

Towards Open Data for Public Accountability: Examining the US and the UK Models

Alon Peled, Hebrew University of Jerusalem
Nahon Karine, University of Washington

Abstract

Since 2009 governments worldwide have been developing Open Government Data (OGD) programs. Our paper examines the ways in which public agencies in the two leading OGD countries, the US and the UK, have released information assets to promote public sector accountability. Theoretically and empirically, we discuss the vision and execution of the OGD policy in these countries since 2009 and demonstrate that the movement took a different path than the one chartered by the original policy makers. We then propose a new path for the OGD movement to better support the democratic ideal of improved accountability of public sector agencies.

Keywords: (Open Data; Accountability; Open Government; Transparency; Disclosure)

Citation: Peled, A., Nahon, K. (2015). Towards Open Data for Political Accountability: Examining the US and UK models. In *iConference 2015 Proceedings*.

Copyright: Copyright is held by the authors.

Contact: Alon.Peled@post.harvard.edu, karineb@uw.edu

1 Introduction

The conceptual origins of the open government data (OGD) movement¹ and freedom of information (FOI) legislation are rooted in a desire to improve public accountability. However, the OGD movement blurred the original intent, combining the goal of accountability with concepts such as participation, collaboration, innovation and improvement of public service. The latter may have no significant impact or even contradict accountability (Peled, 2011; Yu & Robinson, 2012). We therefore ask: Do the designs of US and UK OGD policies demonstrate an understanding of how to harness the potential of OGD to improve public sector accountability? If not, what can be done to fulfill OGD's promise to improve accountability?

We propose that open government programs must be evaluated according to three dimensions: informational, explanatory and enforcement. By empirically examining US and UK OGD datasets, we found a divergence between the OGD vision and its execution since 2009 in these two countries. We explain how the OGD movement grew to its current state of 'release any data now' disconnected from the accountability goal. We also propose to provide incentives to agencies to release even less datasets but ones that are more closely aligned with the goal of accountability.

Critics may propose that 'improved accountability' is but one goal competing against other, no less valuable goals (Robinson et. al. 2009). We propose that the goal of 'improved accountability' is so demanding that it sets the 'OGD bar' very high. To meet this bar, agencies will be forced to publish information that is better linked to other information, better documented, and embedded in richer context. Such information will therefore also contribute to the attainment of additional OGD.

2 Open Government Data: Concise Background (2009-Present)

2.1 The US OGD Program

President Obama's 2009 promise to free non-classified information raised hopes that the new avalanche of open, free and easily accessible government information would unleash improved accountability in the US. On his first full day in office (January 21, 2009), at the height of the worst economic crisis America had experienced since the Great Depression, President Obama signed three memoranda and two executive orders. Four of these five documents promoted open government (Obama, 2009). After March

¹ In this paper, we refer to the collective efforts of governments worldwide to release information to the public since 2009 as "the OGD movement." We use the term "OGD 1.0" to refer to the period between 2009 and 2012 during which legislators asked agencies to release information *voluntarily*. We use the term "OGD 2.0" to refer to the period after mid-2013 when US and EU legislators adopted new *mandatory* OGD guidelines that *compel* agencies to release some information.

2009, a barrage of open government initiatives surfaced including eRulemaking, IT Dashboard, Recovery.gov, and USAspending.gov. The administration showcased open government stories and ensured that senior appointees adhered to open government principles (Office of Management and Budget (OMB), 2013).

On May 21, 2009, a team headed by the CIOs of the US Department of the Interior (DOI) and the US Environmental Protection Agency (EPA) launched www.data.gov as the premier web publishing location of the most important federal datasets. Next, on December 8, 2009, the OMB published the Open Government Directive (OMB, 2009). The directive ordered agencies to make as much information as possible available online. At the outset, agencies were instructed to publish at least three high-value datasets on the web. High-value datasets were defined as containing: "information that can be used to increase agency accountability and responsiveness; improve public knowledge of the agency and its operations; further the core mission of the agency; create economic opportunity; or respond to need and demand as identified through public consultation" (OMB, 2009). High-value datasets were required to have never before been made available or published in a downloadable and open format. The first three high-value datasets were considered a down payment, or a minimum; federal agencies were expected to continually make new datasets available to the public (McDermott, 2010).

In May 2013, on commencing his second term, President Obama raised the open government bar with an open government memorandum that ordered the default state of new and updated government information resources to be made open and machine-readable. He also instructed that government information be managed as an asset throughout its life cycle. Agencies were required to maintain and nurture their datasets as products with real market value: as assets that could be bought, sold, and exchanged. In this article, we adopt the "asset" concept and refer to the OGD information artifacts that agencies generate, maintain, and upgrade as information assets (Executive Office of the President, 2013). To ensure that these new orders would be followed, President Obama enforced a rigorous OGD execution timetable with which agencies were required to comply. In another important expansion of the 2009 directive, President Obama extended OGD policy to all federal agencies by instructing independent agencies to comply with this policy (OMB, 2013).

