



# Open data study

New technologies

Becky Hogge

For more information contact:

## Transparency & Accountability Initiative

c/o Open Society Foundation  
4th floor, Cambridge House  
100 Cambridge Grove  
London, W6 0LE, UK

Tel: +44 (0)20 7031 0200

[www.transparency-initiative.org](http://www.transparency-initiative.org)



Copyright 2010. Creative Commons License.  
This work is licensed under a Creative Commons Attribution  
3.0 Licence: <http://creativecommons.org/licenses/by/3.0/us/>

# Contents

Introduction	4
The UK – data.gov.uk	5
The US – data.gov	7
The three-tiered approach	9
Civil Society	10
The push from the middle	12
Top-level drive	15
Utility not Users – the absent driver?	17
International perspectives	18
The relationship between FOI and Open Data	19
Questions over impact	21
Data characteristics	21
The three-tiered approach: top level drivers	25
The three-tiered approach: the middle layer	27
The three-tiered approach: civil society	29
Donors: a fourth tier?	30
Issues for developing and middle income country users	32
Complementary strategies for open data	33
An aside: advice for funders	34
Avenues for further research	35
An open data strategy checklist	36
Status of FOI	37
Current data collection activities	37
Potential end-users	37
Tier 1: Civil society	37
Tier 2: The middle layer	37
Tier 3: The top-level	38
Tier 4: Donor drivers	38
Conclusion	39
Annexes	41
Annexes I: Methodology	42
Annexes II: List of interviewees	43
Bibliography	44
Acknowledgements	45
About the author	45

# Introduction

There are substantial social and economic gains to be made from opening government data to the public. The combination of geographic, budget, demographic, services, education and other data, publicly available in an open format on the web, promises to improve services as well as create future economic growth.

This approach has been recently pioneered by governments in the United States and the United Kingdom (with the launch of two web portals – [www.data.gov](http://www.data.gov) and [www.data.gov.uk](http://www.data.gov.uk) respectively) inspired in part by applications developed by grassroots civil society organisations (CSOs) ranging from maps of bicycle accidents to sites breaking down how and where tax money is spent. In the UK, the [data.gov.uk](http://data.gov.uk) initiative was spearheaded by Tim Berners-Lee, the inventor of the World Wide Web.

This research, commissioned by a consortium of funders and NGOs under the umbrella of the Transparency and Accountability Initiative, seeks to explore the feasibility of applying this approach to open data in relevant middle income and developing countries. Its aim is to identify the strategies used in the US and UK contexts with a view to building a set of criteria to guide the selection of pilot countries, which in turn suggests a template strategy to open government data.

The report finds that in both the US and UK, a three-tiered drive was at play. The three groups of actors who were crucial to the projects' success were:

- Civil society, and in particular a small and motivated group of 'civic hackers';<sup>1</sup>
- An engaged and well-resourced 'middle layer' of skilled government bureaucrats; and
- A top-level mandate, motivated by either an outside force (in the case of the UK) or a refreshed political administration hungry for change (in the US).

As Tim Berners-Lee observed in interview, 'It has to start at the top, it has to start in the middle and it has to start at the bottom.' The conclusion to this report strengthens that assertion, and warns those attempting to mirror the successes of the UK and US projects not to neglect any of these three layers of influence.

Based on these findings, and on interviews conducted with a selection of domain and region experts to refine these observations for a developing and middle-income country context (where a fourth tier of potential drivers towards open data – in the shape of international aid donors – is identified) the report presents a list of criteria to be considered when selecting a pilot country in order to test this strategy. This checklist includes questions on:

- the status of Freedom of Information law and data collection activities in the country;
- the capacity of civil society, the media and other end-users to access and make use of the data;
- the attitude of high and mid-level political operators and government officials to opening up data;
- the role of multilateral and bilateral donors.

---

<sup>1</sup> For a definition of 'civic hackers', see the 'Civil society' section of this report.

The UK - [data.co.uk](https://data.co.uk)

data.gov.uk is a UK government web portal providing access to data gathered and maintained by UK government and related public agencies.

In June 2009 the then Prime Minister, Gordon Brown, announced that Tim Berners-Lee – the man credited with inventing the World Wide Web – was joining the Government as an expert advisor.<sup>2</sup> Since 1999, Berners-Lee has been developing the idea of a ‘semantic web’ based on linked data facilitated by metadata standards. In March 2009, Berners-Lee had made a call at the influential TED conference for data-gathering bodies around the world to release ‘Raw Data Now!’<sup>3</sup>

For several years previous to Brown’s announcement, the UK Government had been under pressure from civil society – and in particular a vocal group of ‘civic hackers’ responsible for the development of grassroots political engagement websites – to release its data openly and in machine-readable formats. This pressure – sometimes exerted through traditional advocacy, but also exerted through demonstrator projects – had led to internal policy reviews and initiatives such as the Power of Information Review (co-written by mySociety’s Director Tom Steinberg) and the Power of Information Taskforce (championed by then Cabinet Office minister Tom Watson MP and led by Richard Allan, a former Member of Parliament and at the time a policy advisor to Cisco systems).

These initiatives succeeded in fostering collaboration between civil servants in the middle layer of government administration and civic hackers of many different stripes, for example in regular GovDataBarCamps, and through competitions such as ‘Show Us a Better Way’, which attracted around 500 entries, including the budget tracking website WhereDoesMyMoneyGo?<sup>4</sup> However, there was strong resistance from a small number of Government agencies, and in particular the Government’s cartographic mapping agency, the Ordnance Survey, which was funded from revenues from the commercial exploitation of public data (so-called ‘Trading Funds’). The Power of Information Review had recommended a detailed cost-benefit analysis of the Trading Fund model, which was commissioned jointly by HM Treasury and the then Department for Business, Enterprise and Regulatory Reform, and published in February 2008.<sup>5</sup> Despite that study finding strongly in favour of abandoning the Trading Fund model<sup>6</sup>, and despite the Power of Information Taskforce recommending reform on the back of its findings, positions against re-examining the Trading Fund model within government remained entrenched.

This situation eventually changed in early 2010. data.gov.uk was launched in beta (invitation only) on 30 September 2009, with an official, open launch in January 2010. On 1 April 2010, the Ordnance Survey announced that it would make crucial data assets openly available through data.gov.uk<sup>7</sup>. It is unclear whether the change of heart on Trading Funds derived more from mounting evidence – including internal Government development of a database to cater to the UK census – or from simple political will. Nonetheless, the announcement was seen as a victory for the open data community, and the UK is now witnessing a flourishing of postcode-related campaigning and political engagement sites, (in particular of sites geared towards the May 2010 General Election).

As of April 2010, data.gov.uk listed 3,241 datasets and hosted 49 derived applications (although many more applications which make use of the data – and particularly the geospatial data – are hosted elsewhere on the web). Departments are not compelled to publish data via the portal; rather, it is presented as an aspirational choice for government departments wishing to open their data to the possibilities of data mash-ups and the semantic web.

<sup>2</sup> Brown, 2009.

<sup>3</sup> Berners-Lee, 2009.

<sup>4</sup> See: <http://www.wheredoesmymoneygo.org/>

<sup>5</sup> Newbury, Bently and Pollock, 2008.

<sup>6</sup> The study found that in most cases, although making data available without charge would create additional expenses for government, the benefits to society of moving to a marginal-cost regime – effectively zero cost for digital data – outweighed the costs of short-term ‘cost-recovery’ models.

<sup>7</sup> BBC News, 2010.

# The US - data.gov

data.gov is a US government web portal providing the public with access to federal government-created datasets. It was launched in 2009, both to allow 'citizen feedback and new ideas' – enabling transparency, participation and collaboration between state and citizen – and to increase efficiency among government agencies. Most US government agencies already work to codified information dissemination requirements, and data.gov is conceived as a tool to aid their mission delivery.

Like the UK, the US also exhibited a small and influential community of 'civic hackers', working independently and most often without financial reward, to repurpose government-created datasets and present them online in more accessible and enriched formats. For example, in 2004 a student called Josh Tauberer launched GovTrack.us, which repurposes publicly available data about key activities of the US Congress and publishes it in an accessible, searchable form. Perhaps in response to this activity, the Chief Information Officers (CIOs) in key states in the US – most notably the District of Columbia – began to release state-level datasets through data portals. It was the CTO for the District of Columbia, Vivek Kundra, who was hired by Barack Obama in March 2009 to act as the Federal CIO.

One of the first memorandums signed by President Obama when he entered the White House was on openness and transparency. As a Senator, Obama's headline achievements included the Coburn–Obama Transparency Act, which established USAspending.gov, a search engine on federal spending, showing that the future President 'got' the possibilities early on. Following a period of research and development, data.gov was launched on 21 May 2009. It initially contained 76 datasets from 11 government agencies. By late April 2010 that number had risen to 1,284 datasets from 170 government and related public agencies.<sup>8</sup> Fears that not enough data was being channelled through data.gov led to Obama issuing a decree on 8 December 2009 that each government agency should post at least three high-value datasets by a given date.<sup>9</sup> Indeed, after data.gov.uk launched in January 2010, US commentators began comparing US efforts unfavourably, complaining that the UK site had three times as many useful datasets as the US site, even though it lagged six months behind.<sup>10</sup>

The data.gov website is a curated web portal sitting on top of three separate, searchable data catalogues ('Raw Data Catalog', 'Tool Catalog' and 'Geodata Catalog'),<sup>11</sup> each of whose entries include a metadata template based on the Dublin Core standard (Dublin Core Metadata Element Set, Version 1.1). Although data.gov does not host derived applications on its site, it does record the number of times datasets have been downloaded: for example, in the week beginning 19 April 2010, it reported that datasets across all three catalogues had been downloaded over 25,000 times in the previous week.

Unlike data.gov.uk, data.gov does not generally employ data formats associated with semantic web development. However, it does include integration with the emerging semantic web as one of its future goals, and is in the process of commissioning pilot projects around semantic web development.

<sup>8</sup> For the purposes of comparing the UK and US efforts, only datasets that fall within the 'Raw Data Catalog' in data.gov have been counted (see below).

<sup>9</sup> Executive Office of the President, Office of Management and Budget, 2009.

<sup>10</sup> Kirkpatrick, 2010.

<sup>11</sup> As noted above, for the purposes of comparing the UK and US efforts, only datasets that fall within the 'Raw Data Catalog' have been counted.

# The three-tiered approach

*'It has to start at the top, it has to start in the middle and it has to start at the bottom.'*

Tim Berners-Lee

At least three constituencies have been involved in the push for data.gov and data.gov.uk. These constituencies are civil society; the 'middle layer' of professional public administrators; and top-level political leaders. It is arguable that without each one of these constituencies driving progress, we would not be where we are today.

Civil society, as will be elaborated below, was of a very particular kind, and provided bottom-up pressure for change through traditional advocacy, but also through demonstrator projects such as TheyWorkForYou.com (UK) and GovTrack.us (US). The 'middle layer' of civil servants and state and federal administrators were also motivated to open government data, spotting opportunities for efficiency and for greater recognition of the work they did. Without their expertise, top-level political leaders could not have been confident that plans for data portals could realistically be realised. Finally, political leaders could squeeze out institutional inertia, and give what looked like highly technical projects some democratic grounding.

## Civil Society

*'It's not like mySociety were the only people reusing data, but we were virtually the only people reusing it in a sphere that meant that politicians and policy people paid any attention to who we were.'*

Tom Steinberg

TheyWorkForYou.com was launched in the UK in June 2004, having been developed by a loosely-affiliated group of volunteers 'who thought it should be really easy for people to keep tabs on their elected MPs, and their unelected Peers, and comment on what goes on in Parliament.'<sup>12</sup> Between them, the original volunteers had already developed and run a number of civic engagement and other related websites, including PublicWhip.org.uk (a Parliamentary vote tracking website) and AccessibleTrainTimes.com (an accessible version of the National Rail Enquiries website).

As with previous sites created by this group, TheyWorkForYou.com repurposed data already being published online (in this case, via the Hansard website), presenting it in a format that was accessible and searchable, and allowed for user feedback. The site launched despite the fact that the data it was reusing constituted a copyright infringement in the UK: the volunteers did not have the right to reproduce Hansard, which was covered by Crown Copyright. Later on, and in cooperation with some of the TheyWorkForYou.com volunteers, click-use licences were

developed at the Office for Public Sector Information (OPSI) which among other things legitimised the site's activities. In 2006, by mutual agreement, the running of TheyWorkForYou.com was passed to mySociety, because 'the overlap of people and goals with mySociety was so substantial that it was best to pass the running over to mySociety, who are an actual proper organisation with staff and time'.<sup>13</sup>

GovTrack.us was launched in the US in September 2004 by Josh Tauberer, a linguistics postgraduate and software developer. Like TheyWorkForYou.com, it repurposes official data (this time produced by the Library of Congress) and publishes it in an accessible, searchable form. Because of the way the Library of Congress makes that data available, GovTrack.us needs to use automated techniques (for example, pdf-scraping and subsequent information parsing techniques). The datastream this activity creates means that GovTrack.us also acts as a data hub for other civic engagement websites. GovTrack.us lists at least 22 such projects on its website.<sup>14</sup>

In 2009, Josh Tauberer incorporated Civic Impulse LLC, and GovTrack.us became a project of this entity. Advertisements that run on the website provide the company with a small profit, but according to Tauberer 'the point has never been to make this a business'.<sup>15</sup>

Both the volunteers around TheyWorkForYou.com and Tauberer identify themselves as 'civic hackers'. Although the word 'hacking' has, in the context of computers, evolved in meaning so as to attract broadly negative connotations among the general public, among software engineers 'hacking' simply means the experimental development of software and systems to solve particular problems. 'Civic hacking' is therefore understood to mean deploying information technology tools to enrich civic life, or to solve particular problems of a civic nature, such as democratic engagement. Civic hackers might be described as tending away from characteristics associated with other types of more traditional civil society: for example, they might privilege approaching only those civic problems to which they can see an elegant (technical) solution. For more advice on how to spot civic hackers, see the section below on 'Advice for funders'.

