>.

38 Christopher M. Kelty, *Two Bits: The Cultural Significance of Free Software: The Cultural Significance of Free Software and the Internet*, Durham NC: Duke University Press, 2008. Jonathan Gray, 'Towards a Genealogy of Open Data', *European Consortium for Political Research General Conference*, Glasgow, 3-6 September 2014, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2605828; Jonathan Gray, 'Three Aspects of Data Worlds', *Krisis: Journal for Contemporary Philosophy* (2018).

39 Noortje Marres, *Digital Sociology: The Reinvention of Social Research*, Cambridge: Polity Press, 2017.

Fish, Adam, Christopher M. Kelty, Luis F.R. Murillo, Lilly Nguyen, and Aaron Panofsky. 'Birds of the Internet'. *Journal of Cultural Economy* 4 (2011): 157-187.

Gray, Jonathan. 'Towards a Genealogy of Open Data', *European Consortium for Political Research General Conference*, Glasgow, 3-6 September 2014, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2605828.

_____. 'Three Aspects of Data Worlds', *Krisis: Journal for Contemporary Philosophy* (2018).

Guyer, Jane. 'Percentages and perchance: archaic forms in the twenty-first century', *Distinktion, Journal of Social Theory* 15 (2014): 155-173.

Kelty, Christopher M. *Two Bits: The Cultural Significance of Free Software: The Cultural Significance of Free Software and the Internet*, Durham NC: Duke University Press, 2008.

Latour, Bruno. 'Why Has Critique Run out of Steam? From Matters of Fact to Matters of Concern', *Critical Inquiry* 30 (2004): 225-248.

_____. *An Inquiry into Modes of Existence: An Anthropology of the Moderns*, trans C. Porter, Cambridge MA: Harvard University Press, 2013.

Law, John, and Evelyn Ruppert. 'The Social Life of Methods: Devices', *Journal of Cultural Economy* 6 (2013): 229-240.

Le Dantec, Christopher A., and Carl DiSalvo. 'Infrastructuring and the formation of publics in participatory design', *Social Studies of Science* 43 (2013).

Maron, Mikel. 'Let's Build the Global Goals Data Census', *Medium*, 5 October 2015, <https://medium.com/@mikelmoron/let-s-build-the-global-goals-data-census-c38b0458c9a>.

Marres, Noortje. *Digital Sociology: The Reinvention of Social Research*, Cambridge: Polity Press, 2017.

Milan, Stefania and Lonneke van der Velden. 'The Alternative Epistemologies of Data Activism', *Digital Culture & Society* 2 (2016): 57-74.

Montiel, Oscar. 'Our Country Sample and What It Tells Us About Our Contributors', *Open Knowledge International blog*, 3 May 2017, <https://blog.okfn.org/2017/05/03/our-country-sample-and-what-it-tells-about-our-contributors/>.

Open Knowledge International. 'Government data still not open enough - new survey on eve of London summit', *Open Knowledge International blog*, 28 October 2013, <https://blog.okfn.org/2013/10/28/government-data-still-not-open-enough/>.

_____. 'The Future of the Global Open Data Index: Assessing the possibilities', *Open Knowledge International blog*, 2 November 2017, <https://blog.okfn.org/2017/11/01/the-future-of-the-global-open-data-index-assessing-the-possibilities/>.

Pauwels, Pieter-Jan. 'Results of Wiki Survey and final steps', 12 September 2014, <http://openbelgium.be/2014/09/results-of-wiki-survey-and-final-steps/>.

Pollock, Rufus. 'The Open Data Census - Tracking the State of Open Data Around the World', *Open Knowledge International blog*, 20 February 2013, <https://blog.okfn.org/2013/02/20/open-data-census-tracking-the-state-of-open-data-around-the-world/>.

_____. 'G8 countries must work harder to open up essential data', *Open Knowledge International blog*, 14 June 2013, <https://blog.okfn.org/2013/06/14/g8-countries-must-work-harder-to-open-up-essential-data/>.

_____. 'Announcing the Local Open Data Census', *Open Knowledge International blog*, 4 February 2014, <https://blog.okfn.org/2014/02/04/announcing-the-local-open-data-census/>.

Power, Michael. *The Audit Society: Rituals of Verification*, Oxford: Oxford University Press, 1999.

Prime Minister's Office. '2013 Lough Erne G8 Leaders' Communiqué', 18 June 2013, <https://www.gov.uk/government/publications/2013-lough-erne-g8-leaders-communique>.