Following the 2013 memorandum, OMB published a detailed implementation guide instructing agencies on how to follow the President's latest OGD policy. First and most importantly, all federal agencies were ordered to develop a Public Data Listing, which contains a list of all information assets that are or could be made available to the public. This Public Data Listing was to be posted at [www.\[agency\].gov/data.json](http://www.[agency].gov/data.json). The intention was to provide the public with access to an extensive inventory of agencies' information assets, enable the public to track agencies' subsequent progress to clear more assets for publication, and empower the public to re-use open data. In addition, according to the new OGD standard, US federal agencies were instructed to improve the quality of information assets that were previously released to the public and enrich the metadata descriptions of published information assets so the public could discover them more easily. Finally, all agencies were instructed to include on their open data websites, at a minimum, all information assets that were published on data.gov before August 1, 2013 as well as additional representative assets from programs and bureaus.

However, like OGD 1.0, the OGD 2.0 program was ill funded. Agencies did not receive funding or other incentives to pay for the preparation of data for public release (Peled 2013). Then, in mid-2011 the meager US federal e-Government budget was slashed from \$35 million to \$8 million (this amount was shared between several e-Government programs including OGD). Kundra, the first US federal CIO, resigned from his post to protest these budget cuts (Wadheia 2011).

2.2 The UK OGD Program

The UK government has pioneered new techniques to release information to the public². In accordance with this tradition, in January 2010 the national UK OGD portal (www.data.gov.uk) was officially launched. The Transparency and Open Data team in the Cabinet Office spearheaded data.gov.uk and they benefited from the active support and involvement of several of the most famous entrepreneurs in the history of the Internet including Sir Tim-Berners Lee and Professor Nigel Shadbolt. In 2010 the UK

² An interactive timeline that displays these experiments can be found here: <http://timemapper.okfnlabs.org/timdavies/uk-open-data-timeline#13>

Government created the Open Government License (version 2. of this license was released in June 2013); public bodies can publish their Crown Copyright material under this license. An Open Data Institute was created as part of the UK OGD initiative in October 2013.³ Another Open Data Users Group has been in operation since 2012, acting as a conduit for data requests and advising government on priority datasets to release. The Open Data Barometer Report ranks the UK as the most advanced country for open data readiness, implementation and impact, scoring above the US (2nd), Sweden (3rd), New Zealand (4th), and Denmark and Norway (joint 5th).

It is important to note however that, initially, the UK OGD architects oriented www.data.gov.uk towards supporting innovation and economic growth and were less concerned about promoting accountability than their US counterparts.⁴ Only at the beginning of 2014, the Cabinet Office began addressing the issue of using OGD information assets to promote accountability.⁵ At present, it appears that overall the UK OGD movement has overlooked an important opportunity to use OGD to promote public accountability. In addition, critics including the National Audit Office criticized the UK OGD campaign for being overly costly and for not attracting enough citizens to download data.⁶ The latter criticism is an indication that insufficient context is provided to help citizens find the information assets they need.

It is important to note that as a member of the European Union (EU), the UK is subject to the latest EU mandatory OGD policies and guidelines. In April 2013, 27 EU member states formally approved the European Commission's (EC) new rulebook on the re-use of Public Sector Information (PSI). The new rulebook instructed European agencies (including UK agencies) that the disclosure of public data for reuse is obligatory (i.e., such disclosure was optional in the older version of this rulebook published in 2007).⁷

Finally, the UK national OGD program was brought under the Cabinet Office in 2012 with a \$3.12 million budget (smaller than that of the US federal program). However, it appears that UK national departments spend—each—an additional \$83,000 to \$500,000 each year to prepare data for public release (Rogers 2012). This significant department/agency-level OGD commitment may be one important differentiating factor between the relative success of the US and UK OGD national programs.

3 Theory: Transparency, Accountability, and OGD

Transparency and accountability in democracies are the mechanisms, that assures a 'good' government, that is, an honest, efficient and effective government. Increasingly, scholars connect the themes of transparency to accountability (Braman, 2006). For example, Piotrowski claimed that governmental transparency allows the public to develop a more accurate picture of the operation inside governments, and therefore, to evaluate the performance of agencies and hold them accountable for their actions (Piotrowski, 2007). Bertot et al. advocated for transparency as an approach to overcome corruption (Bertot, Jaeger, & Grimes, 2010).

Obviously, the normative benchmark of what constitutes a 'good' government varies from one government to another, as the norms driving political cultures differ. However, the primary way to guarantee a 'good' government is by institutionalizing powerful accountability mechanisms that hold every public official responsible for their actions as public servants, and public institutions responsible for institutional decisions and actions (Ackerman, 2004). Scholars use different analytical schemas to define and classify different types of accountability (O'Donnell, 1998).

This paper uses Schedler's theoretical framework for accountability, as it is easy to operationalize for our OGD needs. According to Schedler, accountability is a process which involves both *answerability*, "the obligation of public officials to inform about and to explain what they are doing" (Schedler, 1999, p. 14)

³ <http://theodi.org/odis-first-year-annual-report>.

⁴ <http://www.opendataresearch.org/dl/odb2013/Open-Data-Barometer-2013-Global-Report.pdf>, page 6, 27.

