

JULY 2015



DEMOCRATISING THE DATA REVOLUTION:

A DISCUSSION PAPER

JONATHAN GRAY

DEMOCRATISING THE DATA REVOLUTION

“New technologies are leading to an exponential increase in the volume and types of data available, creating unprecedented possibilities for informing and transforming society and protecting the environment. Governments, companies, researchers and citizen groups are in a ferment of experimentation, innovation and adaptation to the new world of data, a world in which data are bigger, faster and more detailed than ever before. This is the data revolution.” – UN Data Revolution Group, 2014

What will the “data revolution” do? What will it be about? What will it count? What kinds of risks and harms might it bring? Whom and what will it serve? And who will get to decide?

This brief discussion paper is intended to advance thinking and action around civil society engagement with the data revolution. It looks beyond the disclosure of existing information, towards more ambitious and substantive forms of democratic engagement with data infrastructures.¹ It concludes with a series of questions for discussion about what practical steps institutions and civil society organisations might take to change what is measured and how, and how these measurements are put to work.

WHAT COUNTS?

How might civil society actors shape the data revolution? In particular, how might they go beyond the question of what data is disclosed towards looking at what is measured in the first place? To kickstart discussion around this topic, we will look at three kinds of intervention: changing existing forms of measurement, advocating new forms of measurement and undertaking new forms of measurement.

CHANGING EXISTING FORMS OF MEASUREMENT

Rather than just focusing on the transparency, disclosure and openness of public information, civil society groups can argue for **changing what is measured with existing data infrastructures**. One example of this is recent campaigning around company ownership in the UK. Advocacy groups wanted to unpick networks of corporate ownership and control in order to support their campaigning and investigations around tax avoidance, tax evasion and illicit financial flows.

While the UK company register recorded information about “nominal ownership”, it did not include information about so-called “beneficial ownership”, or who ultimately benefits from the ownership and control of companies. Campaigners undertook an extensive programme of activities to advocate for changes and extensions to existing data infrastructures – including through legislation, software, and administrative procedures.²

¹ In this context we understand data infrastructures as composites of technical, legal and social systems (e.g. software, laws, policies, practices, standards) involved in the creation and management of data.

² See: Gray, J. & Davies, T (2015) “Fighting Phantom Firms in the UK: From Opening Up Datasets to Reshaping Data Infrastructures?”. Working paper available at: <http://ssrn.com/abstract=2610937>

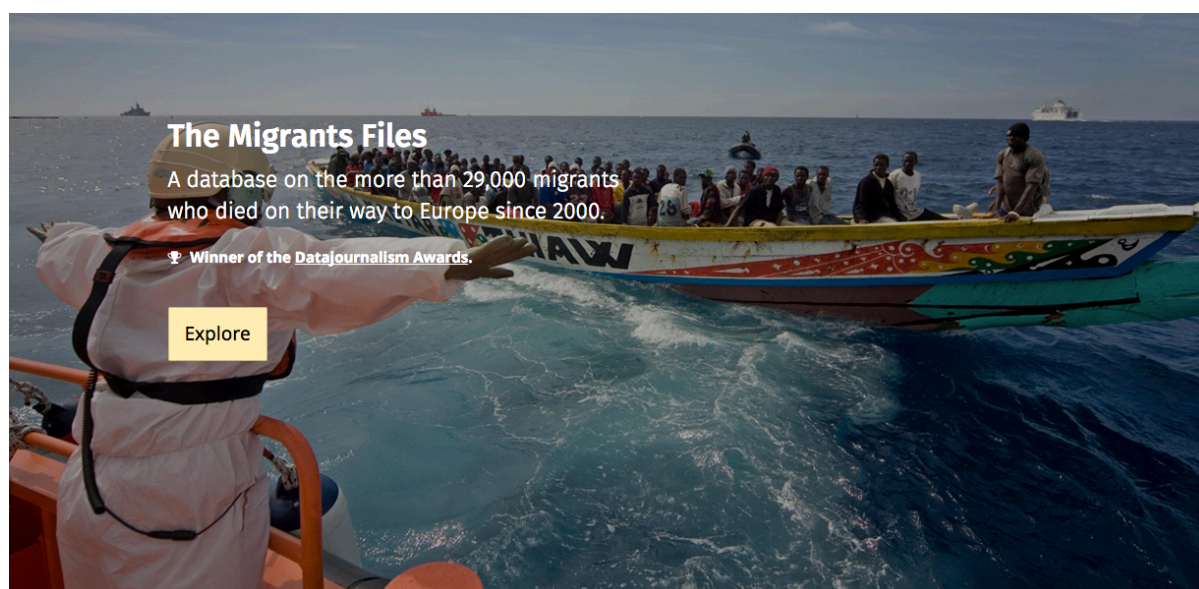
ADVOCATING NEW FORMS OF MEASUREMENT

As well as changing or recalibrating existing forms of measurement, campaigners and civil society organisations can make the case for **measuring things which were not previously measured**. For example, over the past several decades social and political campaigning has resulted in new indicators about many different issues – such as gender inequality, health, work, disability, pollution or education.³ In such cases activists aimed to establish a given indicator as important and relevant for public institutions, decision makers, and broader publics – in order to, for example, inform policy development or resource allocation.

UNDERTAKING NEW FORMS OF MEASUREMENT

Historically, many civil society organisations and advocacy groups have **collected their own data** to make the case for action on issues that they work on – from human rights abuses to endangered species.