Rottenburg, Richard, Sally E. Merry, Sung-Joon Park and Johanna Mugler (eds). *The World of Indicators: The Making of Governmental Knowledge through Quantification*, Cambridge: Cambridge University Press, 2015.

Rubinstein, Mor. 'What should we include in the Global Open Data Index? From reference data to civil society audit', *Open Knowledge International blog*, 18 June 2015, <https://blog.okfn.org/2015/06/18/what-should-we-include-in-the-global-open-data-index-from-reference-data-to-civil-society-audit/>.

Strathern, Marilyn. *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy*, London: Routledge, 2000.

Thorsby, Jeffrey, Genie, N.L. Stowers, Kristen Wolslegel and Ellie Tumbuan. 'Understanding the content and features of open data portals in American cities', *Government Information Quarterly* 34 (2017): 53-61.

Tkacz, Nathaniel. *Wikipedia and the Politics of Openness*, Chicago: University Of Chicago Press, 2014.

Transparency International. 'Corruption Perceptions Index', *Transparency International blog*, 21 February 2018, https://www.transparency.org/news/feature/corruption_perceptions_index_2017.

Verran, Helen. 'Number as an inventive frontier in knowing and working Australia's water resources', *Anthropological Theory* 10 (2010): 171-178.

_____. 'Enumerated Entities in Public Policy and Governance', in Ernest Davis and Philip J. Davis (eds), *Mathematics, Substance and Surmise*, New York: Springer, 2015, pp. 365-379.

Villum, Christian. 'The Open Data Census Challenge on Open Data Day 2013', *Open Knowledge International blog*, 5 March 2013, <https://blog.okfn.org/2013/03/05/the-open-data-census-challenge-on-open-data-day-2013/>.

BIOGRAPHIES

Bruce Baer Arnold is an Assistant Professor at Canberra Law School (University of Canberra). Dr Arnold's doctoral dissertation explored the nature of legal identity, including the information state, autonomy and population-scale data-mining. He has served on OECD health informatics and data protection working parties, alongside Vice-Chairmanship of the Australian Privacy Foundation, the Australia's preeminent civil society body concerned with data protection. His work on new technology law and privacy has appeared in leading journals and collections. He is a contributor to Australia's authoritative privacy/confidentiality practitioner resource. Dr Arnold's current research focus is on regulatory incapacity regarding medical products and data.

Wendy Elizabeth Bonython is an Adjunct Associate Professor at Canberra Law School (University of Canberra). Dr Bonython has graduate qualifications in molecular genetics and law, with a background in health research, public administration and policy development. She serves on government and professional committees, in addition to providing invited expert testimony to a range of legislative and law reform inquiries. Dr Bonython's work has appeared in leading Australian and overseas law journals, including studies of the interaction between tort law and practitioner regulation, tort theory and law's understanding of property in bodies and intangibles. Her research encompasses professional and public understandings of risk, harm and innovation.

Rachelle Bosua is Assistant Professor in the Business Process Management & IS Dept of the Management Science and Technology Faculty at the Open University of The Netherlands. She is also an Honorary Senior Fellow in the Department of Computing and Information Systems (CIS) at the University of Melbourne. She is co-author of the book *Knowledge Management in organizations: a critical introduction* (Oxford University Press 2018). Her research focuses on data and information privacy that relates to human behavior and the use of information and communication technologies.

Ellen Broad is Head of Technical Delivery, Consumer Data Standards at CSIRO's Data61. Previous roles include Head of Policy for the Open Data Institute (ODI), an international non-profit founded by Sir Tim Berners-Lee and Sir Nigel Shadbolt, and adviser to senior UK government minister Elisabeth Truss on data, both in the UK. She has also held roles as Manager of Digital Policy and Projects for the International Federation of Library Associations and Institutions (Netherlands) and Executive Officer for the Australian Digital Alliance. Ellen has written and spoken about AI, open data and data sharing issues in the *New Scientist*, the *Guardian* and a range of technology publications, and is the author of *MADE BY HUMANS: The AI Condition* published by Melbourne University Publishing (MUP) in August 2018.