⁵ <https://www.gov.uk/government/policies/improving-the-transparency-and-accountability-of-government-and-its-services>

⁶ <http://www.theguardian.com/news/datablog/2012/apr/18/uk-open-government-data-national-audit-office>

⁷ <http://www.epsiplatform.eu/content/eu-endorses-new-psi-directive>

and *enforcement*, “the capacity of accounting agencies to impose sanctions on power holders who have violated their public duties” (Schedler, 1999, p. 14). Particularly relevant to our paper, Schedler delineates three dimensions of public accountability: informational, explanatory, and enforcement. The informational dimension asks what agencies or public servants did or plan to do. The explanatory dimension provides reasons, justifications and judgments regarding why agencies or public servants decided to carry out or did carry out a certain action. Finally, the enforcement dimension rewards good or punishes bad agencies’ or public servants’ behavior (Schedler, 1999, p. 15). Other scholars use alternative nomenclature to describe these three dimensions (Allen, 2003) but these three dimensions are a “continuous variable that show up in different degrees” in public sector agencies (Schedler 1999 15).

The OGD programs that were implemented after 2009 emphasized the power of transparency and technological innovations to strengthen public sector accountability (Bass et al., 2010; Gurstein, 2011). The claim was that technology facilitates the process of disclosing information to the public (Noveck 2009; Berners-Lee 2010). However, recently, voices emerged warning of the problematic assumption that technology leads to increased transparency which, in turn, leads to increased accountability (Barry & Bannister, 2013; Fox, 2007; Shkabutar, 2012; Yu & Robinson, 2012). Other scholars warned that embracing OGD programs may serve as a cover to divert energies from Right to Information programs (Davies, 2013; Horner, 2012; Janssen, Charalabidis, & Zuiderwijk, 2012; Rosenberg, 2013). Nonetheless, scholars have provided largely anecdotal evidence to support these critical claims. The need to empirically examine the relationship between OGD programs and accountability is strong. In this paper we will focus on the US and UK cases.

4 Methodology: Harvesting Metadata About OGD Information Assets in the US and the UK

4.1 Methodology: Building and Analyzing a Corpus of Metadata about US and UK OGD Assets

In order to examine whether US and UK OGD programs are designed to promote accountability, we systematically studied the metadata about the information assets released by government agencies in the these two countries. To do so we used a two-step approach: First, we developed a software able to automatically extract the metadata as provided by agencies. Second, we mapped the metadata and its content according to Schedler’s theoretical framework of accountability levels: informational (e.g, what agencies or public servants did or plan to do), explanatory (e.g, reasons why agencies decided to carry out a certain action) and enforcement (e.g, rewarding good or punishing bad agencies’ or public servants’ behavior) (Schedler, 1999).

Many OGD studies focused on the US federal www.data.gov and the UK national www.data.gov.uk portals. In the UK case, scholars could easily harvest metadata about the information assets that British agencies uploaded because the UK OGD model provided a unified and singular Application Programming Interface (API) that all UK agencies must use to upload information assets. However, the story is more complex in the US case. Here too, other scholars found it easier to harvest metadata from a single web portal (www.data.gov) than collect similar data from federal agency-based OGD sites. However, www.data.gov itself is a ‘whole seller service provider’ that harvests metadata about OGD assets from agency-based OGD portals. Therefore, in the case of the US, we decided to harvest the metadata directly from the agency-based OGD portals. By so doing, we successfully harvested metadata directly from the point-of-origin where the OGD information assets were *first* published.

Our software solution crawls, scrapes, collects, cleanses, and registers important metadata⁸ about OGD information assets. In the US case, OGD information assets are harvested from a list of 473 US federal agencies⁹, focusing on agency open government portals. Currently, for the US, the software solution only crawls through the `data.json` format catalogs that US agencies publish on OGD portals. However the,

⁸ Metadata about open government data portals and open government data information assets exists (for example <http://www.opendatainitiative.org> or <http://datahub.io>). Some of it is designed to help end-users discover datasets. A handful of databases address the end-user experience such as the European Engage project (<http://www.engagedata.eu>) or rely on surveys such as the OD Barometer. Nonetheless, not a single corpus exists to help social scientists test OGD hypotheses. Therefore, the first task for this paper was to create this corpus and perform the initial analysis of its contents.

⁹ The number of active federal agencies is a topic of debate. For this matter we used the list generated by the 2013 US Manual <http://www.usa.gov/Agencies/Federal/Executive.shtml>

data.json catalog remains the most important standard and the only mandatory catalog that the OMB instructs agencies to use. For the UK case the software solution harvests metadata and information assets that exist in the CKAN format. Henceforth, we will refer to the US data.json format as the “US catalog” and to the UK CKAN format as the “UK catalog.”¹⁰

Next, we studied and classified 28 metadata fields based on Schedler’s schema of three dimensions of political accountability. Table 1.0 below displays the results of our analysis. Metadata fields that provide primordial context about a given “information asset” were placed under the informational dimension. For example, the “language” of a given information asset is an example of a primordial piece of metadata information because a given information asset is either in English or not. We either know or do not know what the “language” of the information asset is. There is no room to negotiate or ask follow-up questions about the “language” of a given information asset. In contrast, the name of the person who uploaded the information asset and his or her email address are good examples of metadata fields that belong to the explanatory dimension of accountability. Citizens can actually contact this person, ask questions about the information asset and receive additional valuable information that deepens the context of this information asset and makes it more useful. Finally, metadata fields such as legislative budgetary codes or Information Technology (IT) investment codes belong to the enforcement dimension of accountability because legislative aids, auditors, and the media can associate these codes with the codes of funded programs and legislative budgets. Potentially, such associations can provide the auditors with powerful tools to closely track agencies execution of publicly funded programs.