<sup>12</sup> mySociety.

<sup>13</sup> Ibid.

<sup>14</sup> Civic Impulse LLC 2.

<sup>15</sup> Civic Impulse LLC 1.

It was the early interventions of civic hackers which arguably served to kick off the public debate around open data, and led to the emergence in 2009 and 2010 of the open data catalogues data.gov and data.gov.uk. As one observer put it:

*'I would say that they're the source of this debate. Things like data.gov are being created to meet the new demand and the new role that civic hackers present. I just think of Josh Tauberer, who's the University of Pennsylvania graduate student that built GovTrack.us, which is one of the first big sites that was taking on a huge scraping problem, taking Congressional Bills data and scraping it and parsing it and releasing it in XML and RDF, even in 2005 or 2006, much earlier than anyone else was even thinking about RDF. That site really empowered an ecosystem of other sites that use congressional data.'*

#### John Wonderlich, Policy Director, Sunlight Foundation

One question that arises from remarks such as these is whether policy makers got interested in civic hacking because of the obvious utility of sites such as GovTrack.us given the meagre resources needed to produce them, or whether instead it was the widespread adoption of these sites among members of the general public. Several observations indicate it was the former, not the latter, which swung policy makers; for example:

*'(Josh Tauberer) operated almost completely independently and was able to provide a more valuable service than what was coming out of the entire Library of Congress and he was one person who was just a linguistics student'*

#### John Wonderlich

This theme is picked up later in the report.<sup>16</sup>

The incorporation of both these individual projects points to a gradual consolidation of the grassroots movement around open data issues. The two most visible grassroots players in the UK and US open data space are mySociety and the Sunlight Foundation respectively, but this does not mean that open data issues were not being discussed in other forums. The licensing issues around data reuse in the UK (these issues appear to be absent from the US discussion, mainly because of historic laws that proscribed public data and databases in general from attracting copyright protection) led in 2006 to the launch of the Free Our Data campaign,<sup>17</sup> run by *Guardian* newspaper technology editor Charles Arthur. The Open Knowledge Foundation (OKF) was also a key player in convening a range of grassroots and institutional players from across national borders around open data issues:

*'The Open Knowledge Foundation has also been involved, from quite an early phase, in organising events which have brought developers wanting to do civic web services using open Government data, together with Government representatives. John Sheridan from OPSI has come to OKF events and been very supportive since at least 2005 ... Some of the discussions that happened at those kinds of meetings have proved to have been very valuable in the longer term. So for example in autumn 2008 we discussed the idea of a registry of open Government data for the UK, and at that time, data.gov didn't exist yet – people were looking at, for example, things like Vivek Kundra's ... data catalogue<sup>18</sup> [for inspiration].'*

#### Jonathan Gray, Community Coordinator, OKF

data.gov.uk would eventually be built using the CKAN infrastructure,<sup>19</sup> a project developed by the OKF after it became aware of the value of open data registries 'which aims to be a lightweight open source piece of software for helping people to find and reuse open data'.<sup>20</sup>

In interviews, representatives from both mySociety and the OKF stressed how they felt that the culture of their communities contributed positively to getting buy-in from government at an early stage.

*'I think it helped that the data agenda wasn't just and solely associated with anti-politics, basically.'*

#### Tom Steinberg

*'Partly what's made the OKF work is an attempt to keep things quite simple, because I guess in both of those communities grassroots and institutional, there might be a temptation to go off into quite technical language.'*

#### Jonathan Gray

*'There's a sense in which the OKF has been able to mediate between those two groups, but also I would suggest that the OKF has always been quite neutral, the OKF has never been a campaigning organisation.'*

#### Jonathan Gray

One interviewee was keen to stress, however, that the cooperative approach might not translate to all political settings: 'in some countries it's entirely the logical thing to do, to represent yourself as a threat, because there's no other way of doing it'.<sup>21</sup>

<sup>16</sup> See section entitled 'Utility not users: absent drivers?'

<sup>17</sup> <http://www.freeourdata.org.uk/>

<sup>18</sup> The interviewee is referring to the data catalog Kundra pioneered as CTO of the District of Columbia, and not data.gov, which would emerge later.

<sup>19</sup> Where CKAN stands for 'Comprehensive Knowledge Archive Network'. CKAN is a registry of open data and content packages; for more details, visit <http://www.ckan.net/>

<sup>20</sup> Interview with Jonathan Gray.

<sup>21</sup> Interview with Tom Steinberg.

One striking and apparent difference between the UK and US experiences is that in the UK both mySociety and OKF, were in some sense co-opted by the institutional ‘middle layer’ of government as the trajectory towards data.gov.uk progressed: the OKF, to provide the software that would eventually run the data portal; and mySociety in the form of Tom Steinberg’s appointment in 2007 by the Cabinet Office to undertake the Power of Information Review jointly with Ed Mayo, Director of the National Consumer Council (NCC, now Consumer Focus).<sup>22</sup> It is to developments in this layer that this report now turns.

## The push from the middle

*‘I think actually you can put a lot of the success down to very serious competence of UK civil servants - both their understanding of the workings of government and their understanding of the technology.’*

### Tim Berners-Lee

It is in the contribution of the middle layer of Government where the trajectories of data.gov and data.gov.uk appear to diverge. This could well be a function of this report’s methodology: the researcher was unable to conduct interviews with any public administrator directly involved in the launch of either platform, and literature that would give insight into the middle layer experience is necessarily sparse. Whereas civil society respondents from both sides of the Atlantic painted both positive and negative pictures of the attitudes of public servants towards open data issues, the highest praise was reserved for the UK middle layer (see the quote above, which was echoed by other participants). This leads the researcher to speculate whether – on top of the early outreach undertaken by organisations such as the OKF – it was in fact the contentious issue of data licensing that gradually brought together allies from the grassroots and public administrator communities, building a stronger base of expertise and shared goals than perhaps was exhibited in the US. The different characteristics of the top-level pressures might also be a factor, and this is explored in greater detail below (see ‘Top level drive’).

In 2007, together with the director of the NCC, Ed Mayo, Tom Steinberg was commissioned by the Cabinet Office to conduct the Power of Information Review.<sup>23</sup> The review made 15 recommendations, many of which had to do with the future role of open public data. Tom Steinberg identifies two major opportunities offered by the Power of Information Review. The first was to act as a ‘cheerleader’ for opening up government activities to wider civil society:

*‘The Power of Information Review was a combination of cheerleading and one or two really meaningful policies. The cheerleading is because you need to set a tone around agendas in Whitehall, and probably governments generally, that this is of the moment, that this is worth doing, that this is not frightening – and this is in any field, whether it’s technology or agriculture or just whatever. Well, a way in which you get things done is to talk about it as if this is the thing that we’re doing right now, and the reason you have to do that is because some of the people who hold the decision-making power in any such public institution will never understand what you’re pitching. So as well as coming up with good arguments, you have to come up with a more general non-specific emotional vibe that helps people who don’t really understand the arguments, but who for some reason have been given powerful positions. It helps them get over the risk-aversion. So the emotional tone is really important.’*

### Tom Steinberg

The second was to instigate an evidence-based review of the Ordnance Survey’s Trading Fund model:

*‘Getting the Treasury to do an official economic review of the charging model for Trading Funds is the other most important thing to come out of the Power of Information Review.’*

### Tom Steinberg

The Government response to the review was broadly positive, accepting each recommendation outright, in part, or in theory.<sup>24</sup> The Office of Fair Trading had already produced a review of the Trading Fund model in December 2006,<sup>25</sup> but Steinberg remarks that this report ‘was important as a source of reference, but not as a source of authority or power’.

More useful in getting the Government to accept the recommendation and commission a review were personal contacts within the Treasury.<sup>26</sup> Steinberg admits that this scenario is not necessarily replicable, agreeing with the interviewer that it was more or less a function of ‘networks of young people who know about technology, who then go on into different spheres of life, but stay in touch.’<sup>27</sup> However, Steinberg goes on to remark that interventions made by Tim Berners-Lee subsequent to the Power of Information Review may render the need for such personal contacts obsolete (‘he’s not young, he doesn’t know anyone in government, but he’s famous enough that the Prime Minister would pay him attention.’<sup>28</sup>). Indeed, although the report found strongly against the Trading Fund model in terms of the overall economic value it delivered, it was not until Tim Berners-Lee came on board at the Cabinet Office that the report’s findings were seriously acted upon.<sup>29</sup>

<sup>22</sup> Mayo and Steinberg, 2007.

<sup>23</sup> Ibid.

<sup>24</sup> Cabinet Office, 2007.

<sup>25</sup> Office of Fair Trading, 2006.

<sup>26</sup> Newbury, Bently and Pollock, 2008.

<sup>27</sup> Interviewer’s own remarks in conversation with Tom Steinberg.

<sup>28</sup> Interview with Tom Steinberg.

<sup>29</sup> Crabtree, 2010. This sentiment was reflected by the report’s co-author Rufus Pollock in a recent presentation on open government data at the British Council event ‘Copyright’s 300th birthday: What next?’, attended by the researcher.

The other major outcome of the review was the founding of the Power of Information Taskforce, convened by then Cabinet Office minister Tom Watson MP in March 2008, and chaired by early TheyWorkForYou contributor and former MP Richard Allan. Grassroots and industry representatives<sup>30</sup> were joined on the Taskforce by various members of the public sector, from Tom Loosemore (then at OfCom, now head of Channel 4's own response to the web 2.0 agenda, 4IP) to Andrew Stott, Alex Allan (both at the Cabinet Office, and the former mentioned again and again in interviews as a changemaker within the civil service – 'Andrew wears a tie but actually hacks as well', says Tim Berners-Lee) and Natalie Ceeney (National Archives). The secretariat was drawn from the Cabinet Office (William Perrin and Alex Butler) and John Sheridan (OPSI).<sup>31</sup> It is perhaps important to note that the early cross-cultural base of the OKF open data events was repeated in the make-up of the Taskforce.

One of the most visible and striking outcomes of the work of the Power of Information Taskforce was its 'Show Us A Better Way' competition,<sup>32</sup> a prize fund of £20,000 for the best proposals for reuse of public sector information. The competition attracted around 500 entries, and was won jointly by 14 projects.<sup>33</sup> The Taskforce's final report notes that:

*'The Taskforce has been impressed by the extent to which access to geospatial data has been a recurrent theme during its activities. For example, the Show us a Better Way competition had around 500 entries and of these over one third were for ideas around maps and location.'*

#### Power of Information Taskforce Report<sup>34</sup>

This observation, together with mounting evidence – for example, from the Treasury study<sup>35</sup> led the Taskforce to give detailed recommendations around the freeing up of geospatial data and reform of the Ordnance Survey. In total, the Taskforce made 25 recommendations, including the need for a well-resourced public service information regulator, in the form of a radically better-funded OPSI

CITATION Gra09 \l 205.<sup>36</sup> It was after the publication of the Taskforce's final report that Gordon Brown first met with Tim Berners-Lee on the subject of open data,<sup>37</sup> though it is unclear who made the introduction. In June 2009, as part of a statement on constitutional reform made in response to the MPs expenses crisis, Gordon Brown announced the appointment of Berners-Lee as a Government advisor.<sup>38</sup> A Cabinet Office press release<sup>39</sup> confirmed that Berners-Lee would be working with Nigel Shadbolt, Professor of Artificial Intelligence at Southampton University – both men had met with Brown at 10 Downing Street earlier that year.<sup>40</sup>

Tim Berners-Lee describes his time as a Government advisor as being characterised by a series of meetings, including meetings with civil servants producing data. During those meetings, he encountered a high level of enthusiasm at the middle layer:

*'There were people who'd actually got some datasets that they were very happy to give out, they just hadn't been given the okay to do it.'*

#### Tim Berners-Lee

Where people Berners-Lee encountered were unsure of the open data strategy, he introduced them to mash-up projects that were already transforming public sector information:

*'I think when they saw just a few examples of how information is reused by a third party who has never been involved with the process before, that's when they realised the value of the information that they've been producing is really untapped - it's not been exploited. Then they get very keen to be a part of that infrastructure, to be part of putting the data out there. And those two spheres, the data sphere, as I call it sometimes, and the mashup sphere are separate.'*

#### Tim Berners-Lee

This distinction made by Berners-Lee between the 'data sphere' – populated by government departments and other institutions producing data – and the 'mash-up sphere' – populated by civil society and private actors – is a crucial one, as it points to a fundamental shift in strategies around transparency. The distinction is also highlighted by Robinson *et al*, in their paper 'Government Data and the Invisible Hand', which makes a case for third parties repurposing government data along the following lines:

*'The biggest advantage of third party data processing is to encourage the emergence of more advanced features, beyond the delivery of data... Exactly which of these features to use in which case, and how to combine advanced features with data presentation, is an open question. Private parties might not get it right the first time, but we believe they will explore more approaches and will recover more rapidly than government will from the inevitable missteps. .. For those desiring to build interactive sites, the barriers to entry are remarkably low once government data is conveniently available.'*

#### Government Data and the Invisible Hand<sup>41</sup>

<sup>30</sup> The latter in the form of Richard Sargeant, formerly an official at HM Treasury but by that time working as a public policy advisor to Google.