Anna Bruce is a Senior Lecturer in the School of Photovoltaic and Renewable Energy Engineering at UNSW and Research Coordinator (Engineering) at the Centre for Energy and Environmental Markets. Her research focuses on energy transitions, including renewable and distributed energy system and energy market modelling. Current research projects include Energy Data for Smart Decision Making through the Australian Government's Smart Cities and

Suburbs Program, and Integrated Smart Home Energy Management Technologies through the CRC-P program. She participates as an Australian expert in the IEA's Photovoltaic Power Systems and Demand Side Management technology collaboration programs.

Stephanie Carroll Rainie (Ahtna Athabascan), MPH, DrPH, is Assistant Professor, Public Health Policy and Management at the Community, Environment and Policy Department, Mel and Enid Zuckerman College of Public Health; Assistant Research Professor, Udall Center for Studies in Public Policy (UC); Associate Director and Manager - Tribal Health Program for the Native Nations Institute (<http://nni.arizona.edu/>) in the UC; and Assistant Director for the Center for Indigenous Environmental Health Research(<http://ciehr.arizona.edu/>), MEZCOPH at the University of Arizona (UA). Stephanie's research explores the links between governance, health care, the environment, and community wellness. Stephanie is the co-founder of the United States Indigenous Data Sovereignty Network and the International Indigenous Data Sovereignty Interest Group at the Research Data Alliance.

Donna Cormack (Kati Mamoe, Kai Tahu) is a Senior Lecturer and researcher with joint positions at Te Kupenga Hauora Maori, University of Auckland and Te Ropu Rangahau Hauora a Eru Pomare, Univeristy of Otago. Donna has had a long-standing interest in data issues, particularly as they relate to measuring and monitoring Maori health and ethnic health inequities. She is a member of the Te Mana Raraunga Maori Data Sovereignty Network.

Karin Clark is a Senior Fellow (Melbourne Law Masters) at the University of Melbourne. Previously, she was Special Counsel with the Melbourne office of the Allens law firm and practised in the firm's Communications, Media and Technology Practice Group specialising in advising on compliance with privacy laws.

Tyng-Ruey Chuang is an Associate Research Fellow at the Institute of Information Science, Academia Sinica (Taipei, Taiwan) with a joint appointment at both the Research Center for Information Technology Innovation and the Research Center for Humanities and Social Sciences. He was trained as a computer scientist (PhD, NYU 1993) and has been working with ecologists, historians and legal scholars to make better use of research data. He serves on the Executive Committee of CODATA (The Committee on Data for Science and Technology) at the International Science Council. He is also a member of CODATA's International Data Policy Committee.

Angela Daly is a critical socio-legal scholar of the regulation of new technologies. She is currently based in the Chinese University of Hong Kong Faculty of Law and holds adjunct positions at Queensland University of Technology Faculty of Law (Australia) and the Tilburg Institute of Law, Technology and Society (Netherlands). She is the author of *Private Power, Online Information Flows and EU Law: Mind the Gap* (Hart 2016) and *Socio-Legal Aspects of the 3D Printing Revolution* (Palgrave Macmillan 2016).

S Kate Devitt is a philosopher and cognitive scientist working as a social and ethical robotics researcher for the Australian Defence Science and Technology Group. She is an Adjunct Fellow in the Co-Innovation Group, School of Information Technology and Electrical Engineering,

University of Queensland. Her research includes: the ethics of data, barriers to the adoption of technologies, the trustworthiness of autonomous systems and philosophically designed tools for evidence-based, collective decision making.

Martin Flintham is an Assistant Professor of Computer Science at the University of Nottingham, where he is a member of the Mixed Reality Lab and Horizon Digital Economy Research Institute. His research focus on deploying and studying interactive experiences and disruptive technology probes, including domains such as the Internet of Things, ubiquitous computing, mixed reality and games. He has led successful 'in the wild' collaborations with the creative industries, leading to several highly-cited publications at ACM CHI, and winning the Prix Ars Electronica and two BAFTA nominations.

Murray Goulden is an Assistant Professor of Sociology at the University of Nottingham, and current holder of a Nottingham Research Fellowship. Through the Fellowship he is exploring the sociological implications of Internet of Things technologies for domestic life. His most recent work addresses the political economy of the 'smart home', and the implications of IoT-generated 'interpersonal data' for intimate social groups. He has worked extensively on research applying novel digital technologies to real world settings, with a focus on networking, digital data, and smart energy, their role in reconfiguring associated social practices, and the implications for policy making and design.