4.2 Methodology: Limitations

Three drawbacks characterize our methodology: its exclusive quantitative nature, tight focus on the ‘OGD supply side’ dimension, and reliance on the types of catalogs used by UK national agencies and US federal agencies. First, we did not interview, survey, or observe public sector officials involved with OGD activities. We plan to do so in our future work. However, we believe that there is important value in empowering OGD information assets to ‘tell their stories.’ Governments often contain a handful of enthusiasts who are passionate about OGD. Yet, these enthusiasts cannot overcome some of the deep, structural problems that the OGD movement is currently experiencing. Our analysis in this paper reveals some of these problems.

Second, the metadata we collected and analyzed is exclusively about the ‘OGD supply side’ (e.g, the information that agencies release) and it contains no information about the ‘OGD demand side’ (e.g, to what extent citizens downloaded OGD information). Nonetheless, by now, we have sufficient evidence to suggest that neither US citizens nor UK citizens did much with the published OGD information since 2009. Only 1% of all www.data.gov visitors have downloaded a dataset. Likewise, the UK Comptroller discovered that 80% of all visitors to www.data.gov.uk left the site without downloading data (Bannister & Connolly, 2011; Davies, 2010, 2011; Fioretti, 2012; Janssen, 2012; McClean, 2011; Peled, 2011; Rogers, 2012; The White House, 2012).

Finally, our software currently only knows to harvest metadata published by the CKAN catalog (which we call the “UK catalog”) and the data.json catalog (dubbed the “US catalog”). This is no small achievement because CKAN is the mandatory OGD catalog used by UK national agencies and data.json is the mandatory one for US federal agencies. Nonetheless, we know that many government agencies publish OGD information assets using older types of catalogs that our software currently cannot access. We will extend our analysis to cover additional OGD catalogs in the near future.

5 Findings

The US and UK OGD catalogs are poorly documented, technically oriented, and difficult to decipher and navigate. The UK OGD portal directs visitors to an outdated version of the guidelines to the Application Programming Interface (API) of the UK catalog standard, through which developers and users approach

¹⁰ We believe that the choice of a particular OGD catalog is political rather than technical in nature. During the early OGD period (2009-2013), OGD architects implemented the catalogs they themselves were familiar with (Kundra and data.json in the US and Sir Tim-Berners Lee and Professor Nigel Shadbolt and CKAN in the UK). OGD catalogs have proliferated. We urge other researchers to pay close attention to the initial OGD catalog choices that agencies make.

public information¹¹. The UK catalog also uses technical examples to instruct visitors on using the catalog¹². The US catalog provides more complete documentation accessible to non-technical visitors¹³. The open data project remains part of US federal policy and therefore the US catalog is better documented than the UK one (including mappings between different fields in different systems).

In order to learn about the way in which disclosure of data enhances accountability, we analyzed the metadata fields and classified them according to Schedler's theoretical framework, and then analyzed the content of the metadata fields which summarize on an aggregative level the content of the published UK and US OGD assets.

5.1 Accountability Classification of Open Government Metadata

The public servants who release information are required to add metadata, data about the released data. Currently, there are 29 metadata fields¹⁴. Table 1 shows how the 29 metadata fields in the US catalog are mapped into three levels of accountability.

Table 1.0 "Metadata Classified by the Three Levels of Accountability"

Dimension	Informational	Explanatory	Enforcement
Common Core Required Fields	Title	Contact Name	
	Description	Contact Email	
	Tags	Public Access Level	
	Last Update		
	Publisher		
	Unique identifier		
Common Core Required-if-Applicable Fields	Download URL	Access Level Comment	Bureau Code
	Endpoint		Program Code
	Format		
	License		
	Spatial		
	Temporal		
Expanded Fields	Distribution	Category	Primary IT Investment UII
	Frequency	Data Dictionary	
	Homepage URL	Data Quality	
	Language	Related Documents	
	Release Date		
	System of Records		
Total	18 fields	8 fields	3 fields

¹¹ The updated version of the API of CKAN is <http://docs.ckan.org/en/ckan-2.2/api.html>; the older version pointed to is <http://docs.ckan.org/en/ckan-2.0.2/api.html> (accessed on Sept. 7, 2014)

¹² http://demo.ckan.org/api/3/action/package_show?id=adur_district_spending

¹³ <http://project-open-data.github.io/schema>

¹⁴ More information about the metadata can be found at <http://project-open-data.github.io/schema/>