<sup>31</sup> Power of Information Taskforce 3.

<sup>32</sup> Power of Information Taskforce 2.

<sup>33</sup> For the full list, see <http://www.showusabetterway.co.uk/call/2008/11/and-the-winners-are.html>

<sup>34</sup> Power of Information Taskforce 1, 2009.

<sup>35</sup> Newbury, Bentley and Pollock, 2008.

<sup>36</sup> Power of Information Taskforce 1, 2009.

<sup>37</sup> Crabtree, 2010.

<sup>38</sup> Brown, 2009.

<sup>39</sup> Cabinet Office, 2009.

<sup>40</sup> Crabtree, 2010.

<sup>41</sup> Robinson, Harlan, Zeller and Felten, 2009.

As observed above, it is the ease with which applications can be developed by the 'mashup sphere', and not the mass adoption of such applications, which provides the decisive argument for opening up government data by the 'data sphere'. This theme is picked up in subsequent sections of this report.<sup>42</sup>

Berners-Lee particularly applauds the decision of Cabinet Office civil servants to release data.gov.uk 'in beta' in September 2009 to a self-selected community of data mashers and other civic hackers before the official release in January 2010, a move he saw as crucial to the success of the official launch:

*'I think that was rather clever, the way they created this community, and that was done largely by the Cabinet Office putting together some open source software and skinning it with CSS files pretty promptly.'*

#### Tim Berners-Lee

It is tempting to draw parallels with this lightweight and inclusive technical strategy – which stands in stark contrast to the vast majority of Government-commissioned technology projects – and the strategies introduced to the Cabinet Office by the Power of Information Taskforce the previous year.

Many of the people interviewed in the UK praised the dedication of civil servants. This comment is typical:

*Really, the civil servants have been what has made this happen. They've understood what the benefits might be, they're really on it technically speaking, in terms of understanding the web, understanding some of these new internet technologies, understanding the community-driven nature of some of these projects. One can't emphasise enough how on it they've been.*

#### Jonathan Gray

And indeed, in the US, where the role of the 'middle layer' in the run-up to the launch of data.gov is less clear, stakeholders still praise the attitude of at least some public administrators:

*'It's a frequent occurrence for us to write about some database or piece of information and be ignored by the people who can make decisions about it, yet be silently and secretly appreciated - and contacted - by the people who actually administer the database. They're our biggest allies but they are hamstrung as far as their ability to advocate for improvement. It's not uncommon to have meetings near Congress or in an agency with someone who doesn't want their boss to know because they don't know how to appropriately advocate for an improvement or for a data release... So sometimes we'll put up a blogpost and six hours later have a really appreciative email from someone who's happy to see that they're being noticed... And they usually understand the value of their data better than anyone else.'*

#### John Wonderlich

Although it is questionable how far upwards the concept of a 'middle layer' of government extends, it is worth noting that Wonderlich identifies actors at the most senior level of the Secretary or Minister as being choke points in the release of open data on occasion:

*'Overcoming that challenge - part of it is buy-in at the secretary level. There are really exciting things happening, for example, at HHS, our Health and Human Services Agency. That's got a variety of new responsibilities with the new healthcare legislation that passed, so I think it's great that they're on the ball because they really have to be. If you look at the Department of Defense, however, the Open Government Plan they released yesterday was really, for the most part, boiler plate language and plans to make plans about other plans and to set up a new commission that will report to another commission. I don't know if that's characteristic of the Defense Department or the people who are involved or if it's the secretary's fault. I'm not sure how to assign the blame.'*

#### John Wonderlich

Wonderlich goes on to describe the attitudes he encounters from public servants unwilling to open up data:

*'Excuses you hear are "if we release the data, you'd find out how bad it was, so it's not that useful anyway," and "if we release too much data, then all the data becomes less useful" I think they're really just cover for insecurity about relinquishing control because when you have a dataset that only you can access, then you have a certain power that comes along with that. And when you release it, you have to take responsibility for the privacy problems or the errors and everything else.'*

#### John Wonderlich

<sup>42</sup> See section entitled 'Utility or users: absent drivers?'

One interviewee who had had direct experience of working inside public administration was Nathaniel Heller, now Managing Director of Global Integrity, who spent a short time from 2002 as part of the US State Department. He characterises his experiences there as follows:

*'I think the challenge is not so much in a way a cultural friction or the government being offended or annoyed by civil society, it's that all of the incentives within a bureaucracy – at least the government bureaucracy – tend to be aligned and oriented around the rest of the bureaucracy and not the public. So, for instance, I mean this was a weird era because the interagency system in the US was so broken during the Bush administration, but we spent 98% of our time basically fighting with the Pentagon over various issues and trying to limit the damage. That was our whole universe, it was centred around dealing with interagency process, trying to promote the department vision as opposed to somebody else's. The engagement, talking with the public on these things, just never happened, partly because I was in a policy portfolio that was very superior-oriented and fairly classified - it's not like you would go out and talk about these things on a daily basis. But the point was that everything I did, my success as a bureaucrat, my promotion, my pay raises, were tied almost entirely to how successful I was working within the bureaucracy. It was not at all tied to the public or to public policy outcomes and I think that's the challenge.'*

**Nathaniel Heller, Managing Director, Global Integrity**

This frank and intriguing insight into the inner workings of government may offer clues to both the successes and the challenges of both data.gov and data.gov.uk.

## Top-level drive

*'The idea of open data as a thematic issue has been enormously raised by the Obama administration and the creation of data.gov'*

**John Wonderlich**

Whereas the trajectory towards data.gov.uk is a story of policy reform – especially with regards to the licensing of public data – the data.gov data portal faced no such obstacles. There was, however, a historic lack of political will to realise the value of public sector information:

*'Before Obama was elected, there were a bunch of structures that were supposed to do this, but didn't. So there was a clear lack of political will to take on this kind of issue. Perennially, we have laws that try to fix eGovernment or create more transparency or electronic access to information and they've often failed in the past, or they've had one or two successful components and the rest of it ends up being ignored. So I think at its heart, what's happening now is a cultural movement and it's a change in public expectations. The structures in Government are largely secondary to that.'*

**John Wonderlich**

Obama's campaign was characterised by the employment of new technology in radical new ways. But as Tom Steinberg points out, the technologies employed by the Obama campaign were of a different hue to those advocated by the open data community:

*'You probably want to try and ask, specifically, how did Andrew McLaughlin get his job, and how did [Beth Noveck] get her job? Within what was essentially a hyper-political campaign, who decided this was worth having? Given that it's very, very different from the technology stuff that got Obama elected – that was very high-tech, but it wasn't very open-data-y, or even open source. So that's an interesting question, and I don't know the answer.'*

**Tom Steinberg**

Absent from Steinberg's remarks, but surely key to the direction of data.gov, was Obama's hiring of Vivek Kundra as the United States Chief Information Officer in March 2009, just before the data.gov portal was announced. As CTO for the District of Columbia, Kundra was a figurehead in the open data community. As mentioned above, Kundra's data catalogue<sup>43</sup> had been the inspiration for early OKF-sponsored discussions of a data registry for the UK's public sector information.

<sup>43</sup> <http://data.octo.dc.gov/>

John Wonderlich goes some way towards answering Steinberg's question:

*'I don't know when Obama first started up a relationship with a lot of the Google officials, but it's a pretty deep one: he gave speeches at the Google headquarters pretty early on in the campaign. And Andrew McLaughlin, for example, was one of the top Google people and he's now in the Office of Science and Technology Policy in the White House. But beyond that, I think there's a couple of factors. One is a generational sense, that he wanted to be viewed as a change agent, to be associated with technological savvy. And beyond that, throughout his Senate career, he also had a particular appreciation for the power of what it is to post information when the broader public has access to that information. So one of the Bills he was responsible for in the Senate created access to the first Federal database of Grants and Contracts, which is USAspending.gov, so there is that link pretty early.'*

**John Wonderlich**

However, Wonderlich notes that the long-term impact of Obama's clear vision for data.gov is yet to be established:

*'I think right now what we're waiting to see is to what degree the President's clear political will translates to actual change through civil servants and throughout the bureaucracy, and it's interesting to see, you know, some agencies, that political will can transfer rather quickly down through the ranks. In others, it gets caught right below the secretary and doesn't turn into very much, at least not yet.'*

**John Wonderlich**

If Obama's decision to back data.gov came as he rode a wave of political optimism, the contrast with Brown's position when he announced the appointment of Tim Berners-Lee could not be starker. An unelected Prime Minister<sup>44</sup> speaking in the shadow of one of the worst scandals to hit British public life in decades – over MPs' bogus expenses claims – Brown's was a very different sort of political moment. Tim Berners-Lee gives his impression of Brown's thinking:

*'I think he's excited about making the country more efficient. Transparency was a big thing, in the sense of allowing the British public to hold the British government to a high level of effectiveness. The way Gordon Brown talks about transparency, it's not about trying to find people cheating. It's about looking at the state of the nation and seeing how it changes and thinking about how it can be better and then providing feedback at every level.'*

**Tim Berners-Lee**

There is no doubt that the combination of a world-famous inventor, bestowed with the backing of the highest Minister in the land, was a catalyst to change across all Government departments, and particularly in the case of the still-contentious Ordnance Survey data.

*'So the sequence was to go and talk to people at a ministerial level and get them to urge the different departments to do this, and simultaneously to talk to developers, the people who had been working for a while on this sort of stuff in a relatively quiet way...if you look at the ministers, each minister had to basically command each department to say, "Yeah, we're going to do this".'*

**Tim Berners-Lee**

The success of Berners-Lee in bringing Ministers on board, in contrast to reports from the US experience that data.gov aspirations often got stuck at or just below Secretary level, leads the researcher to speculate on the strategic importance of the Prime Minister's chief advocate for the project being at one remove from traditional political power. Did the presence of Berners-Lee – famous and celebrated in a sphere completely removed from politics – encourage Ministers to tune down their political sensitivities, put aside personality politics and point-scoring, and just get on with it? It is tempting to believe so.

Finally, those who have followed the debate sound a note of caution to anyone wishing to suppose that it was Tim Berners-Lee alone who created the success of data.gov.uk:

*'The received picture in the press, certainly in the last year, has been one of Tim Berners-Lee coming in and revolutionising everything in Government. Which of course is, to a large extent, true in terms of the excitement and the catalytic potential of him coming in and making all of this work more high-profile. And also in terms of communicating it to the public, saying that the inventor of the World Wide Web is involved adds credence to something that could otherwise be viewed as quite obscure. Bringing Tim Berners-Lee on board at that time, I think, was crucial, in terms of the political value that's been obtained. That said, recent developments in the last year or so with data.gov.uk have cast a much longer historical shadow, going back to the Show Us A Better Way competition, but before that too. The fantastic work that OPSI have been doing for a long time, which have their roots in, for example, recommendations made in studies such as the Cost Cutting Review, back in the late nineties or early 2000s.'*

**Jonathan Gray**

<sup>44</sup> Brown took over the position when he became Labour Party leader after Tony Blair resigned in 2007. His election as party leader was unopposed and automatically bestowed on him the

role of Prime Minister, and he declined to call a General Election afterwards, as might have been expected.

## Utility not users – the absent driver?

Was end-user take-up a key driver in the journey towards data.gov and data.gov.uk? Certainly, the potential for broadening civic engagement of applications developed by the small communities of civic hackers that congregated around, say, GovTrack.us in the US, or MySociety in the UK, came up frequently in interview. But more often, it was the utility of applications (in contrast with the resources expended to produce them) and not their broad user bases, which seem to have inspired officials further up the line to engage with the open data agenda.

In the Power of Information Review – which covered many aspects of the interface between Government and Web 2.0 beyond open, linked data – Steinberg and Mayo highlight the existence of large communities of ordinary British citizens congregating around user-generated information websites, such as the parenting website Netmums (described in the report as having 275,000 users) and MoneySavingExpert (2.5 million users).