Timothy Graham is Postdoctoral Fellow in Sociology and Computer Science at the Australian National University. His research combines computational methods with social theory to describe, explain, and predict the dynamics of social systems. He is currently a Chief Investigator of an ARC Discovery Project that utilises large-scale hyperlink network analysis and web experiments to comparatively assess the government web presence of 10 high-tech countries. He develops open source software tools for the analysis of socially-generated data, and has published in journals such as *Information, Communication & Society*, *Information Polity*, *Big Data & Society*, and *Critical Social Policy*.

Daniel Gray (@DanielGray00) is a PhD researcher based at Cardiff University, United Kingdom. His undergraduate and postgraduate studies have focused on critical discursive approaches to social media. At the time of writing this takes the form of an analysis of neo-liberal and misogynistic discourse on Twitter, blending big data and computational methods with critical, qualitative research. His research interests are focused on Marxist and radical approaches to digital sociology, particularly political, ethical and methodological issues on social media. More generally, his interests also include the intersections of radical left politics and Christianity, and teaching in university.

Jonathan Gray is Lecturer in Critical Infrastructure Studies at the Department of Digital Humanities, King's College London, where he is currently writing a book on data worlds. He is also Cofounder of the Public Data Lab; and Research Associate at the Digital Methods Initiative (University of Amsterdam) and the médialab (Sciences Po, Paris). More about his work can be found at jonathangray.org and he tweets at @jwyg.

Miren Gutiérrez (@gutierrezmiren) holds a PhD in Communication Studies. She is the Director of the postgraduate program 'Analysis, Research and Data Communication' and lectures on Communication at the University of Deusto and is a guest lecturer at the University of Navarra. She is also an Associate Researcher at the Overseas Development Institute (ODI) in London and DATACTIVE at the University of Amsterdam. Her interest is data activism or how people and organizations use data infrastructure, in combination with other technologies, for social change, equality and environmental conservation. She is also the author of the book *Data activism and social change* (Palgrave Macmillan 2018).

Chih-hsing Ho is Assistant Research Fellow at Academia Sinica, Taiwan. Her research focuses on the nexus of law and medicine in general, with particular attention to the governance of genomics and newly emerging technologies, such as big data, biobanks and artificial intelligence (AI). She is currently a Co-PI in law for a health cloud project in Taiwan and is responsible for designing an adequate regulatory framework for the secondary use of health data and data linkage. She holds a PhD in law from the London School of Economics (LSE), a LLM from Columbia Law School, and a JSM from Stanford University.

Becky Kazansky is a PhD candidate with the DATACTIVE project at the University of Amsterdam. She has a joint affiliation with the Media Studies and Political Science departments, additionally serving on the ethics committee of the Informatics Institute. Her research focuses on resistance to surveillance amidst datafication, looking specifically at practices that try to predict and preempt surveillance. Alongside her academic research, she has worked for a decade with different human rights and social justice organisations on technology issues.

Declan Kuch is a Research Fellow in the School of Humanities and Languages at UNSW. His research is motivated by problems of reconciling public values, technoscience and democracy in an era characterized by profound environmental constraints. His research spans climate change, energy policy and the social dimensions of the life sciences, and his primary role is as a co-leader of the Social Dimensions stream of the ARC Centre of Excellence in Convergent Bio-Nano Science and Technology. He is the author of *The Rise and Fall of Carbon Emissions Trading* (Palgrave Climate Energy and Environment Series 2015).

Tahu Kukutai (Ngati Tipa, Ngati Kinohaku, Te Aupouri) is Professor of Demography at the National Institute of Demographic and Economic Analysis, University of Waikato. Tahu specialises in Maori and Indigenous demographic research and has written extensively on issues of Maori population change, Maori identity and official statistics. Tahu is a founding member of the Maori Data Sovereignty Network Te Mana Raraunga and co-editor (with John Taylor) of *Indigenous Data Sovereignty: Toward an Agenda* (ANU Press 2016). She was previously a journalist.

Danny Lämmerhirt is researcher and research coordinator at Open Knowledge International. His work focuses on the sociology of quantification, metrics and policy, data ethnography, and the data commons. Among other activities, he leads the methodology around the Global Open Data Index and co-chairs the measurement and accountability working group at the Open Data Charter. Working at the intersection between practice and theory, he explores the

'social life of metrics' and how metrics intervene and reshape organisational cultures and vice versa. He shares his work on Twitter @danlammerhirt.