According to table 1.0, most of the metadata elements are mapped at the lowest level of accountability: the informational level. The US catalog defines 29 metadata fields that agencies must or can use to annotate their information assets before they upload these assets to OGD portals. Most of these template-metadata-fields (18 fields or 62% of all the predefined metadata fields) belong to the informational dimension of accountability. For example, Metadata such as “Title”, “Tags,” and “Language” empower agencies to provide facts about their information assets. Some of these template-metadata-fields (8 fields or 28% of all the predefined metadata fields) empower agencies to inform citizens about the information assets they upload at the explanatory dimension of political accountability. For example, fields such as “Contact Name,” “Contact Email,” “Access Level,” and “Public Access Level” provide agencies an opportunity to inform citizens which official uploaded the information asset, how to contact this official (i.e., by providing an email address), and learn what public accessibility level was designated to the information asset and why. Finally, only a small number of template-metadata-fields (3 fields or 10% of all the predefined metadata fields) provide agencies with the opportunity to tell the public something meaningful about the information assets at the enforcement level of political accountability. For example, fields such as “Bureau Code,” “Program Code,” and “Primary IT Investment UII” provide agencies with an opportunity to anchor each information asset that they upload to a specific sub-unit within the agency, to one or more programs that the legislature approved and funds, or to a unique Information Technology (IT) investment that the agency is approved to use.

It is critical to note here that the OGD architects could have pre-defined a larger number of metadata fields at the explanatory or enforcement dimensions of political accountability. The fact that agencies do not possess or are unprepared to provide metadata information about the explanatory and enforcement dimensions of political accountability does not detract from the responsibility of OGD architects to establish a high threshold for annotating information assets according to the desired goal (i.e., increased accountability). In fact, setting the bar high in all matters regarding the publication of information assets is critical in designing and directing the evolution of the OGD movement.

Finally, It is also important to note that the three leading OGD catalogs each provide an “extra” section that enables agencies to upload information assets in any form. Numerous agencies exploit this “miscellaneous” option. For example, the overwhelming majority of information from the State of Illinois is deposited in this “extras” section..

5.2 Accountability Classification of Agencies’ Information Assets

The second part of the analysis includes studying the content of the metadata, and how it relates to accountability. The empirical findings are based on our analysis of 49,436 information assets that US federal agencies and UK national agencies uploaded to the US OGD portal (www.data.gov) and the UK OGD portal (www.data.gov.uk), respectively. 29,798 of these assets or 60% of the total assets we analyzed were uploaded by US federal agencies while the remaining 19,638 of these assets, or 40% of the total assets we analyzed, were uploaded by UK national agencies. About 60% of all the US information assets were created or last updated after 2009, the year when President Obama first unleashed the OGD campaign. UK agencies reported that 75% of their information assets were last updated in 2014 and the remaining 25% were last updated in 2013.

Table 2.0 captures how far UK and US agencies are willing to go to comply with new OGD policy. At the informational level of accountability, both UK and US agencies provide minimal metadata information to help citizens understand what the information asset is about, such as “title,” “description,” and “keywords.” However, the willingness of UK and US agencies to comply with OGD dictates dissipates quickly when we ascend to examine metadata attributes at the explanatory and enforcement levels of political accountability. Important information that potentially could tie the published information assets to budgets and programs is completely missing.

Table 2.0 “US and UK Agencies’ Compliance with OGD Catalogs”¹⁵

Dimension	Informational	Explanatory	Enforcement
Common Core Required Fields	Title (US: 29,795, 99.99%) [UK: 19,638,100%]	Contact Name (US: 9,483, 31.82%) [UK: 96,0.0049%]	
	Description (US: 29,782, 99.95%) [UK: 19618,100%]	Contact Email (US: 29,795, 99.99%) [UK: 155,0.0079%]	
	Tags (US: 13,374, 44.88%) [UK: 19,618,99.90%]	Public Access Level (US: 29,795, 99.99%) [UK: 0,0%]	
	Last Update (US: 29,168, 97.89%) [UK: 19,638,100.00%]		
	Publisher (US: 29,795, 99.99%) [UK: 19,636, 99.98%]		
	Unique Identifier (US: 0, 0%) [UK: 0,0%]		
Common Core Required-if-Applicable Fields	Public Access Level (US: 29,798, 100.00%) [UK: 19,638, 100.00%]	Access Level Comment (US: 612, 0.03%) [UK: 0,0%]	Bureau Code (US: 29,061, 97.5%) [UK: 19,636, 99.98%]
	Format (US: JSON) [UK: CKAN]		Program Code (US: 8,796, 29.52%) [UK: 0,0%]
	License (US: 3,496, 11.73%) [UK: 15,247,77.64%]		
	Spatial (US: 27,458, 92.15%) [UK: 0,0%]		
	Temporal (US: 24,035, 80.66%) [UK: 0,0%]		
Expanded Fields	Distribution (US: 0, 0%) [UK: 0,0%]	Category (US: 23,942, 80.35%) [UK: 19,638,100%]	Primary IT Investment UII (US: 0, 0%) [UK: 0,0%]
	Frequency (US: 23,293, 78.17%) [UK: 0,0%]	Data Dictionary (US: 0, 0%) [UK: 0,0%]	
	Homepage URL (US: 0, 0%) [UK: 0,0%]	Data Quality (US: 3,660, 12.28%) [UK: 0,0%]	
	Language (US: 2,447, 0.08%) [UK: 0,0%]	Related Documents (US: 0, 0%) [UK: 0,0%]	
	Issued (US: 22,963, 77.06%) [UK: 0,0%]		
	System Of Records (US: 0, 0%) [UK: 0,0%]		

¹⁵ Per cent numbers for UK assets are computed against the total 19,638 British information assets analyzed. Per cent numbers for US assets are computed against the total of 29,798 information assets analyzed. Per cent numbers are rounded to the second number to the left of the decimal point (e.g., 2.48889% rounded to 2.49%). So, for example, under the “Title” metadata-field, we can see that 99.99% of all published US OGD assets and 100% of all published UK OGD assets were published with a title.