*‘The proportion of people using such sites to help themselves and others is now on a par with the friendly societies and mutuals of the nineteenth century’*

Power of Information Review<sup>45</sup>

And yet where user-generated information websites were promoted on the basis of existing large user communities, websites which repurposed government information were promoted by advocates on the basis of the utility they provided when compared with their low cost; on the basis of their potential for wide and beneficial usage, rather than on existing usage patterns. This is despite the fact that one case study of TheyWorkForYou.com shows that it attracts between 250,000 and 500,000 users each month<sup>46</sup> – a not inconsiderable number – while a separate study shows that another mySociety civic engagement tool, WriteToThem.com, had around 100,000 users in 2008.<sup>47</sup>

A discussion of barriers to broad use of civic engagement websites in developing and middle-income country settings takes place in the following sections of this report. Among other topics, this discussion highlights the roles of intermediaries, such as traditional civil society groups, in taking information from open data-driven web applications and disseminating it across populations less ‘wired’ than those found in the US and UK. Clearly, to ensure real impact, a potential user base should be a prerequisite when it comes to deploying such initiatives elsewhere. But from a strategic standpoint, it is worth noting that the existence of this user base was not a key part of the rhetoric used to talk up the open data idea. Rather, it was utility, and not users, which won the arguments in the UK and US.

<sup>45</sup> Mayo and Steinberg, 2007.

<sup>46</sup> Tactical Technology Collective, 2009.

<sup>47</sup> Escher, 2009.

# International perspectives

*'The UK structure and the way this happened in the UK is totally unlikely to work in the developing world. And I think we should just let go of it.'*

**Ethan Zuckerman**

*'This may be one of those things where putting a certain amount of effort in now will have a disproportionate effect.'*

**Tim Berners-Lee**

The researcher spoke to a number of regional and domain experts (see list of interviewees) in order to understand how the strategies employed to realise data.gov and data.gov.uk might translate to a middle-income or developing country setting, and to uncover additional factors that might come into play should such initiatives be attempted. Interviewees ranged from experts on the introduction of freedom of information (FOI) laws and computerised administrative systems around the world, to experts in fiscal transparency and budgetary monitoring, to civil society practitioners in India, Kenya, Tanzania, South Africa and the central and eastern European (CEE) region.

A high level of scepticism was encountered on a number of issues. Questions were raised around whether privileging open data initiatives over information access rights was a desirable strategy in countries where such rights were not yet enshrined in legislation. Interviewees also questioned whether datasets equivalent to those being released on data.gov and data.gov.uk existed in some of the countries discussed. Interviewees also questioned the impact of data.gov and data.gov.uk, which remained, for them, unproven.

## The relationship between FOI and open data

One interviewee, who had spent considerable chunks of his career advocating globally for FOI laws, felt that any move away from this reform programme towards open data-type initiatives would represent an unfortunate change of priorities.

*'I think that for those countries which don't have a law yet ... and these are not mutually exclusive by any means, but certainly getting a law is a major priority and that if there are limited resources available for the country, I think for me that would be a greater priority.'*

**Toby Mendel, Executive Director  
Center for Law and Democracy**

Such scepticism was not limited to the 'old guard'. One interviewee at the centre of the open data movement wished to make the following comment on an unattributed basis:

*'There is a gap between initiatives that are based on governments giving out things that they want to give out, and governments creating rights that mean that they give things out all the time that they maybe don't want to give out. The Freedom of Information Act is a right, it gives people a right, whereas these data initiatives in America and Britain tend to be not rights, but more like gifts. And that doesn't make these data websites a bad thing, not at all – it just makes them not as good as rights. Because there are two problems: firstly, they're not necessarily as strong; and secondly, they assume to some degree that the government knows what the public wants. And in some cases that's really obvious, because [campaigners] are shouting all the time that the public wants maps. But when it turns out that actually the really valuable data exists in something that only one person knows about, and if they could get hold of it, then they could make a billion pounds, if they don't have a right to get that data, and they don't have any leverage, then we're not likely to see that value added.'*

**Anonymous**

Nonetheless, a countering viewpoint was expressed by an activist working in Kenya, who felt the need to call time on expensive advocacy efforts in favour of FOI, given the limited success of such efforts:

*'For a long time the focus was on getting the Freedom of Information Act passed. Let's do that first and then once that law has passed, then we'll have the data. And what you see is that I think there's, even [after all the] money and advocacy, there's still only two countries in Africa that have Freedom of Information Acts that have gone through. In Uganda they passed the Bill but they didn't operationalise it.'*

**Ory Okolloh, co-founder Mzalendo  
co-founder, Ushahidi (Kenya)**

Notwithstanding this contrast in viewpoints, there were common threads to discussions on the relationship between FOI efforts and open data initiatives. For example, the idea that open data catalogues might capacitate bureaucracies<sup>48</sup> in those countries which did have FOI laws, but whose public servants struggled to meet both the proactive and reactive responsibilities enshrined in those laws, proved to have some traction:

*'If you look at the right to information legally, there are two major sides to it: one is proactive disclosure, so all of these laws or most of these laws place an obligation on governments to push out information... And then there is a request-driven side of it. And what you find is that governments in many, many countries - I would say most countries outside the highly-developed country zone - have a great struggle meeting their proactive obligations. I think we need to recognise the very important role of capacitating the public sector in this area and not working with civil society initiatives: it almost sometimes may even undermine the government... [But] I think that trying to link these kinds of initiatives to promoting efficiencies and perhaps even building back capacity in government would be a good antidote to some of the problems that I'm highlighting.'*

**Toby Mendel**

A story of **bureaucracies struggling to meet disclosure responsibilities** shifting to a proactive publication agenda was told in the context of India:

*'So after this law was enacted, thousands of requests have reached Government officers and Government officers are snowed under by these transparency requests. So some of them are trying to deal with this deluge of requests by proactively publishing datasets on their websites.'*

**Sunil Abraham, Executive Director  
Centre for Internet and Society (India)**

A broader context of an emerging democracy struggling under the weight of including all stakeholders in public life was related in Tanzania:

*'Tanzania's been going through reforms for many years now, and... there were all these committees created and a very elaborate and rather cumbersome machinery set up to have everybody involved in the process. And I think now there's a disillusionment with that and the emphasis has partly shifted to look at [whether], if everything was open then different people in their own time could get on with analysing and using the data as and when needed.'*

**Rakesh Rajani, Founder  
Twaweza (Tanzania)**

<sup>48</sup> Robinson, Harlan, Zeller and Felten, 2009.

## Questions over impact

Open data catalogues as a valid initiative to pursue in a setting outside of established, higher-income democracies was not the only concept to attract scepticism. Interviewees questioned the **impact** of the US and UK's own data portals:

*'We all think the techniques are kind of sexy and exciting, but does it matter ... what's the impact?... The corruption drivers in the US, particularly around money and politics, are completely unsolved by greater amounts of money and politics information being put out there... I think the political economy context in which one is working matters hugely. There's going to be a limit at which point it doesn't matter anymore - this information is getting out there, everybody knows the problem, everybody sees it, we don't need more databases, more mashups, more maps - there is something else missing, which is some nasty combination of a lack of leadership and the very real political economy choke point that simply prevents these kinds of reforms from being possible.'*

**Nathaniel Heller**

*'I think the whole open data thing is a bit... I'm not saying it's emperor's new clothes, but it's verging on that... I believe in the idea that stuff should be open and I believe that in principle if you make it open then... people will use it for socially positive ends as well as possibly [the complete opposite]. But... the thing that annoys me about it, isn't so much the people who are doing it, but [that] people think that's it, that's the job done. You know... the idea that if we make the data open there is this layer of civically savvy hackers who will then do the necessary job. And I don't think that's the case basically. Because the necessary job isn't TheyWorkForYou or even WhatDoTheyKnow. That's not reaching out into communities or engaging with... social agendas. And it could do, open data really could do [that]. I think it could be part of a picture of civic activation, but it isn't at the moment and I think it's not because what's happening is a bad thing, but because it's only 50% of the picture. And the worry I have is that that 50% of the picture is being portrayed as 100% of the picture.'*

**Dan McQuillan, Social Innovation Camps (Central and Eastern Europe)**

Although the researcher has a lot of sympathy with such viewpoints, she notes that it might be too early to measure the impact of open data portals on wider social and political issues.

## Data characteristics

**Data availability** in developing and middle income countries was also flagged as an issue by global experts.

*'The idea that this stuff exists, is in digital format of some kind and is anywhere close to standardised is probably just an impossible leap ... I just don't think it's there.'*

**Nathaniel Heller**

*'My simplest example for this would be years ago, talking with the government in Senegal and trying to plan an intervention based on electronic property records and... the Senegalese government was at first very enthusiastic. And then we started talking about the physical challenge of it and what we ended up discovering is that before we built an electronic property records system we actually had to build a property records system. It wasn't clear that that data existed in paper form and that to build the sort of government data transparency system we needed, in many cases we would have to do the basic data collection.'*

**Ethan Zuckerman, Senior Researcher Berkman Center for Internet and Society**

*'Kenya is probably one of the most open societies in sub-Saharan Africa, but it is still a very serious challenge to get that sort of data, and I think... it's a challenge because the data doesn't exist and, in some cases, it is a challenge because people understand just how important that data actually is and what it could be used for.'*

**Ethan Zuckerman**

However, regional civil society actors pointed to various sources of data that they used in their everyday practice. As might be expected, availability, timeliness and digital accessibility varied widely:

*'Broadly speaking there are three kinds of datasets that come to mind. One is governance or administrative data that is collected usually year on year from schools, health centres and so forth. Then there is survey data..., and there the key actor would be the National Bureau of Statistics, whereas for the administrative data the key actors would be the Ministries. And then the third one is the data related to ad hoc studies as well as audits that are... often donor-driven reports such as expenditure tracking surveys. Now there's push to try to see if all three can be available.'*

**Rakesh Rajani (Tanzania)**

'Apart from these types of data which is the law and GIS and census and infrastructural data, [for] all the other Department-related data, availability in digital format varies from state to state. There's no standardisation, very little standardisation.'

**Sunil Abraham (India)**

'The Auditor General whose reports we do rely quite heavily upon for the purposes of reflecting on service delivery in prior years, that information is available in PDF format'

**Jay Kruise, Head of Monitoring and Research Programme, Centre for Social Accountability, South Africa**

'All the work we do [with Mzalendo] is manual, so we have to literally cut and paste information if we can find it. It's gotten a lot better from when we started. Now things like the Hansard are on the website pretty much in soft copy and up-to-date. When we started you couldn't get the Hansard at all online. So it's improved but it's still either in a PDF or in a Word document that you can't crawl or extract information from.'

**Ory Okolloh (Kenya)**

Vivek Ramkumar of the International Budget Partnership provides the following data<sup>49</sup> with regards to the availability of government documents in 85 countries that the organisation regularly scrutinises as part of its monitoring efforts. (See table below).

He made the following remarks about trends towards online availability across the world:

'[For] almost every government that we assessed in the previous round of our survey that was published in 2008, the Ministries of Finance have functional websites. So it's increasingly catching on. I think there are a couple of countries where they didn't have that, so it's the exception where information is [not] available online... [but] in many countries what we've noticed they do is that they make PDF versions of these documents available. In some cases they are scanning the documents and making them available in a PDF version after it's been scanned. And so that obviously is indicative that they're not really working on these documents necessarily online or at least they're making a choice here to make it available by scanning it rather than making the original electronic document available.'

**Vivek Ramkumar, Manager, Open Budget Initiative**

KEY BUDGET DOCUMENTS	NO. OF COUNTRIES MAKING AVAILABLE ONLINE	NO. OF COUNTRIES MAKING AVAILABLE ON REQUEST	NO. OF COUNTRIES PRODUCING THE INFORMATION BUT NOT PUBLISHING IT	NO. OF COUNTRIES NOT PRODUCING THE INFORMATION
Pre-budget statement	27	3	29	26
Executive's budget proposal	49	13	23	0
Citizens budget	13	4	1	67
Enacted budget	68	13	4	0
In-year reports	63	4	13	5
Mid-year review	18	4	21	42
Year-end report	50	14	14	7
Audit report	50	8	21	6

<sup>49</sup> International Budget Partnership.

Several interviewees agreed that centralised information on the current publishing practices of governments – either in the context of proactive requirements enshrined in FOI law, or in a looser context – was lacking. A call was made to map the availability of datasets (see ‘Avenues for further research’ section):

*‘I think it would be great to start mapping what datasets exist within governments, but I’m going to stand by my scepticism; I think a lot of the data that you want ... data is not an easy thing to generate ... and, you know, I’ve spent a whole lot of time in a whole lot of African ministries and a lot of the critical records are still on paper ... it’s not clear to me that those records are getting digitised or are getting digitised in a meaningful way ... I don’t think it’s as simple as just taking a database and publishing it in many cases.’*

#### Ethan Zuckerman

It is worth noting that many interviewees spoke of **personal data**, such as census data, in the same context as non-personal data. Indeed, in interview, Tim Berners-Lee appeared to exclude the idea that government data portals might encompass personally identifiable information only in the first six months of the project:

*‘the rule of that first six months was to get the low-hanging fruit, to show that online data was valuable but to do it without attempting anything which was questionable, like not going anywhere near personally identifiable information.’*

#### Tim Berners-Lee

The researcher notes that funders who wish to foster public goods through the better provision of digital information should consider the dramatic implications for the right to privacy of datasets containing personally identifiable information being released to the public.