Vanessa Lee (Yupungathi and Meriam people, Cape York and the Torres Strait), is a social epidemiologist and senior academic within the discipline of Behavioural and Social Sciences in the Faculty of Health Sciences at University of Sydney. Vanessa's overarching focus addresses the social issues of the burden of disease to break the cycle of inequality that potentially lead to suicide in First Nations communities, and to strengthen the health and wellness of Aboriginal and Torres Strait Islander cultural identity, particularly for women. She is a founding member of the Indigenous Data Sovereignty group in Australia - Maiam nayri Wingara.

Raymond Lovett (Wongaibon/Ngiyampaa) is an Associate Professor and leader of the Aboriginal and Torres Strait Islander health program at the National Centre for Epidemiology and Population Health, Research School of Population Health, the Australian National University. Ray is a social epidemiologist with extensive experience in health services research, public health policy development and health program evaluation. He is a founding member of the Indigenous Data Sovereignty group in Australia (Maiam nayri Wingara) and is a member of the International Indigenous Data Sovereignty Interest group at the Research Data Alliance.

Iain MacGill is an Associate Professor in the School of Electrical Engineering and Telecommunications at the University of New South Wales, and Joint Director (Engineering) for the University's Centre for Energy and Environmental Markets (CEEM). Iain's teaching and research interests include electricity industry restructuring and the Australian National Electricity Market, sustainable energy generation technologies, renewable and distributed energy integration, energy efficiency and demand-side participation, and broader energy and climate policy opportunities and challenges.

Daniel McNamara is a PhD candidate in the Research School of Computer Science at the Australian National University. He is affiliated with the Machine Learning Research Group at CSIRO Data61. Daniel's research interests include fair machine learning and representation learning. He was a visiting student researcher during 2016-17 at the Carnegie Mellon University Machine Learning Department, supported by a Fulbright Postgraduate Scholarship. His work experience includes roles as a campaign data specialist at the Australian Labor Party, as a management consultant at the Nous Group, and as an intern at Microsoft and Google.

Monique Mann is the Vice Chancellor's Research Fellow in Technology and Regulation at the Faculty of Law, Queensland University of Technology. Dr Mann is advancing a program of socio-legal research on the intersecting topics of algorithmic justice, police technology, surveillance, and transnational online policing.

Luke Marshall is a PhD student in the School of Photovoltaics and Renewable Energy Engineering and Centre for Energy and Environmental Markets in the University of New South Wales. His research centres on the intersection between energy market rules, and the integration of solar resources into the electricity grid. He loves to code (including projects such as <http://energyopensource.org/>) and also to surf.

Nikita Melashchenko is a PhD Candidate at Victoria University of Wellington. He focuses on intellectual property and regulatory theory in application to the knowledge economy. His dissertation examines data localization policies under the WTO law. Nikita is a Fulbright alumnus and earned degrees from Vanderbilt University (LLM) and SPbU (LLB & LLM). Before joining the academy he practiced in the field of law and technology.

Stefania Milan (stefaniamilan.net) is Associate Professor of New Media at the University of Amsterdam. Her work explores the intersection of digital technology, governance and activism, with emphasis on critical data practices and autonomous infrastructure. Stefania is the Principal Investigator of the DATACTIVE project, funded with a Starting Grant of the European Research Council (639379) and exploring the evolution of citizenship and participation vis-à-vis datafication and surveillance (data-activism.net), and of the spin-off ALEX project (ERC Proof of Concept 825974). She is the author of *Social Movements and Their Technologies: Wiring Social Change* (Palgrave Macmillan 2013/2016) and co-author of *Media/Society* (Sage 2011).

Cheng Soon Ong is a principal research scientist at the Machine Learning Research Group, Data61, CSIRO(<http://data61.csiro.au/>). He is also an Adjunct Associate Professor at the Australian National University. His PhD in Computer Science was completed at the Australian National University in 2005. He was a postdoc at the Max Planck Institute of Biological Cybernetics and the Fredrich Miescher Laboratory in Tübingen, Germany. From 2008 to 2011, he was a lecturer in the Department of Computer Science at ETH Zurich, and in 2012 and 2013 he worked in the Diagnostic Genomics Team at NICTA in Melbourne. Since 2014, Cheng Soon is researching ways to enable scientific discovery by extending statistical machine learning methods with the Machine Learning Research Group in Data61, CSIRO in Canberra. Prior to his PhD, he researched and built search engine and Bahasa Malaysia technologies at Mimos Berhad, Malaysia. He obtained his B.E. (Information Systems) and B.Sc. (Computer Science) from the University of Sydney, Australia.