In fact, the picture is even bleaker than it appears at first sight. Table 2.0 only captures a high-level view of how US federal agencies and UK national agencies chose to upload their information assets. Unfortunately, when we begin looking deeper into the published data and the metadata describing it we discover even more problems. The insufficient metadata published by both US and UK agencies is of poor quality. For example, at the explanatory level of political accountability, only 3,660 US information assets (12.28% of the total number of information assets published by US federal agencies) have a value for “data quality.” However, when we examine these 3,660 information assets more closely, we discover that agencies marked 99.48% of them as being of good quality.

Likewise and also at the explanatory level of political accountability, we are initially impressed by the fact that no less than 29,795 US information assets (or 99.99% of the total information assets that US federal agencies uploaded) contain an email address associated with them. This email address, as mandated by the US catalog must be a “Contact person’s email address.” Potentially, this email address can be an important metadata attribute to help citizens who download the data contact the agency to receive further explanation and context about downloaded information assets. However, when we examine more closely the *actual* email addresses that US federal agencies provide we discover that the majority of these email addresses are generic, departmental email addresses. In fact, the generic NODC.DataOfficer@noaa.gov email address is associated with no less than 16,531 information assets or 55.48% of all the information assets that American agencies uploaded.

On diving deeper into the metadata about the information assets that US agencies published, we discovered that no less than 82% (33,862 of 41,702) of all the downloadable files associated with these information assets were not machine-processable as required by all OGD policies published in both the US and the UK since 2009. Put differently, we discovered that it did not matter much if end-users (after much trouble) managed to navigate their way through the poor metadata that US federal agencies published. In more than 80% of the cases, such navigation efforts end in frustration when end-users discover that the “data” they seek is saved in a non-machine-processable filetype such as pdf, html, jpg, or tiff. In addition, we discovered that 6% of all the download-URLs (i.e., the URL provided by the agency to empower the end-user to download the actual data) are broken links.

Finally, there are some differences between the US and the UK “national OGD style.” US federal agencies tend to publish metadata in more metadata-fields than UK agencies. However, in general, UK national agencies tend to do a better job than US federal agencies in the provision of higher quality metadata information. Thus, for example, overall, UK agencies do a better job than their American counterparts in publishing “keywords” that help citizens find useful information assets. Likewise, UK agencies do a better job than US agencies in associating information assets with a specific office or bureau within a large governmental ministry or agency. Though insufficient by itself, the close association of information assets to sub-units within large governmental entities is important for establishing accountability at the explanatory level because citizens can more accurately identify the origin of the data.

6 Discussion

6.1 The OGD Metadata Problem

Before 2009, US federal agencies and UK national agencies were content to conduct ‘business-as-usual’ in all matters regarding governmental information: not dishing out information, not explaining their decisions, not suffering punishments due to bad behavior. However, at the beginning of 2009, the creation and popular endorsement of the global OGD movement by governments caught US and UK agencies by surprise. OGD policy forced them to start changing their traditional “disclosure behavior, where the default is a ‘closed’ government and disclosure is on a ‘need to know only’ basis.

Initially, several US and UK agencies agreed to publish some OGD assets. Our study demonstrates that the vast majority of these assets fall into the lowest information dimension of public accountability. For example, the US and UK agencies almost exclusively published information assets about the domains they regulated. They published almost no information about their inner workings, decision-making processes, or performance. Therefore, since 2009, in both the US and the UK, we find almost no information assets that promote improvement at the explanatory or enforcement dimensions of accountability

With time, this practice (i.e., the release of information assets at the informational level only) grew to become the new norm. Agencies learned that it was relatively easy, fast, not expensive, and unthreatening to publish large quantities of OGD assets that are unrelated to the explanatory or enforcement dimensions of political accountability. We have no evidence to support the argument that US and UK agencies purposefully published poorly documented data at the lowest, informational level of accountability. We believe that agencies released poor quality information because they did not receive incentives to behave differently. It has been demonstrated in academia too, that scholars who do not receive incentives tend to develop poor metadata (Griffiths 2009). Alternatively, agencies may still be learning to adjust to the new OGD policy and, possibly, might develop better metadata in the future. Either way, we believe that legislators must provide specific incentives to agencies to develop improved OGD metadata.

For their part, politicians were pleased to see any information released by agencies, and eased the pressure put on public officials. Also OGD architects, such as Noveck and Kundra in the US and Berners-Lee and Shalblot in the UK, were so eager to see information released, that they praised the release of *any* information asset by any agency. In the memorable words of Berners-Lee during his 2012 Ted presentation: “Raw Data, Now!”. The result *de facto* was that the technical staffers working to implement the vision of the OGD architects built a modern, electronic ‘Tower of Babel’ based on incompatible or poorly compatible catalogs through which agencies release information. The OGD goal was and unfortunately remains ‘release something now!’ Therefore, currently, the OGD movement appears to be running forward with little design or purpose or attempt to connect democratic ideals such as improved accountability of public agencies to the vision and practices of OGD.