Other issues around the character of government-maintained data were raised in interviews. The **specification of standard formats for data publication** was one such issue. The format in which a government releases data can have a positive or negative impact on that data’s reuse by third parties, particularly if the government chooses to release data in proprietary, as opposed to open format. The following guidance is taken from a techno-legal manual on open data in the context of international aid, but applies universally to open data projects:

*‘Proprietary, non-proprietary and open file formats: On the one hand some file formats are non-proprietary and open, which means they can be used or implemented by anyone with little or no restriction. Prominent examples include HTML/XHTML, OpenDocument, PDF, TXT, XML7.*

*On the other hand some file formats are proprietary, which means that there may be restrictions on how the format may be used, and certain software packages may be required to read the files. Prominent examples include MPEG Audio Layer 3 (MP3), Windows Media Video (WMV), Microsoft Word (DOC/DOCX) and Microsoft Excel (XLS/XLSX)8.*

*...If the file format specification is publicly available, then there is less risk that prospective re-users will be required to use a particular piece of software, or that, in the worst case scenario, the format will become obsolete and unreadable without software that is no longer supported.’*

#### Unlocking the potential of aid information<sup>50</sup>

<sup>50</sup> Gray, Hatcher, Hogge, Parrish and Pollock, 2009.

Whereas proprietary standards lead to monopolies and lock-in, open standards allow for open competition, and so are vital in encouraging vibrant competition in the ‘mash-up sphere’ of open data – both in the context of private sector application development, and civil society application development. However, the **dominance of commercial information technology vendors** around discussions of format standards, each one keen to have their format approved as the national standard and to reap the benefits of the monopoly conditions that would flow from such a decision, risks the development of national and international open data standards in some regions:

*‘The Open Standards Policy document is in [its third] draft and there have been two rounds of consultation over the last two years. There have been two rounds of consultation with the corporate sector and civil society, but there are several areas of contention. Some corporates representing proprietary interests want the Open Standards Policy to have space for patented technologies which might have royalty implications. So that’s the first point of contention and the second point of contention is that some corporate sector organisations believe that in a single domain, multiple standards should be allowed to operate, there shouldn’t be a mandate on a single standard. Over these two particular disagreements the policy has been delayed and even today it’s very difficult to say whether that policy would ever be fully published.’*

#### Sunil Abraham (India)

*‘If the donor and the ministry are not specifying data export standards, the idea that this data is open, what you’re going to get instead is someone coming in and [proposing] a proprietary system, heavy focus on security, a system where it is almost impossible to squeeze the data out of it in the end. I mean, I think where there might be an opportunity in all of this is thinking about how you construct a data standard and policy for donors... What I’m not sure about is how long it might take for international or local contractors to figure out how to respond to that and figure out how to build good bids around that.’*

#### Ethan Zuckerman

*‘Recently in Kenya there was a seminar on connected government. There was huge vendor representation there and I was following the conference on Twitter and all of them were pushing a solution – integrated, customised solution, but solutions that are internal to Government, nothing really about the public.’*

#### Ory Okolloh (Kenya)

**Commercial models of cost recovery** around public data – of the sort that caused so many problems for the open data movement in the UK – were also described in at least two of the regions covered by the expert interviews:

*‘Yeah, so GIS – geographic information – is available... But until very recently the Government always saw this as a revenue stream because certain private organisations were willing to pay for these datasets.’*

#### Sunil Abraham (India)

*‘Definitely with geographical information, I think maps, they sell maps or they sell that kind of data, so that might be an issue.’*

#### Ory Okolloh (Kenya)

A new type of commercial/cost-recovery model was also encountered, involving civil society practitioners in South Africa. The Parliamentary Monitoring Group (PMG)<sup>51</sup> is a consortium of NGOs in South Africa working to publish data about the activities of South African Parliamentary committees. The data it publishes is based on reports and recordings made by volunteer monitors it enlists from the general public and who attend the committee meetings in this capacity. Although PMG is funded by various grant-giving organisations (it lists current funders on its website as the Open Society Foundation, the Royal Danish Embassy and USAID), it cross-subsidises free access to its materials for South African civil society by charging commercial organisations a subscription fee:

*‘PMG will have minutes of meetings of committees. It’ll have audio recordings. It’ll have submissions by interested parties and business and political parties and it’ll have draft Bills and Gazettes which are being considered by Committees. But not everything is accessible....At this stage because we’re a non-profit we are not being charged, but I believe that there are certain charges that are applicable. But thankfully because we meet the requirements for exemption we don’t need to.’*

#### Jay Kruuse (South Africa)

<sup>51</sup> <http://www.pmg.org.za/>

The implications of this for open data portals are unclear, but the existence of PMG points to a different path being taken by civil society actors around government data than that seen in the UK and US. Those planning interventions around open data in South Africa should consider the impact of such interventions on existing civil society initiatives.

The issue of **data quality** was also raised:

*'There is a series of public interest disclosures that are expected from every single electoral candidate. However, when the candidates make this disclosure they usually send it in paper form, not in electronic form and what they do is they write their names differently. So it's very difficult for you to run [any] analysis without spending a lot of time... manually cleaning up the data.'*

#### Sunil Abraham (India)

Finally, a set of three questions to use when evaluating different countries' suitability for a data.gov-type intervention was proposed:

*"Does the data exist?" "Is it released and does it threaten someone?" ... and then, "Is it usable by journalists or somebody else?"*

#### Ethan Zuckerman

The researcher agrees that the existence of digitised data seems like a necessary prerequisite to the success of a project like data.gov, and recommends that further research be conducted to establish rates of administrative digitisation across the world. In subsequent sections of this report, the researcher goes on to examine who open data might threaten (in particular see the final paragraph of this section on 'high stakes' faced by some government bureaucrats), and who might be the user-end beneficiaries of open data projects.

## The three-tiered approach: top-level drivers

Given the clear emergence of a three-tiered approach in both the data.gov and data.gov.uk trajectories, interviews with domain and regional experts focussed on their experiences with each of these layers of society. Several issues around the top level were discussed, including how a character like Tim Berners-Lee might wield influence at this level, and how different political cultures might characterise the top level response to open data strategies in different regions.

A push from the **top level** was seen as a good strategy for overcoming wider inertia around opening up to citizens in Tanzania:

*'I think the larger constraint is Government is simply not in the habit of, it hasn't quite got to understanding that this is public data. As far as they're concerned this is their stuff and 'Why the hell should we give it to anybody?'. Now the way to get around that, if one got a sufficiently high level commitment from the President – and I wonder whether there's also a role here for Parliament and so forth and possibly legislation – that was on the one hand broad enough to not get stuck in some detailed problem that somebody might have a beef with, but sufficiently detailed so it doesn't just become a 'wishy-washy' commitment in principle, but has some teeth to it. So if one got that balance right then the technical piece can kick in, meaning the middle level can then come in and start doing their part provided there was this high-level commitment that was publicly made.'*

#### Rakesh Rajani (Tanzania)

The role of '**superstar**' power, such as that clearly fulfilled by Tim Berners-Lee in the UK experience, was generally greeted positively, with some caveats. Other potential top-level personalities were also mentioned:

*'[India] is a country where film stars get voted into power, so there is a bit of personality cult for sure. And that's why they're using Nandan Nilekani who's the former boss of Infosys to run the national identity project.'*

#### Sunil Abraham (India)

*'[Tim Berners-Lee] would not be viewed by the Government as someone like me or other lobbyists who have been working on this theme. They're like 'Obviously we know where she's coming from'. But that leverage and the fact that he could point to other examples I think it would make a huge difference. I've also been encouraging Google which engages with a lot of Governments to leverage their access and their conversations, even as a vendor. They could be pushing the Cloud or whatever it is that they sell to Governments. But, as part of that process, they could certainly influence the standards or the format of information that's going into the Cloud.'*

**Ory Okolloh (Kenya)**

*'Tim BernersLee, I doubt whether there are more than 50 people in the whole country who even know who he is.'*

**Rakesh Rajani (Tanzania)**

*'One needs to be careful that it doesn't sound like the bad old imperialist US. There are ways of getting it right and there are ways of botching it. But, done well, I think there's nothing that can match Obama's star power particularly in East Africa to get this done.'*

**Rakesh Rajani (Tanzania)**

Different trends in **political structures and cultures** were also highlighted. The impact of nepotism, and the 'big man' culture of African politics, were mentioned negatively in the context of capacitating the middle layer of government administrations:

*'It's highly centralised. There are high levels of control over what can be released and in many instances it's completely unnecessary but it's part and parcel of where we've come from. And I think the fact that we have a lot of people within Government who are appointed to a position in the Public Service on the basis of connections. We do have high levels of nepotism and patronage in the public sector and so people are very much aligned to the wishes of those who hold power and that results in a culture of secrecy or "I cannot do this without approval from on high". And the result is that information which really isn't sensitive and which should be distributed quite freely is caught up in the red tape. So we would need to change that climate.'*

**Jay Kruise (South Africa)**

*'Yes, absolutely, there is a culture of the big man, and the big man culture isn't just a governmental culture, it's also a commercial culture. The biggest problem we had doing software development work in sub-Saharan Africa from 1999 to 2004 was management, and it was basically how do you persuade an entrepreneur who runs his own company that he needs to treat his software developer as a precious resource because they could go elsewhere and, at a certain point, you actually ended up with companies essentially saying, "Well, I don't want my employees to be any smarter, that would be bad for me".'*

**Ethan Zuckerman**

*'It's still a big man culture in most African countries. So they would not take well to feeling like they've been left out of the loop. It's a protocol-type issue. So maybe you do [the two streams of advocacy] in parallel, but I'm not sure it would work as well ...if the Minister finds out people are working, the junior guys, without him even being aware. It's more of an ego management issue than anything else and I think that will be a bigger problem in African countries maybe than it was in the UK.'*

**Ory Okolloh (Kenya)**

## The three-tiered approach: the middle layer

In many ways, the importance of the **middle layer** of government bureaucrats in the success or failure of an open data-type initiative was as present in the experiences of interviewees as in the US and UK experience:

*'I think there would have had to be some baseline of support at the mid-levels of management in order to get this done'*

**Nathaniel Heller**

*'The idea that independent groups out there may want to analyse the data themselves and come up with perhaps different conclusions from what Government comes up with is not quite accepted.'*

**Rakesh Rajani (Tanzania)**

Messaging to change attitudes at this layer might also play out in a similar fashion:

*'Something else that I find useful is saying... "Hey you guys do so much good work" or "The perception out there in the public is that you guys sit around and do nothing. By putting this data out, it's in your interests".'*

**Ory Okolloh (Kenya)**

Question marks were raised over the capacity of this layer of government to meet the open data challenge:

*'Government bureaucrats, unlike British bureaucrats, they move around a lot. They don't have the domain expertise or they don't know computer science or information sciences'*

**Sunil Abraham (India)**

*'One of the things to think about is that in the computerisation of African government ministries, the computers usually go first to the minister and the deputy secretaries, and those aren't necessarily the people who are going to open things up.'*

**Ethan Zuckerman**

*'I think sometimes they just don't know any better or there's just not... someone making that compelling argument about why a PDF is not enough.'*

**Ory Okolloh (Kenya)**

But other interviewees saw capacity-building as a simple challenge to overcome:

*'I think the primary constraint is not going to be technical. So even where there is a technical gap, it could be easily filled.'*

**Rakesh Rajani (Tanzania)**

The journey towards greater **collaboration between civil society and the middle layer** of government seems to have already begun in some of the regions discussed:

*'I think the non-profit sector is also beginning to realise that it is not very useful to be in an antagonistic position with the Government. It would be best if the sector also took some responsibility for building the necessary data ecosystem.'*

**Sunil Abraham (India)**

*'Even those NGOs who have been working on electoral transparency have realised that they cannot do all the data cleaning themselves, that they have to work together. So I think there will be more and more working with the Government in the near future and various top-down and bottom-up projects that try and build consensus around data standards.'*

**Sunil Abraham (India)**

*'You need to change mindsets. I belong to a civil society network of access to information practitioners who are requesting information, but there is also the odd information officers' forum so people who are in Government who are appointed by law to consider requests for information, there are regular meetings of those forums to try and break down the history of secretiveness, defensiveness and unwillingness to provide information.'*

**Jay Kruise (South Africa)**

In others, there was still much ground to cover:

*'I think, what I haven't seen enough of is people engaging not even necessarily with the top officials, just with the guys who have to deal with this on a day-to-day basis. So a lot of Departments now have CIOs, CTO type guys. Those are the people, or the e-Government Directorate and not necessarily the Minister himself because he can just say 'Yeah okay, we agree. Thanks, bye.'"*

**Ory Okolloh (Kenya)**

And one region-expert interviewee saw the separation of civil society and government, when it came to new movements for social change that included technological innovation, in a fairly positive light:

*'I think the incumbents haven't woken up to [social innovation camps]. I don't know how much you've been to these places, but the incumbent situation is pretty hardcore. Power is a very live issue. And networks of influence are pervasive, and corruption is pandemic. I don't think there's any interest at all [at that layer of government]. Which is nice in a way because it means there's a strong contrast with this stuff. Yeah, there's none of this tokenistic doo-dab, so there's still a possibility to believe in its radical potential.'*

**Dan McQuillan, (Central and Eastern Europe)**

Finally, the **high stakes faced by bureaucrats** in the least developed parts of the world were highlighted:

*'In 2003, I was asked by the new USAID administrator in Kigali to go sit down with the head of the Rwandan Customs service and my task was to try to figure out why the Rwandan Customs service hadn't computerised. They'd gotten all the money from the US government, the system had been on the ground, installed for 18 months; no one was using it. And I spent four hours with the director of the Customs service and, for three hours, he basically told me it was hard, it was difficult, people didn't want to use it, there were problems with it. Finally, he got up, he closed the door to his office, and he said, "Look, I don't want to get killed and I don't want you to get killed" and then proceeded to explain to me that corruption in the Customs service was so rife within Rwanda that the computerisation of it would reveal who's stealing what and that there was so much money at play that he literally feared for his life if he used the system. So, it's important to realise that in some of these cases, it's not just something simple.'*

**Ethan Zuckerman**

## The three-tiered approach: civil society

Where the existence of civil society actors pushing for open data was remarked upon by interviewees, such groups were characterised as small, a situation consistent with that in the UK and US:

*'I think the specific push for data being transparent is driven again by a relatively small group of people. There's a broader push for Government to become transparent, but that isn't focused on data.'*

**Rakesh Rajani (Tanzania)**

*'The main constraint is that it all comes down to people and there are very few people who have the wherewithal to really do this within Tanzanian civil society'*

**Rakesh Rajani (Tanzania)**

*'So there is engagement, there are people who get it and there is certainly a group of developers who, if the data was open, perhaps could be incentivised in the way they've done in the US and UK with Apps for Democracy type competitions. But not necessarily on the advocacy.'*

**Ory Okolloh (Kenya)**

*'Ushahidi has gotten so much attention because it is one of the very few players in the space; it is one of the few really innovative ideas to come up.'*

**Ethan Zuckerman**

Even in the US context, the need to stimulate civil society to participate in open data projects was highlighted:

*'I think the bar has been set in many ways almost too high and there's an implicit expectation or assumption that if we just get it out there 'people' will figure out how to use it and automatically push a demand side agenda that's for the good of the public and that leap, to me, is really a big one. There has been very little evidence that that's ever been done in a meaningful way.'*

**Nathaniel Heller**

Although the ability to stimulate civil society was highlighted, it was in the context of general transparency advocates, and not the focussed, tech-minded actors witnessed in the US and UK contexts:

*'Getting movements going to push for openness is almost the easiest advocacy thing you can do ... I mean you go into a country, bring together ranges of potentially interested stakeholders, and you're almost certain to launch a campaign.'*

**Toby Mendel**

However, the attraction of new technology was described as something universal, leading the same interviewee to conclude that finding such people in any relatively technologically-developed country would not be a great barrier:

*'People just love to play around with this stuff, and that is a potent force to harness and it's not susceptible of those problems with parachuting in. Of course, local people, if it's uninteresting, won't do it. But I think what drives whoever puts that stuff on the UK and the US sites, I think that the similar kind of person will exist in most of the more technologically developed countries where this thing might be a goer'*

**Toby Mendel**

Certainly, practices of civil society actors in US and UK, in particular the hijacking of data outside of the permitted licensing terms but in the public interest, were already being mirrored in Kenya:

*'In my work I've always appropriated stuff that even we shouldn't really have access to, but our argument has been that it's public information and we have the right to it.'*

**Ory Okolloh (Kenya)**

## Donors: a fourth tier?

The centrality of international aid donors in much of the developing world was highlighted on a number of occasions during this round of interviews, leading the researcher to conclude that these actors might represent a 'fourth tier' in the drive towards opening up government data.

*'[Government] will grudgingly accept [that] donors [who] give a lot of money [can] raise questions, but they'll be less open to Tanzanians, particularly if it's done in an open space.'*

### Rakesh Rajani (Tanzania)

*'I think in many cases NGOs hold just as much data as the Government. And so I've always argued that similar efforts [should be directed at NGOs], especially in Africa where NGOs are so involved in health or water, or disease. You've seen the World Bank starting to release its datasets now, but they're sometimes just as bad as Government in terms of holding onto information.'*

### Ory Okolloh (Kenya)

*'I've argued for a long time that opening up data is really the way to go, both with Government and especially in Africa with NGOs as well or multilateral institutions like the World Bank. You'll probably find the World Bank has better stats on health than our Government does. And so making sure that they're not left out of the conversation [is important] as well.'*

### Ory Okolloh (Kenya)

One interviewee pointed out that getting donors to open their books was only one half of the problem – it would be necessary to identify where money had eventually ended up in order to create a useful system. Rather than look to national governments to provide this data, crowd-sourcing it from individual beneficiaries on the ground might be just as, if not more effective. Nonetheless, issues persisted:

*'You can't just look at how the donor hands the money out, you have to have some sort of confirmation from the other side as far as where the money ended up going. Because the real problem in many of these cases is you hand out a large sum of money and then it's possible that big percentages of it disappear before it hits any of the desired recipients.'*

### Ethan Zuckerman

*'So, one way you could do this is you could have the donor essentially disclose what their funding matrix looks like and then who their beneficiaries are and then, ideally, you go to the beneficiaries and say, 'Well look, how much of this did you get?' and then try to figure out how you resolve this back and forth. Then, what you're doing is building a very complex map of who the donors are funding and questions about how that money is actually getting to recipients. To the extent that that money is going into governments as far as governmental support of the situation, like Tanzania, I'm curious [whether] you really would be able to get the budgetary data that you're looking for. I think what you're going to find is that a lot of that money is fairly general support ... you know, it is money that goes to support and strengthen a particular agency.'*

### Ethan Zuckerman

On several occasions, interviewees observed that donors could tie their aid to the opening up of government data. Good arguments were made about how what looked like technocratic (and therefore fairly unthreatening) reforms on the surface, could lead to transformative outcomes:

*'The technocratic [reforms] are the only ones that the donors - at least the multilaterals and the bilaterals - can usually engage on. I mean, the World Bank being a very good example: by mandate, it can't touch politics or it can't touch the justice sector, for instance. Those are huge areas that are in desperate need, but they literally cannot fund anything. So, instead, they can do budget reform, they can do public financial management, they can do capacity building of the various civil servants, all that kind of stuff, which... starts to smuggle in a culture of accountability and transparency.'*

### Nathaniel Heller

Depending on a particular country's situation, donors could take one of two approaches:

*'What we have found is that, in a large number of countries, the lack of public information is due to two reasons. One is because the governments don't produce the information even for internal purposes. Now this is where there is lack of technical capacity or there's a lack of an institutional system that requires government to produce that particular kind of information on a routine basis. The role of donors there is to be able to... support public finance management systems, the development of infrastructure, providing technical assistance to governments in order to develop their capacities to start producing this information. So that's one set of countries. But in another set of countries, an almost equal number of countries, we find that the governments don't publish information, but this information is being produced for internal purposes and for their donors. So donors do have access to these reports, budget reports. In those countries the lack of public availability of these documents is due to political will, it's a political decision taken that "we are not going to be transparent" or "we're not going to publish this document" or "we've not done it in the past and it's not a priority for us at this moment". There donors have a different role, because we know they have access to these documents or could very easily get access to these documents if they wanted to. There we are encouraging donors to actually use their own influence over government to start encouraging them.'*

#### Vivek Ramkumar

Some countries, however, might resist such approaches on account of perceived hypocrisy:

*'The hypocrisy issue is not trivial ... even if it's blown out of proportion, I think ... the perception of it is very real in low-income countries. So for the World Bank to walk in and talk about the broader transparency agenda rings hollow, I think, for a number of governments and they say, 'We don't even know what's going on at the executive board level'*

#### Nathaniel Heller

And it was suggested that the strategic value of some aid recipients as allies (around, for example, security issues), might mean some donors **hesitate to enforce transparency conditions attached to aid**:

*'The chances that the US is going to put too much pressure on some of its key allies in this environment is highly unlikely. They're going to continue to give aid to certain countries irrespective of whether those countries are transparent and meeting certain benchmarks and that's the way with other countries too.'*

#### Vivek Ramkumar

Having donors tie aid could also lead to unintended consequences, either through poor execution of open data projects (which could be expected in the absence of a strong local movement able to guide and monitor open data initiatives made in response to donors) or through countries rejecting aid tied to opening up data, in favour of aid which is tied in a less burdensome fashion:

*'Yes, donor pressure has worked, but it has also backfired, I think ... it's something you have to be careful with. Probably the leading example of a backfire is Pakistan. Basically the Asian Development Bank imposed adoption of an FOI law as a loan conditionality, and Pakistan adopted an FOI law. But the bank didn't know what it was doing and was unable to assess the quality of that law, so it accepted it at face value. It was actually an ordinance passed by a military government, so it had no credibility among the people anyway, but also it was not a progressive rule and it's been ... I mean, not completely ... it's being very ineffective and, to some extent, I think it is ... I mean, this is arguable, but I think, to some extent, it has blocked further progress.'*

#### Toby Mendel

*'I can imagine some governments or some departments refusing that as a condition and looking for essentially less demanding aid, which is to say that if the US and the UK come in and say, "This is how we're tying our aid" and the Japanese say, "Well, we've got a very different tie ... we're actually not going to touch transparency, we're going to tie it to the use of Japanese contractors". That would be my opinion.'*

#### Ethan Zuckerman

*'I think we are still in such early days that it's hard to tell just how systematic or systemic the shift may be and, frankly, what the harm is or isn't. I think there are some negative externalities that would come out of that shift, but I think we may not know for a decade or two just how serious they are or aren't.'*

#### Nathaniel Heller

## Issues for developing and middle income country users

In terms of the widespread existence of **end-users** for applications derived by civil society or other actors from Government data, the researcher received mixed reports:

*'There are only 200 postings on Vote Report India of some kind of noncompliance or some kind of irregularity with particular voting booths. That's a very small number if you look at India as a country with one billion people.'*

**Sunil Abraham (India)**

*'Things are changing. In the last municipal elections that I went to vote in, outside where we had the polling booths there were young people with laptops and they were holding onto spreadsheets of the voter lists and they were verifying or helping voters find their election ID before they went into the polling booths. So things are slowly changing, but as far as the internet goes we are still a very long way off from the West.'*

**Sunil Abraham (India)**

*'A lot of us in this country were completely disempowered due to an apartheid state where you didn't question, you weren't entitled to ask, you were completely disenfranchised. And now we've gone towards a country which has these advanced laws but a lot of people remain illiterate, impoverished and disempowered to ask those questions that they're entitled to ask and obtain explanations from Government.'*

**Jay Kruuse (South Africa)**

The importance of **institutional users** of derived government data was highlighted:

*We've got a number of organisations that we work with who are beginning to realise the benefit of evidence-based monitoring and evidence-based advocacy, where you engage with assurances and promises or plans by Government to improve their services to communities.'*

**Jay Kruuse (South Africa)**

*'It's certainly a long road, speaking from my experience with Mzalendo.... Mzalendo's site will never reach more than a couple of hundred hits a day unless we do a huge media campaign. But then [we can produce] information in a format that if you're an advocacy group you can come to our website, print out really easily and use to do your advocacy work.'*

**Ory Okolloh (Kenya)**

The importance of **SMS and other mobile platforms** in the ecosystem of developing world communications was also remarked upon:

*'There are only 20 million broadband connections in India and only 71 million people claim to have used the internet last year. So I think using the internet to exercise your citizenship, that's quite a mature usage and in India the internet is still very, very recent, at least the PC and desktop version of the internet. Maybe the next generation of open data e-governance tools will have mobile interfaces.'*

**Sunil Abraham (India)**

The researcher was interested in how the cost structure of SMS, as opposed to internet communication, would impact on user takeup of open data applications designed for mobile. One interviewee had already been working in partnership with mobile providers to reduce costs at the supply end:

*'The commitment we have is really at a pilot level now. What we're trying to do with the mobile phone companies is to say to them 'Look, if you allow us to send SMSs to citizens at your off-peak hours ...' – so we would send them just to subscribers within their own network so they don't have to pay interconnection charges, in other words it doesn't cost them anything additional – 'So if you let us do that you do two things. One is that you reach citizens and you provide a service at virtually minimum cost to yourself, so you're just using your unused capacity. But secondly when citizens either forward those SMSs to others or respond with monitoring information, in fact what they're doing is they are paying normal rates, normal SMS rates and in this way you are generating revenue for yourself.'*

**Rakesh Rajani (Tanzania)**

Another interviewee, when challenged on cost, pointed out that engagement through SMS still represented value for money, when compared with other advocacy interventions:

*'If we wanted to send out SMSs also, there's potentially a cost implication, as you say, with that as well. But I always argue, I say "Compared to a lot of the other transparency projects I see [it's cheap, if you] really look at how much has been spent on advocacy around the Freedom of Information Bill. [And it] has not really gone anywhere".'*

**Ory Okolloh (Kenya)**

The pace of change on the network was also highlighted:

*In a matter of years, in a matter of three/four/five years we will get to a situation whereby we are not dealing with SMS anymore, we are dealing with internet via mobile phone in East Africa which I think will open up the possibility for citizen engagement remarkably.'*

**Rakesh Rajani (Tanzania)**

## Complementary strategies for open data

During the second round of interviews, a number of complementary strategies were suggested for pushing open data initiatives in developing and middle income countries. As observed in the UK and US contexts, Obama and Brown, in their endorsement of data.gov and data.gov.uk respectively, were operating in distinct political moments. Toby Mendel drew on his experience of promoting FOI laws around the world to describe a number of other political moments worth looking out for.