Sefa Ozalp is a research assistant and a PhD student at the Social Data Science Lab, Cardiff University. His main research interests are computational sociology and computational criminology, particularly focusing on dissemination of multiple typologies of cyberhate and xenophobia on social media platforms. He is also interested in regulation of digital platforms and abuse of personal (meta)data by oppressive regimes, focusing particularly on Turkey.

Rob Passey is a Senior Research Associate at the Centre for Energy and Environmental Markets (CEEM), and Postdoctoral Fellow at the School of Photovoltaic and Renewable Energy Engineering (SPREE), both at UNSW Sydney. He is also a Senior Consultant with ITP Renewables, and Policy Analyst and Treasurer at the Australian PV Institute. He focuses on policy research and analysis, particularly renewable energy, distributed energy resources, energy efficiency and carbon pricing. He has published widely on market-like policies to reduce greenhouse emissions, the uptake of low emission technologies and energy efficiency, as well as distributed generation and its integration into electricity networks, including smart grids.

Chiara Poletti (@ccpollon) is an ESRC funded PhD researcher in the School of Social Sciences at Cardiff University. Her research interests cover digital rights, digital labour, cybersecurity, digital methodologies and ethics of research on big data. In the last years, she has been working on Internet governance, social media platforms and big data. Her PhD project focuses on the discursive and socio-economic aspects of social media platforms' content regulation. Previously, she studied the public discourse on social media platforms in the aftermath of terrorist attacks. She is one of the co-founders and chair of the Digital Stuff Post Graduate research group in Cardiff University (<https://digitalstuffcardiff.com>).

Colin Porlezza is Senior Lecturer in Journalism with the Department of Journalism at City, University of London. He is also a Visiting Professor at the Institute of Media and Journalism at the Università della Svizzera italiana. His research focuses on data journalism, media accountability and transparency, as well as innovation in digital journalism. His research so far has resulted in more than 30 peer-reviewed publications. He is a Board Member of the Swiss Association of Communication and Media Research (SACM) and a founding member of the European Journalism Observatory. He has featured on the the Swiss Public Service Broadcaster SRG SSR, Poynter, the Neue Zürcher Zeitung and other news outlets.

Dominic Price is a researcher at the Horizon Digital Economy Research Institute at the University of Nottingham. His research focuses on the software engineering aspect of Human-Computer Interaction and the development of platforms that bridge the gap between system level computing and human interfaces. He has most recently been involved in work in the areas of crowd-sourcing and social-networking.

Paula Restrepo is an Associate Professor at the Department of Social Communication, Faculty of Communication, Universidad de Antioquia, Medellín, Colombia. She has a degree in Anthropology and a Masters in Philosophy from Universidad de Antioquia, and a PhD in Philosophy from the Basque Country University. She is a member of the Research Group 'Grupo de investigación Comunicación Periodismo y Sociedad' at Facultad de Comunicaciones, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellín, Colombia. This chapter is a product of the Project "Prácticas comunicativas de la agricultura urbana en Medellín" (cod. 2016-12689) approved by Convocatoria Programática 2016, from the Universidad de Antioquia.

Megan Richardson is a Professor of Law, Co-Director of the Centre for Media and Communications Law, and Director of the Intellectual Property Research Institute of Australia at the University of Melbourne. Recent books include *The Right to Privacy: Origins and Influence of a Nineteenth-Century Idea* (Cambridge University Press 2017), and she is currently writing an *Advanced Introduction to Privacy Law* (Edward Elgar, forthcoming 2020).

Gerard Jan Ritsema van Eck is a PhD researcher at the Security, Technology, and e-Privacy (STeP) research group at the University of Groningen. His research focuses on participatory surveillance and privacy issues in public spaces. In particular, he is interested in how new and rapidly evolving technologies, such as smartphones and predictive analytics, enable states and large corporations to enrol citizens as data gatherers. His work combines insights

from sociology and surveillance studies with work on European (Union) data protection law to generate new insights on the changing nature of urban public spaces.

Mike Roberts is a Research Associate at the Centre for Energy and Environmental Markets (CEEM) and in the School of Photovoltaic and Renewable Energy Engineering (SPREE) at UNSW Sydney. His research interests include technical and policy aspects of integration of distributed renewable generation, storage and demand management, co-operative approaches to energy planning, including embedded networks, micro grids and community renewable energy, and solar potential assessment. Mike's doctoral thesis explored opportunities for photovoltaics and storage in multi-occupancy residential buildings.