6.2 Returning to the Original OGD Vision

At the beginning of the second period of the OGD Movement (from 2013), which we call “OGD 2.0,” OGD architects pushed towards ‘linked open data’ referring to end users’ need to find and link together different information assets from different OGD web portals. US and UK agencies then released new information that was more of the same: -- poorly documented information assets that cannot be linked to each other, are devoid of useful context, and cannot inform citizens about why agencies acted in certain ways (the explanatory dimension of accountability) or how we can hold agencies accountable to their actions (the enforcement dimension of accountability).

Yet ‘OGD 1.0’ (2009-2012) architects themselves created the ‘linking’ problem that ‘OGD 2.0’ (2013-) architects now seek to solve. Worst still, there is no solution in sight for this ‘linking’ problem because ‘OGD 2.0’ provides more-of-the-same “OGD 1.0.” Therefore, the only solution is to return to the drawing table and pay careful attention to the key theme of this article--how OGD information assets can and must contribute to democratic ideals such as improved accountability. Again, the growing number of OGD assets on the web since 2009 does not equate to an ability to hold governments accountable for their actions; for this much more is needed. We must stop following ‘raw data, now!’ because present OGD practice is leading us down a dangerous path towards finding ourselves swimming helplessly and without aim in a vast ocean of meaningless data.

Some scholars have already given up on open data all together. In a recent paper, Clarke and Margetts proposed that big data rather than open data might be a “better long-term bet for more substantively transparent, efficient, and citizen-focused government” (2014, pp. 412-413). We believe that it is premature to give up on the original open data ideal. There is still hope that open data will meet its original improved political accountability goal if OGD architects re-design OGD policies to support improved accountability.

To re-align OGD policies with the original OGD accountability goal, OGD architects must return to the normative ideal and ask what they expect to receive in return for instructing agencies to spend precious resources to release information assets to the public. Improved accountability as a goal may be translated into better templates or catalogs that instruct agencies on preparing information assets before they release them. These architects then can design mechanisms that monitor and measure agencies compliance with these templates. If OGD architects re-design OGD with accountability as the key goal, we may end up releasing less OGD information but citizens will be empowered to more easily find and link together the released information. For example, mandating that every new OGD information asset be

released with metadata information that ties this asset to specific legislative programs and public sector budget articles will empower citizens to more easily and effectively link numerous disparate OGD assets.

7 References

- Ackerman, J. (2004). Co-Governance for Accountability: Beyond "Exit" and "Voice." *World Development*, 32(3), 447–463.
- Allen, A. (2003). *Why Privacy Isn't Everything: Feminist Reflections on Personal Accountability*. Rowman & Littlefield.
- Bannister, F., & Connolly, R. (2011). The Trouble with Transparency: A Critical Review of Openness in e-Government. *Policy & Internet*, 8(3), 1-30.
- Barry, E., & Bannister, F. (2013). Barriers to Open Data Release: A View from the Top. Presented at the European Group for Public Administration, Edinburgh: The University of Edinburgh. Retrieved from https://www.scss.tcd.ie/disciplines/information_systems/egpa/docs/2013/BarryBannister.pdf
- Bass, G., Brian, D., Fuchs, M., Schwartz, A., McDermott, P., Miller, E., & Weismann, A. (2010). Letter Encouraging the Administration to Improve Its Open Government Efforts.
- Berners-Lee, T. (2010). The year open data went worldwide. TedGlobal. Retrieved August 13, 2013, from http://www.ted.com/talks/tim_berniers_lee_the_year_open_data_went_worldwide.html
- Bertot, J. C., Jaeger, P. T., & Grimes, J. M. (2010). Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. *Government Information Quarterly*, 27(3), 264–271.
- Braman, S. (2006). *Change of State: Information, Policy, and Power*. MIT Press.
- Clarke, Amanda, and Helen Margetts. (2014). "Governments and Citizens Getting to Know Each Other? Open, Closed, and Big Data in Public Management Reform." *Policy & Internet* 6, no. 4: 393–417.
- Davies, T. (2010). Open data, democracy and public sector reform. A look at open government data use from data.gov.uk. Social Science of the Internet. Retrieved September 20, 2013, from <http://www.opendataimpacts.net/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf>
- Davies, T. (2011). Evaluating the Autumn Statement Open Data Measures. Retrieved September 20, 2013, from <http://www.timdavies.org.uk/2011/12/02/3090/>
- Davies, T. (2013). How might open data contribute to good governance? In R. Jones-Parry & A. Robertson (Eds.), *Commonwealth Governance Handbook 2012/13: Democracy, Development and Public Administration* (pp. 148–150). Cambridge, UK: Commonwealth Secretariat.
- Executive Office of the President. (2013). Open Data Policy - Managing Information as an Asset (Memorandum No. M-13-13). Washington D.C. Retrieved from <http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>
- Fioretti, M. (2012). Open Data: Emerging Trends, Issues and Best Practices - A Research Project About Openness of Public Data in EU Local Administration. In G. Bottazzi (Ed.), *Open Data, Open Society*. Pisa: Laboratory of Economics and Management.
- Fox, J. A. (2007). The uncertain relationship between transparency and accountability. *Center for Global, International and Regional Studies*. Retrieved from <http://escholarship.org/uc/item/8c25c3z4>
- Griffiths, Aaron. (2009). "The Publication of Research Data: Researcher Attitudes and Behaviour." *International Journal of Digital Curation* 4, no. 1 (June 29): 46–56.
- Gurstein, M. (2011, July 11). A Data Divide? Data "Haves" and "Have Nots" and Open (Government) Data. Retrieved from <http://gurstein.wordpress.com/2011/07/11/a-data-divide-data-%E2%80%99Chaves%E2%80%9D-and-%E2%80%99Chave-nots%E2%80%9D-and-open-government-data/>
- Horner, C. C. (2012). *The Liberal War on Transparency: Confessions of a Freedom of Information "Criminal."* New York, NY: Threshold Editions.
- Janssen, K. (2012). Open Government Data and the Right to Information: Opportunities and Obstacles. *The Journal of Community Informatics*, 8(2), 1-15.
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, 29(4), 258–268.
- McClean, T. (2011). Not with a Bang but a Whimper: The Politics of Accountability and Open Data in the UK. American Political Science Association Annual Meeting. Retrieved September 20, 2013, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1899790
- McDermott, P. (2010). Building open government. *Government Information Quarterly*, 27(4), 401–413.