The first (which is arguably similar to Obama's 'moment') was the context of a fresh administration brought in on a popular mandate to replace a corrupt or otherwise politically dicey old regime:

*'Mexico – when the Fox administration came in – very much embraced transparency as a very important value and set up a fantastic system: great law and a great oversight institution. But one of the reasons they were interested in that was because all of the skeletons in the closet, as it were, didn't belong to them.'*

**Toby Mendel**

*'In 1974 there were very, very progressive amendments to the US FOIA, falling on the heels of the impeachment of Nixon'*

**Toby Mendel**

The second was against a background of the fear or reality of corruption, in countries focussed on **rapid economic development**:

*'China has a... sensible worry that their economic growth will be stunted by massive corruption and they see openness as an antidote to that. In Indonesia, too - the anti-corruption angle in Indonesia had a stronger practical precedent than in China because corruption was seen as having brought about the economic collapse that eventually led to Suharto's downfall.'*

**Toby Mendel**

Regional **'peer pressure'** could also play a role:

*'Of course, there is the whole process of developing international standards and that includes regional pressure. So the Philippines is about to adopt a law and Indonesia has just adopted one, and Thailand has had one for a while and... once the Philippines does, then Malaysia will start to say, 'Well, you know we're falling behind in the Asian region' so there's a larger political and peer pressure.'*

**Toby Mendel**

Finally, **sub-national laws** were also identified as important drivers:

*'Malaysia is a country which is very, very secretive. It has very, very draconian secrecy rules. It has very repressive media rules, it's a country where information is highly controlled, but it's not a tinpot dictatorship like Burma or North Korea or even Vietnam. And there, civil society groups working with opposition politicians have taken advantage of the loss to the governing party of, I think, five states in the last election, and they worked at the state levels, so the sub-national level. And actually that's been a pattern across the larger Asian countries. So if you look at India, Japan, China, Indonesia - all of those countries had sub-national laws first and then the countries started to develop some sub-national experience. Because a lot of this is about fear and the fear of the unknown... and I think a practical experience at the city level, at a provincial level, or a state level can really help with that.'*

**Toby Mendel**

In the context of forcing administrations to meet their proactive publishing requirements where an FOI regime is nominally in place, Jay Kruuse – whose Public Service Accountability Monitor organisation has in the past brought proceedings against government departments for failing to meet publishing requirements - identified the judiciary as an important ally:

*'The most positive signals that we see are... pronouncements by our judiciary on these issues. We've had some really landmark decisions in recent times which have vindicated these issues and pointed to the fact that democracy presupposes participation and dissemination of information so that people can make informed choices.'*

**Jay Kruuse (South Africa)**

Prizes were also highlighted as a means to encourage best practice in the middle layer of government:

*'The Human Rights Commission who have a mandate in terms of the current access to information legislation in South Africa, they look to hold a Golden Key Award every year to recognise information officers who have excelled and who provide guidance to other information officers. That has been one way of improving access.'*

**Jay Kruuse (South Africa)**

*'One thing I suggested to people who are trying to work on this within Kenya is to have an award for the most open Ministry or something, get them prizes for the best Information Department, also incentivise so there's something in it for them rather than them feeling that they're just being pressured to change.'*

**Ory Okolloh (Kenya)**

Finally, at least two interviewees recognised the importance of **individual political actors** in their domain spheres, and extrapolated to suggest strategies targeted at identifying and engaging such transformative individuals:

*'the other possibility is that you might start looking for ministers who embrace this. So, for instance, the former Nigerian finance minister, who was a fierce anti-corruption advocate... when she was in power she would have been an excellent minister to try this with. It's probably the sort of thing that if we're going to do it, we would want to make bets on a small number of ministers who want it and who get it.'*

**Ethan Zuckerman**

*'There are lots of pockets of where the Government is putting that information out there. It just needs to be pushed to the next step or to find a champion who gets it. The new Speaker of the Kenyan Parliament is much better than the previous one and so he's been okay with allowing TVs to broadcast... So again, it's not really an issue of secrecy or the Civil Service [not wanting to] change. It was just someone who got it. But now the next step becomes then if you find that person, educating them and saying 'Oh okay, it's not just about putting it out there but you can do X, Y and Z and this is why it's important.'*

**Ory Okolloh (Kenya)**

## An aside: advice for funders

During the second round of interviews, advice was offered directly to the funders who commissioned this report:

*'If [funders] want to make headway on these issues, they need to do two things. One is they need to be very clear about the overall end goal, a big-picture, strategic goal. They then have to be open to dealing with people who can make things happen and those people, they often are not your usual development sets. They might be quite different characters, often in the private sector. You might find that the best way to move this is to hire some really expensive private sector guy who's going to charge you \$50,000 for 20 days of work or 25 days of work, but that \$50,000 may end up getting you much farther than some long-term project where you sink in half a million dollars.'*

**Rakesh Rajani (Tanzania)**

Another interviewee backed up this assertion that 'civic hackers' might look fundamentally different to the sorts of civil society actors with whom funders are used to engaging:

*'[They're] possibly technocrats actually... Maybe [civic hackers] have a slightly cybernetic view of systems: they can work well, and [for civic hackers] value and good is about removing the obstacles to the smooth functioning of your cybernetic system. Whatever. I don't know. But it's very particular, it's great, it's inspiring, it's internationally inspiring and also real. But that's not the same as social impact.'*

**Dan McQuillan (Central and Eastern Europe)**

Common donor practices, such as avoiding project replication, were also highlighted as an issue:

*'But there's a systemic problem which is all donor agencies, including the ones you're preparing the report for, incentivise uniqueness and incentivise unique propositions. They don't incentivise replication of good ideas. They don't incentivise joining the bandwagon, right?'*

**Sunil Abraham (India)**

Further, it's worth noting that the transition towards data.gov and data.gov.uk took place across a longer timescale – at least four or five years – than the traditional two year funding cycle followed by many of the major grant-giving organisations. It is the researcher's view that funders should consider the 'long historical shadow' cast by these positive final outcomes when setting goals for similar interventions elsewhere in the world.

## Avenues for further research

This report has thrown up two valuable avenues for further research. The first is around the social impact of existing data catalogues such as [data.gov](http://data.gov) and [data.gov.uk](http://data.gov.uk). Much hope has been invested in the transformative potential of these catalogues to open issues in public life to a wide audience thanks to the interactivity and ubiquity of the World Wide Web in the US and UK. In order to build on their call for open data to be pursued across the world, funders have an interest in tracking over time to what extent open data catalogues deliver on this promise.

The second, and perhaps more pressing, research need is for a systematic evaluation of government data collection activities across all regions being considered for [data.gov](http://data.gov)-type interventions. Such an evaluation should capture not just what data is being collected, but also:

- how that data is being maintained (ie digitally or on paper);
- to what extent data collection practices are standardised across a country, or whether they are district- or region-specific;
- what formats the data is stored in, and whether those formats are open or proprietary; and (perhaps most importantly)
- what licensing practices are associated with redistribution or reuse of that data.

During the interview process, the researcher encountered several international monitoring NGOs with existing networks of in-country researchers which could be leveraged to complete such a study, perhaps in conjunction with an organisation such as mySociety or the OKF, or any other organisation that has special expertise in open data contingencies.

Finally, it's worth mentioning that several research projects which seek to provide technical guidance to those wishing to build open data portals are ongoing, including a nascent project by the OKF to produce an open data manual,<sup>52</sup> and a project to map open government data initiatives around the world that is being conducted jointly by [access-info.org](http://access-info.org) and the OKF.<sup>53</sup>

<sup>52</sup> See <http://lists.okfn.org/pipermail/okfn-discuss/2010-April/007244.html>. The manual will be based on the OKF's pamphlet 'Unlocking the Potential of Aid Information' (Gray, Hatcher, Hogge, Parrish and Pollock, 2009).

<sup>53</sup> See <http://opengovernmentdata.org/>

# An open data strategy checklist

Based on the evidence uncovered in this report, the following checklist is proposed for those considering initiating a campaign for data.gov and data.gov.uk-type projects in other countries.

## Status of FOI

- Does the country have a Freedom of Information law? Has it been implemented? Is it delivering on transparency? Is it delivering on social outcomes?
- Does the Freedom of Information regime include proactive disclosure/ publishing requirements as well as reactive disclosure, or is it based on request-only access?
- How is the administration coping with the proactive disclosure and reactive disclosure elements of the Freedom of Information regime that is in place?
- Has the judiciary been active in upholding Freedom of Information obligations?
- If the country does not have a Freedom of Information law, is there an active movement advocating for one?
- If the country does not have a Freedom of Information law, are there sectoral freedom of information laws, transparency provisions within the country's constitution, or applicable international standards (e.g. Aarhus convention) that are relevant and implemented?

## Current data collection activities

- What level of data collection is undertaken by the government?
  - Is this data collected in a systematic and timely fashion?
  - Is this data stored digitally or on paper?
  - In what format is digitised data collected and stored? Are these open or proprietary formats? Are they machine-readable?
  - What is the state of e-government activities in the country? Does the government have an e-government strategy? Is it implemented?
- Are government data or compilation of government data currently protected by copyright or another intellectual property-like regime?
- Is the data subject to any licences that restrict reuse? Are fees charged for access (for instance, to aid cost recovery)?
- To what extent are conversations around the digitisation of government subject to vendor capture?
- What sorts of privacy laws are operational in the country and how might they impact on open access to government data?

## Potential end-users

- How free is the press? How wired?
- Is there a user base of traditional civil society groups that may make use of targeted data?
- Are there specific examples of those groups using data in their advocacy/monitoring or other civic engagement activities?
- Are there specific examples of take-up of data by end users that may inform open data initiatives?
- In what ways did that data need to be made accessible in order for it to be used?
- What level of internet penetration is there across the country?
- What level of mobile penetration is there and how are people accessing mobile data services (SMS, 3G etc)?

## Tier 1: Civil society

- How technically literate is civil society? Are 'civic hackers' present?
- Are there instances in-country where local civil society is appropriating government data already?
- Is there an organised, technology-led local group, such as the Sunlight Foundation, or MySociety, in the country?

## Tier 2: The middle layer

- How empowered is the middle layer by the current political environment, in which departments and at what level?
- Is openness a high-stakes issue at the middle layer? How threatened might individuals advocating for openness within this layer feel? To what degree and how might middle level civil servants resist opening government data that was not collected with the intention of being released?
- How technically competent is the middle layer?
- How much capacity is there within the middle layer? How far has e-government penetrated?
- Does an existing network / community exist that bridges the gap between the middle layer and civil society?

### Tier 3: The top-level

- Is the country at a particular stage in the political cycle that would make opening government data advantageous in a political sense?
- Politically, who might the winners and losers be if government data is opened?
- To what degree and how is the expense of opening significant amounts of government data likely to be an issue?
- Regionally, could the country be described as lagging behind neighbouring countries in terms of openness and transparency?
- Could regional peer pressure have an impact on top level political will to open data (e.g. ASEAN, SADC)?
- What are the country's economic aspirations? Could an argument for efficiency be made?
- Are there exemplars of open data practice at the sub-national level?
- How is the top political tier likely to react to advocacy from any one of the following actors:
  - o Tim Berners-Lee?
  - o Barack Obama?
  - o Commercial operators (eg Google)?
  - o Entertainment industry: film stars?
- How does the country rank in corruption assessments? Do allegations of serious corruption reach the highest levels of government, or are they restricted to middle and lower levels of government?
- Is there an anti-corruption movement in the country, and could it be an effective ally?

### Tier 4: Donor drivers

- What levels of multilateral and bilateral donor involvement are there in the country's budget and governance?
- Are donors already releasing their own data openly?
- How thoroughly does the administration report on aid spending?
- How has the country reacted to previous tied aid? Is there scope for positive conditionality?
- Are there private donors (local or international) active in the country who could be useful allies?

# Conclusion

This report has been produced on the premise that there are substantial social and economic gains to be made from opening government data to the public, and that the combination of geographic, budget, demographic, services, education and other data, publicly available in an open format on the web promises to improve services as well as create future economic growth. Several interviewees approached during the research that went into this report have sought to challenge that premise, and those challenges have been noted. It is beyond the scope of this report to investigate those challenges in any further detail. However, the researcher suggests that those who wish to take this research further would do well to monitor the wider impact of data.gov and data.gov.uk, if only to be able to respond to such challenges in the future.

This research has sought to explore the feasibility of applying the approach to open data taken by the US and UK administrations (with the open data catalogues [www.data.gov](http://www.data.gov) and [www.data.gov.uk](http://www.data.gov.uk) respectively) to relevant middle income and developing countries. It has sought to identify the strategies used in the US/UK contexts, and it has proposed a set of criteria to guide the selection of pilot countries, criteria which in turn suggest a template strategy to open government data in a middle income or developing country.

The resulting checklist appears to give equal weight to each factor, and indeed one of the most striking aspects of the success of the two open data projects studied is how, in each case, that success was brought about by a broad range of concurrent factors and events. Nonetheless, the researcher would like to suggest that one aspect of the open data strategy deserves special highlighting: the existence, in both the US and the UK, of highly established data collection activities operated by a well-resourced, broadly independent and highly skilled 'middle layer' of government administrators. This is one aspect of the trajectory towards [data.gov](http://data.gov) and [data.gov.uk](http://data.gov.uk) that has perhaps attracted the least attention of onlookers, for two reasons: the activities of this section of society do not generally attract attention (at least not when they are functioning well), nor do its members seek it; and secondly, the activities of another group that has contributed to the open data initiatives – the so-called 'civic hackers' – are fresh and exciting, and therefore likely to be more attention-grabbing. All this is worth highlighting since, as several regional and domain experts have made clear during interviews, these 'middle layer' activities – both the data collection and the political position the administrators who undertake it exist within – are potentially weak or absent in most middle income and developing countries.