Recently there have been several data journalism projects which highlight gaps in what is officially counted. The Migrant Files is an open database containing information about over 29,000 people who died on their way to Europe since 2000, collated from publicly available sources.⁴ It was created by a network of journalists who were concerned that this data was not being systematically collected by European institutions. In a similar vein The Counted project from *The Guardian* records information about deaths in police custody in the US, explicitly in response to the lack of official data collection on this topic.⁵



³ See: Bruno, I., Didier, E., and Vitale, T. (eds) (2014) Statistics and Activism. Special issue of *Partecipazione e conflitto*. *The Open Journal of Sociopolitical Studies*. Available at: <http://siba-ese.unisalento.it/index.php/paco/issue/view/1248>

⁴ See: <https://www.detective.io/detective/the-migrants-files/>

⁵ See: <http://www.theguardian.com/us-news/series/counted-us-police-killings>

THE ROLE OF THE OPEN DATA MOVEMENT

The nascent open data movement has often focused on the release of pre-existing information about things which are already routinely measured by public institutions. Advocates have pushed for the release of datasets under open licenses in machine-readable formats to facilitate widespread re-use – whether to develop new applications and services, or to facilitate new forms of journalism and advocacy. Datasets are often published via data portals, of which there are now hundreds around the world at local, regional, national and supranational levels.⁶

As well as opening up new datasets, some public institutions have implemented mechanisms to gather input and feedback on open data release priorities, such as:

- **Advisory panels and user groups** – e.g. as the UK's Open Data User Group (ODUG);⁷
- **Dedicated staff** – e.g. public engagement or “Chief Data Officer” roles;
- **User engagement channels** – e.g. social media accounts, forums and mailing lists;
- **Data request mechanisms** – e.g. Data.gov.uk's dataset request service or the EU Open Data Portal's “Suggest a Dataset” form;⁸
- **Consultation processes** – e.g. Open Government Partnership National Action Plans;
- **Solicitation for input around data standards** – e.g. the US's Federal Spending Transparency issue tracker on GitHub.⁹

In principle these kinds of mechanisms could be used not just to inform priorities for the release of existing datasets – but also in order to facilitate engagement between institutions and civil society actors around what should be measured by the public sector and how.

To use a metaphor, if data can be compared to photography, then might the open data movement play a role in intervening not just around access and circulation of snapshots taken by public institutions, but also around what is depicted and how it is shot?

⁶ See: <http://dataportals.org/>

⁷ See: <https://www.gov.uk/government/groups/open-data-user-group>

⁸ See: <http://data.gov.uk/data-request> and <https://open-data.europa.eu/en/suggestDataSet>

⁹ See: <https://fedspendingtransparency.github.io/>

QUESTIONS FOR DISCUSSION

We would like to catalyse discussion and gather input about how to increase civil society engagement around the data revolution and questions about **what** should be measured and **how**. To this end, we invite advocacy groups, journalists, public institutions, data users, researchers and others to respond to the following questions.

WHAT CAN CIVIL SOCIETY GROUPS DO?

- What can civil society organisations do to engage with the data revolution?
- What role might the nascent open data movement play in mediating between civil society organisations and public institutions around what should be measured?
- What opportunities does the data revolution present for civil society organisations?
- What are the best examples of democratic interventions to change, advocate or create new forms of measurement (both present and past)?
- What are the biggest obstacles to greater civil society engagement with the data revolution? How might these be addressed?
- Which kinds of transnational challenges and issues (e.g. climate change, tax base erosion) are currently inadequately dealt with by national data infrastructures?
- What areas might new kinds of measurement make the biggest difference, and how?
- What factors are most important in ensuring that data leads to action?
- What might civil society groups do to flag potential risks and unwanted consequences of data infrastructures as well as their benefits?

WHAT CAN PUBLIC INSTITUTIONS DO?

- What can public institutions do to better understand the interests and priorities of civil society organisations around what should be measured?
- Are there examples of where open data initiatives have facilitated significant changes to existing datasets, or the creation of new kinds of datasets?
- Which kinds of mechanisms might be most effective in understanding and responding to the interests of civil society organisations around what is measured and how?
- What are the biggest obstacles to public institutions responding more effectively to the data needs and interests of civil society groups? How might these be addressed?

HOW TO RESPOND

We welcome responses on these and other topics via the channels below:

- Join the discussion at: <http://bit.ly/ourdatarev-forum>
- Fill in the form at: <http://bit.ly/ourdatarev>
- Tag your responses with the [#ourdatarev](#) hashtag on Twitter

REFERENCES AND FURTHER READING

- Gray, J. & Davies, T (2015). Fighting Phantom Firms in the UK: From Opening Up Datasets to Reshaping Data Infrastructures?. Available at: <http://ssrn.com/abstract=2610937>
- Bruno, I., Didier, E. & Vitale, T. (2014). Statactivism: forms of action between disclosure and affirmation. In *Partecipazione e conflitto. The Open Journal of Sociopolitical Studies*, Vol. 7, No. 2: 198-220. Available at: <http://ssrn.com/abstract=2466882>
- Bruno, I., Didier, E., and Vitale, T. (eds) (2014) Statistics and Activism. Special issue of *Partecipazione e conflitto. The Open Journal of Sociopolitical Studies*. Available at: <http://siba-ese.unisalento.it/index.php/paco/issue/view/1248>
- UN Data Revolution Group (2014). A World That Counts: Mobilising the Data Revolution for Sustainable Development. Available at: <http://www.undatarevolution.org>

© Open Knowledge, 2015.
St John's Innovation Centre
Cowley Road
Cambridge
CB4 0WS
United Kingdom
Cover Design: Sam Smith

This work is licensed under a
Creative Commons Attribution
4.0 International License