Adam Steer is a serial multi-disciplinarian, whose career has ranged from neurophysiology laboratories to web site management to professional bicycle repair to Antarctic logistics, sea ice field research and most recently working on massive high resolution geospatial data management. The intersection of technology and human evolution is a key driver in Adam's outlook on science and data and life - looking at when we cast human problems as engineering problems, and how we can become better humans in order to build a better future. He runs the consultancy Spatialised (<http://spatialised.net>).

Naomi Stringer is undertaking a PhD in the School of Photovoltaics and Renewable Energy Engineering and Centre for Energy and Environmental Markets in the University of New South Wales. Her research is driven by a fascination with the opportunities and challenges afforded by distributed energy resources (namely rooftop PV and home battery systems) and research interests include the technical, social and regulatory aspects of distributed energy integration. Her current work examines the behaviour of rooftop solar PV systems during major electricity system disturbances. Undertaken in collaboration with the Australian Energy Market Operator and solar monitoring company, Solar Analytics, her work is heavily data driven.

Guillén Torres joined DATACTIVE as a PhD candidate to develop a project on how data mediates the interaction between organised citizens and the state. His main areas of interest are transparency, resistance and the anthropology of the state. He also serves as an advisor to the Mexican NGO ControlYourGovernment. He has a background in Urban Studies and Sociology and is a recidivist Actor-Network Theory enthusiast.

Claire Trenham has research and data experience spanning a range of physical sciences, including radio astronomy, ocean waves, and climate models, as well as substantial experience more broadly managing large scale 'nationally significant' research data collections at the Australian National Computational Infrastructure. Claire currently works for CSIRO's Climate Science Centre in the Sea level rise, waves and coastal extremes team alongside the regional climate team. Claire is heavily involved in climate modelling, data processing, data preparation and software to enhance science capabilities, as well as supervising student volunteers in historical data digitisation. Claire has a passion for making maths and science fun, relevant and accessible.

Lachlan Urquhart is a Lecturer in Technology Law at the University of Edinburgh. He is also a visiting researcher at the Horizon Digital Economy Research Institute, where he was a Research Fellow in Information Technology Law from 2016–2018. He is a multidisciplinary researcher, having degrees in both law and computer science, with an LL.B, (Hons) from University of Edinburgh, an LL.M in IT & Telecoms Law (Distinction) from University of Strathclyde and a Ph.D in Human Computer Interaction from University of Nottingham. He works at the boundaries of information technology law (mainly privacy & information security), HCI, and computer ethics. His research mainly considers the technical, legal, social, and ethical implications of living with ambient interactive computing.

Juan Carlos Valencia is an Associate Professor of Communication at Universidad Javeriana in Bogotá, Colombia. He holds a degree in Electronics Engineering and a Masters in Communication from Universidad Javeriana, and a PhD from the Department of Music, Media and Cultural Studies at Macquarie University (Australia). He is a member of the Research Group 'Comunicación, medios y cultura' at Universidad Javeriana.

Karin van Es is an Assistant Professor at Utrecht University and coordinator of the Datafied Society research platform. Her current research interests revolve around tool criticism and public values in the digital age. Karin is author of the book *The Future of Live* (Polity Press 2016), co-editor of the volume *The Datafied Society* (AUP 2017) and the issue 'Big Data Histories' (2018) for TMG-Journal for Media History. She has published in outlets such as Television & New Media, Media, Culture & Society, M/C Journal and First Monday. Her publications cover critical data studies, social media, and the concept liveness.

Daniela van Geenen is a Lecturer and researcher in data journalistic practice and data visualization at the University of Applied Sciences Utrecht and an affiliated researcher at the Datafied Society research platform. Her work tackles the question of the scholarly conduct that the work with digital methods demands, challenged by the need to design accountable software tools. She wrote her MA thesis on tool criticism investigating the case of Gephi. Daniela published on the role of social and technical actors on social media platforms, and their meaning for social and political practices such as public debate and cultural consumption.

Jelmer van Nuss is a developer at the Digital Humanities Lab at Utrecht University and a computing science master's student at Utrecht University. He specializes in artificial intelligence with a particular interest in data science. His work at the Digital Humanities Lab consists of supporting humanities researchers by providing software and tools for their research.