This report has focussed in some detail on the difficulty experienced in the UK of opening up geospatial data that was at the time commercially licensed in order to aid cost recovery. Such commercial / cost recovery activities have also been highlighted in the checklist. However, the relationship of commercial licensing with the eventual success of a data catalogue like [data.gov.uk](http://data.gov.uk) strikes the researcher as complex. It has been suggested in the course of this report that the barrier these activities imposed in the UK may have served as a common call to action among both civil society and the middle-layer government administrators, which in turn served to strengthen the crucial communication between these two groups in the trajectory towards [data.gov.uk](http://data.gov.uk), and ultimately enrich the final proposition when compared with [data.gov](http://data.gov).

Of course, without the intervention of Tim Berners-Lee, it is unlikely whether the UK government would have ever sanctioned the opening up of geospatial data. But there is a risk that funders, seeing the impact of this intervention,

might conclude that such an intervention is all that is needed to establish similarly vibrant open data initiatives in other countries. This would be a mistake. It is this researcher's opinion that without existing and meaningful activity around open data on all three tiers of public life identified in this report (with a potential fourth tier, that of international donors, also coming into play in developing and middle income countries) funders could be at risk of seeding a kind of cargo cult. The researcher agrees with Ethan Zuckerman, who, when asked to qualify his scepticism with a vision of how things might eventually play out around open data in Africa, made the following remark:

*'I think African governments end up opening their data through a combination of citizen pressure in countries where that matters and donor pressure in countries where that matters. So, for instance... in Ghana, you actually have an extremely engaged citizenry, and a sceptical and technically competent press starting to demand certain critical pieces of information. What that might require in Ghana is focusing on the journalists and essentially saying, "Why don't we have this data and this data and this data?", "What could we do, what questions could we ask and answer if we have it?", "Why can't we get it?" and then suing the government to get it if they don't have it. And I could imagine that leading gradually toward the release of data in that country. There are other countries where the government is a lot less [responsive] to popular pressure. I'm going to go out on a limb and say that Rwanda, which is functioning as a one-party dictatorship, is one of those countries. But donors are very powerful and Kagame is extremely concerned with being perceived as a modern, forward-looking benevolent dictator. And, I think in that case, using the pressure of the aid industry which represents a remarkable part of the government budget, and essentially saying, "For our sake, for your sake, for your citizens' sake, this is going to become mandatory, and we're going to release that information and it's going to be available not just within the international community, but to the world at large and particularly the journalists". I can imagine that working if only because we've got so much leverage in that situation.'*

**Ethan Zuckerman**

# Annexes

## Annex I: Methodology

This research was conducted over ten days in April 2010.

It drew on published research studies, commissioned reports, evaluations and other grey literature (listed in the bibliography) as well as telephone interviews.

The researcher was guided to a core group of interviewees who acted as primary sources of information (listed below). An additional interviewee – Jonathan Gray – was added at the researcher's discretion, on the basis of concurrent work he was conducting on the topic (which had been flagged by the research's commissioners), and on his status as a coordinator of an international network of open data enthusiasts.

Interviews were conducted in two rounds. The first round focussed on individuals who had been directly or indirectly involved with the development of data.gov and data.gov.uk. The researcher then used information garnered from this process, and from the literature review, to shape questions for the second round of interviews. The second round of interviews focussed on individuals with wider domain and region expertise.

## Annex II: List of interviewees

### Round one interviewees

- **Tim Berners-Lee**  
[www.w3.org/People/Berners-Lee/](http://www.w3.org/People/Berners-Lee/)
- **Steve Bratt**, CEO  
World Wide Web Foundation  
[www.w3.org/People/Bratt/](http://www.w3.org/People/Bratt/)
- **Jonathan Gray**, Community Coordinator,  
Open Knowledge Foundation Network  
[www.okfn.org/](http://www.okfn.org/)
- **Tom Steinberg**, Director, mySociety  
[www.mysociety.org/](http://www.mysociety.org/)
- **John Wonderlich**, Policy Director,  
Sunlight Foundation  
[sunlightfoundation.com/](http://sunlightfoundation.com/)

### Round two interviewees

#### Domain experts

- **Sunil Abraham**, Executive Director,  
Centre for Internet and Society (India)  
[www.cis-india.org/publications/cis/sunil](http://www.cis-india.org/publications/cis/sunil)
- **Nathaniel Heller**  
Managing Director Global Integrity  
[www.globalintegrity.org/aboutus/team.cfm#nheller](http://www.globalintegrity.org/aboutus/team.cfm#nheller)
- **Toby Mendel**, Executive Director  
Center for Law and Democracy
- **Vivek Ramkumar**  
International Budget Partnership  
[www.internationalbudget.org/](http://www.internationalbudget.org/)
- **Ethan Zuckerman**, Senior Researcher  
Berkman Center for Internet and Society  
<http://ethanzuckerman.com/>

### Round two interviewees

#### Regional experts

- **Jay Kruuse**, Head of Monitoring  
and Research Programme  
Centre for Social Accountability  
South Africa  
[www.psam.org.za](http://www.psam.org.za)
- **Dan McQuillan**  
Social Innovation Camps (CEE)  
[www.sicamp.org/](http://www.sicamp.org/)
- **Ory Okolloh**, co-founder Mzalendo  
co-founder, Ushahidi (Kenya)  
[www.kenyanpundit.com/](http://www.kenyanpundit.com/)
- **Rakesh Rajani**, Founder  
Twaweza (Tanzania)  
[www.twaweza.org/](http://www.twaweza.org/)

# Bibliography

- Allan, R.** (2009, November). *The Power of Government Information State of the eUnion: Government 2.0 and onwards* Retrieved April 16, 2010 from <http://21gov.net/wp-content/uploads/e-book.pdf>
- BBC News.** (2010, April 1). *Ordnance Survey offers free data access.* Retrieved April 16, 2010, from <http://news.bbc.co.uk/1/hi/8597779.stm>
- Berners-Lee, T.** (2009, February). *Tim Berners-Lee on the next Web.* (T. Berners-Lee, Performer) TED Conference 2009.
- Brown, G.** (2009, June 10). *Oral Answers to Questions – Prime Minister.* Retrieved April 16, 2010, from House of Commons debates: [www.theyworkforyou.com/debate/?id=2009-06-10c.795.1](http://www.theyworkforyou.com/debate/?id=2009-06-10c.795.1)
- Cabinet Office.** (2009, June 10). *Pioneer of the World Wide Web to advise the government on using data.* Retrieved April 16, 2010, from Cabinet Office: [www.cabinetoffice.gov.uk/newsroom/news\\_releases/2009/090610\\_web.aspx](http://www.cabinetoffice.gov.uk/newsroom/news_releases/2009/090610_web.aspx)
- Cabinet Office.** (2007). *The Government's Response to The Power of Information Review.* Retrieved from [www.official-documents.gov.uk/document/cm71/7157/7157.pdf](http://www.official-documents.gov.uk/document/cm71/7157/7157.pdf)
- Civic Impulse LLC 1.** (n.d.). *About This Website.* Retrieved April 16, 2010, from GovTrack.us: [www.govtrack.us/about.xpd](http://www.govtrack.us/about.xpd)
- Civic Impulse LLC 2.** (n.d.). *Sites that Use GovTrack Data.* Retrieved April 16, 2010, from GovTrack.us: [www.govtrack.us/downstream.xpd](http://www.govtrack.us/downstream.xpd)
- Crabtree, J.** (2010, January 27). *Mash the State.* *Prospect* .
- Escher, T.** (2009, March 5). *Citizens contacting representatives.* *Presentation to the Oxford Internet Institute Internet and Democracy Workshop.*
- Executive Office of the President,** Office of Management and Budget. (2009, December 8). *Open Government Directive.* Retrieved April 16, 2010, from [www.whitehouse.gov/omb/assets/memoranda\\_2010/m10-06.pdf](http://www.whitehouse.gov/omb/assets/memoranda_2010/m10-06.pdf)
- Gray, J., Hatcher, J., Hogge, B., Parrish, S., & Pollock, R.** (2009, December). *Unlocking the Potential of Aid Information Version 0.2.* Retrieved April 23, 2010, from [www.unlockingaid.info/wp-content/uploads/2010/02/UnlockingAidInformation.pdf](http://www.unlockingaid.info/wp-content/uploads/2010/02/UnlockingAidInformation.pdf)
- International Budget Partnership. (n.d.).** *Open Budget Survey 2008.* Retrieved from [www.openbudgetindex.org/index.cfm?fa=fullReport](http://www.openbudgetindex.org/index.cfm?fa=fullReport)
- Kirkpatrick, M.** (2010, January 20). *UK Launches Open Data Site; Puts Data.gov to Shame.* Retrieved April 16, 2010, from ReadWriteWeb: [www.readwriteweb.com/archives/uk\\_launches\\_open\\_data\\_site\\_puts\\_datagov\\_to\\_shame.php](http://www.readwriteweb.com/archives/uk_launches_open_data_site_puts_datagov_to_shame.php)
- Mayo, E., & Steinberg, T.** (2007, June). *The Power of Information: An independent review.* Retrieved from <http://195.99.1.70/advice/poi/power-of-information-review.pdf>
- mySociety.** (n.d.). *About Us.* Retrieved April 16, 2010, from TheyWorkForYou.com: [www.theyworkforyou.com/about/#1](http://www.theyworkforyou.com/about/#1)
- Newbury, D., Bently, L., & Pollock, R.** (2008, February 26). *Models of Public Sector Information Provision via Trading Funds.* Retrieved from [www.berr.gov.uk/files/file45136.pdf](http://www.berr.gov.uk/files/file45136.pdf)
- Office of E-government and IT; Office of Management and Budget.** (2009, December 3). *Data.gov Concept of Operations: Draft.* Retrieved April 13, 2010, from [www.ideascale.com/userimages/sub-1/736312/ConOpsFinal.pdf](http://www.ideascale.com/userimages/sub-1/736312/ConOpsFinal.pdf)
- Office of Fair Trading.** (2006, December). *The commercial use of public information.* Retrieved from [www.oft.gov.uk/shared\\_oftr/reports/consumer\\_protection/oft861.pdf](http://www.oft.gov.uk/shared_oftr/reports/consumer_protection/oft861.pdf)
- Power of Information Taskforce 1.** (2009, February). *Power of Information Taskforce Report.* Retrieved from <http://poit.cabinetoffice.gov.uk/poit/wp-content/uploads/2009/03/poit-report-final-pdf.pdf>
- Power of Information Taskforce 2.** (n.d.). *Show Us A Better Way.* Retrieved April 16, 2010, from [www.showusabetterway.co.uk/](http://www.showusabetterway.co.uk/)
- Power of Information Taskforce 3.** (n.d.). *Taskforce Membership.* Retrieved April 16, 2010, from Power of Information Taskforce: <http://powerofinformation.wordpress.com/taskforce-membership/>
- Robinson, D., Harlan, Y., Zeller, W., & Felten, E.** (2009). *Government Data and the Invisible Hand.* *Yale Journal of Law and Technology* , 11, 160.
- Tactical Technology Collective.** (2009). *Case Study: mySociety.* Retrieved April 26, 2010, from Ten Tactics for turning information into action: [www.informationactivism.org/viewtactic1](http://www.informationactivism.org/viewtactic1)
- Torkington, N** (2010, February 2). *Rethinking Open Data.* Retrieved April 16 2010 from O'Reilly Radar: <http://radar.oreilly.com/2010/02/rethinking-open-data.html>

## Acknowledgements

The researcher would like to thank the Transparency and Accountability Initiative (a donor collaborative) for providing her with the opportunity to conduct this research, and in particular Martin Tisne, Programme Manager at the Transparency and Accountability Initiative, as well as Janet Haven and Darius Cuplinskas of the Open Society Institute's Information Programme, for their advice and guidance during the research. Thanks too to Steve Bratt of the Worldwide Web Foundation and Stephen King of the Omidyar Network. Finally, the researcher would like to extend her heartfelt thanks to all of those who agreed to be interviewed for this project, for giving up their valuable time to share their expertise and insights with others.

This work is licensed under a Creative Commons Attribution 3.0 Licence:  
<http://creativecommons.org/licenses/by/3.0/us/>

## About the author

Becky Hogge is a freelance writer and broadcaster. Her writing on information politics, human rights and technology appears regularly in the UK press, and she has produced research papers on internet governance, online campaigning and online markets for various international NGOs and grant-giving organisations. She has served on several national awards panels, and has been invited by the UK and other European governments to consult on information policy. Her opinions on internet governance and internet censorship have been broadcast around the world. Becky is the former Executive Director of the UK's Open Rights Group, and she sits on the board of the Open Knowledge Foundation. Her first book, *Barefoot into Cyberspace*, will be published in summer 2011.