- Nahon, K., & Peled, A. (2015). Haunted Ships: An Empirical Examination of Open (Closed) Data. Presented at the Hawaii Conference on System Science (HICSS-48), Hawaii.
- Noveck, B. S. (2009). *Wiki Government: How Technology Can Make Government Better, Democracy Stronger, and Citizens More Powerful*. Washington, D. C: Brookings Institution Press.
- O'Donnell, G. A. (1998). Horizontal Accountability in New Democracies. *Journal of Democracy*, 9(3), 112–126.
- Obama, B. (2009). *Transparency and Open Government* (Memorandum No. M-09-12). Washington D.C. Retrieved from http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_fy2009/m09-12.pdf
- Office of Management and Budget (OMB). (2009). *Open Government Directive* (Memorandum No. M-10-06). Washington D.C. Retrieved from http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf
- Office of Management and Budget (OMB). (2013). Supplemental Guidance on the Implementation of M-13-13 "Open Data Policy – Managing Information as an Asset." Retrieved from <http://project-open-data.github.io/implementation-guide/>
- Peled, A. (2011). When transparency and collaboration collide: The USA Open Data program. *Journal of the American Society for Information Science and Technology*, 62(11), 2085–2094.
- Peled, A. (2013). Re-Designing Open Data 2.0. *JeDEM - eJournal of eDemocracy and Open Government*, 5(2), 187–199.
- Pina, V., Torres, L., & Royo, S. (2010). Is E-Government Leading to More Accountable and Transparent Local Governments? An Overall View. *Financial Accountability & Management*, 26(1), 3–20.
- Piotrowski, S. J. (2007). *Governmental Transparency in the Path of Administrative Reform*. Albany: State University of New York Press.
- Robinson, D., Yu, H., Zeller, W., & Felten, E. (2009). GOVERNMENT DATA AND THE INVISIBLE HAND. *Yale Journal of Law and Technology*, 11(1). Retrieved from <http://digitalcommons.law.yale.edu/yjolt/vol11/iss1/4>
- Rogers, Simon. (2012). "UK Open Government Data: The Results of the Official Audit." *The Guardian*, sec. News. Accessed December 23, 2014. <http://www.theguardian.com/news/datablog/2012/apr/18/uk-open-government-data-national-audit-office>.
- Rosenberg, L. (2013). A Decidedly Different Obama on Transparency. Retrieved September 7, 2014, from <http://sunlightfoundation.com/blog/2013/01/22/a-decidedly-different-obama-on-transparency/>
- Schedler, A. (1999). Conceptualizing Accountability. In A. Schedler, L. J. Diamond, & M. F. Plattner (Eds.), *The Self-restraining State: Power and Accountability in New Democracies* (pp. 13–28). Lynne Rienner Publishers.
- Shkabutar, J. (2012). Transparency With(out) Accountability: Open Government in the United States. *Yale Law & Policy Review*, 31(1). Retrieved from <http://ylpr.yale.edu/transparency-without-accountability-open-government-united-states>
- Wadhwa, Vivek. (2011). "The Coming Death of Open Government." *The Washington Post*, June 21, 2011. http://www.washingtonpost.com/national/on-innovations/the-coming-death-of-open-government/2011/06/21/AGPK3afH_story.html.
- White House. (2012). The Obama Administration's Commitment to Open Government: A Status Report. Retrieved September 20, 2013, from http://www.whitehouse.gov/sites/default/files/opengov_report.pdf.
- Yu, H., & Robinson, D. (2012). The New Ambiguity of "Open Government." *UCLA Law Review Discourse*, 59, 178–208.