Lonneke van der Velden is a postdoctoral researcher with DATACTIVE and a Senior Lecturer at the Department of Media Studies at the University of Amsterdam. Her work deals with surveillance and data activism, with a special focus on Open Source Intelligence. She is part of the editorial board of *Krisis: Journal for contemporary philosophy in the Netherlands* and is on the Board of Directors of the Dutch digital rights organisation Bits of Freedom.

Jennifer Walker (Haudenosaunee), PhD, is a Canada Research Chair in Indigenous Health and Assistant Professor, School of Rural and Northern Health, Laurentian University. She

is the Indigenous Lead at the Institute for Clinical Evaluative Sciences in Ontario, Canada. Jennifer works to support and advance the governance and use of population health data by Indigenous nations to address community health and wellbeing. Her program of research integrates Indigenous perspectives on multi-morbidity and culturally safe care for Indigenous older adults. Jennifer is a member of the Indigenous Data Sovereignty Interest group at the Research Data Alliance.

Jeb Webb is a research fellow at Australia's Oceania Cyber Security Centre, which works in partnership with the University of Oxford's Global Cyber Security Capacity Centre to conduct maturity reviews for countries in the Oceania region. He received a Bachelor of Arts in Political Science from the University of California at Berkeley, a Master of Arts in Intelligence Studies from American Military University, and a PhD from the University of Melbourne's School of Engineering (Department of Computing and Information Systems).

Maranke Wieringa is a PhD candidate at Utrecht University and has a background in Cultural Studies and Media Studies, with a specialization in software and data. Her dissertation investigates algorithmic accountability in Dutch municipalities, and, following from this insight, Maranke develops a toolkit to empower municipal data professionals to give testimony of the (decisions around the) algorithm. At Utrecht University, Maranke is part of the Datafied Society research platform, and teaches various courses on (scholarly) data analysis. Her academic interests lie at the intersection of software and data (e.g. tool criticism, algorithm studies).

Kersti Wissenbach (kerstiwissenbach.com) is a researcher and senior consultant working on the crossroads of communication, governance, responsible data, and civic tech. She is specialised in participatory methods and has worked with activist groups, NGOs, and government institutions in over 15 countries. Since 2011 Kersti runs her own consulting firm providing strategy and policy advice and facilitating co-creation processes. Kersti lectures in the Communication for Development Master of Malmö University. She is a researcher with the DATACTIVE project where she merges social movement and communication for social change scholarship for her study of power dynamics within transnational civic tech activism.

Sharon Young has recently submitted her PhD thesis to the University of New South Wales, investigating the influence of tariffs on the impacts of decentralised energy resources on the Australian electricity industry and is currently working with the Centre for Energy and Environmental Markets (CEEM).

Andrea Zeffiro is an Assistant Professor in the Department of Communication Studies and Multimedia, and Academic Director for the Lewis & Ruth Sherman Centre for Digital Scholarship at McMaster University.

DATA

Theory on Demand #29

Good Data

Moving away from the strong body of critique of pervasive ‘bad data’ practices by both governments and private actors in the globalized digital economy, this book aims to paint an alternative, more optimistic but still pragmatic picture of the datafied future. The authors examine and propose ‘good data’ practices, values and principles from an interdisciplinary, international perspective. From ideas of data sovereignty and justice, to manifestos for change and calls for activism, this collection opens a multifaceted conversation on the kinds of futures we want to see, and presents concrete steps on how we can start realizing good data in practice.

Angela Daly is a transnational and critical socio-legal scholar of the regulation of new technologies. She is currently based in the Chinese University of Hong Kong Faculty of Law and holds adjunct positions at Queensland University of Technology Faculty of Law (Australia) and the Tilburg Institute of Law, Technology and Society (Netherlands).

S. Kate Devitt is a philosopher and cognitive scientist working as a social and ethical robotics researcher for the Australian Defence Science and Technology Group. She is an Adjunct Fellow in the Co-Innovation Group, School of Information Technology and Electrical Engineering, University of Queensland. Her research includes: the ethics of data, barriers to the adoption of technologies, the trustworthiness of autonomous systems and philosophically designed tools for evidence-based, collective decision making.

Monique Mann is the Vice Chancellor’s Research Fellow in Technology and Regulation at the Faculty of Law, Queensland University of Technology. Dr Mann is advancing a program of socio- legal research on the intersecting topics of algorithmic justice, police technology, surveillance, and transnational online policing.

Printed on Demand
ISBN: 978-94-92302-28-1

Institute of
network cultures

DATA