Ask Not What Your Country Can Do For Open Data,

Ask What Open Data Can Do For Your Country:

Examining the Rapid Expansion of National Open Data Initiatives Globally

Theo Jaeho Goetemann

Georgetown University Department of Government Honors Program

Defended with Distinction on 02 May 2017

Mentor: Professor Irene A. Jillson

Reader: Professor Diana M. Owen

Abstract

Over the past eight years, governments around the world have created initiatives aimed at fostering public and private use of government data through open data portals. For this thesis I examined the factors driving open data initiatives from four angles: cultures of transparency and accountability, government involvement and support for technology, education attainment and technology access and adoption, and the involvement of international organizations promoting open data initiatives. I employed a mixed-method approach: critical review of literature (qualitative) and analysis of secondary data (quantitative). Data were extracted from publically available datasets created and maintained by international organizations that review open data at the national level. Variables include levels of civil and political liberties, corruption, government participation in international organizations promoting open data, and technology adoption. The findings of the literature review and data analysis evaluate normative claims describing the paths to successful open data initiatives, and point to further qualitative study of the regional, political, cultural, and societal factors that influence national open data portals.

Acknowledgements

This thesis would not have been possible without all of the early morning meetings, rapid and thoughtful email responses, and continual support from Professor Irene Jillson. Our conversations on music, adventures and anecdotes from abroad, while not explicitly written out in this paper, made the early mornings all the better. I am also grateful to the time that Professor Diana Owen has devoted to reading and revising my initial thesis proposal, and for giving even more of her time as my reader two years later. I owe many thanks to Professor Langenbacher for pushing me to develop literary angles of approach, Dr. Nancy Crowell for sharing her expertise and patience while walking me through tests of association step-by-step, and Alexander Podkul for fielding questions on quantitative methods.

Table of Contents

Chapter 1: Introduction

1.1 Overview

In 1999, Alasdair Roberts argued that while "the burden was once on proponents of access rights to make a case for transparency; today, the burden is on governments to make the case for secrecy."[1] A decade later, in 2009, this became a reality. Under President Obama, Data.gov, the first open data portal for federal agencies in the world, went live. By 2014, the number of countries with open government data catalogues rose to 46, and by 2016, the number had risen to 106 countries, more than double the 2014 count.[2] Experts who promote open data and officials who have been involved in implementing open data initiatives have touted the potential civic and economic benefits open data can bring. However, with little empirical evidence, it is unclear why 106 countries have been so quick to divert government money and time towards this largely untested venture, other than a perceived need to strengthen citizen engagement and democracy. Furthermore, although instituting open data initiatives in countries with high levels of government corruption and little citizen input seems counterintuitive, seven countries with open data portals are cross-listed with the World Bank's list of fragile states.[3]

The overall purpose of this thesis is to explore the political, cultural, and societal factors that may have shaped governments' paths to join the global open government data movement.

The remainder of Chapter 1 introduces the research aims and provides definitions of key terms.

---

[1] Alasdair Roberts, "Access to Information: A Key to Democracy," *The Carter Center,* November 2002, pg. 9, https://www.cartercenter.org/documents/1272.pdf.

[2] Wyatt Kash, "Number of Countries Embracing Open Government Surges," *FedScoop,* August 1, 2016, http://fedscoop.com/number-of-countries-embracing-open-government-surges.

[3] Of the 92 countries in Open Data Barometer's open data initiative index, Cameroon, Haiti, Myanmar, Sierra Leone, Sudan, Yemen, and Zimbabwe, are all cross-listed with the World Bank's FY09 list of fragile states. *World Bank,* "FY06 LICUS List (Low-Income Countries Under Stress), http://pubdocs.worldbank.org/en/359521410886172040/FY6toFY9Fragile-States-List-formerly-LICUS.pdf.

1.2 Research Aims

Research Aim 1: Explore the association between the strength of democracy in a country and the strength of its national open data initiative.

- Specifically examining fulfillment levels of FOI requests, political rights and civil liberties, and levels of corruption.

Research Aim 2: Explore the association between the level of support an open data initiative receives from its national government and the strength of the country's national open data initiative.

- Specifically examining open data initiatives that involved the president/prime minister's office and the legal support provided to initiatives.

Research Aim 3: Explore the association between the level of education and e-government familiarity and the strength of a national open data initiative.

- Specifically examining mean years of schooling, internet usage, and e-government familiarity.

Research Aim 4: Explore the association between the involvement of international organizations in a country's open data initiative, the participation of a government in open data organizations and programs, and the strength of its open data initiative.

1.3 Definitions

While the definition of open data is still being shaped, as organizations and programs around the world continue to redefine its criteria, for the purposes of this thesis, I use the 2016 Open Data Handbook (ODH) definition: "Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and share alike."[4] For government agencies, this definition means that data must be openly-licensed, in that both private and public entities and individuals can reuse the datasets published on the government open data portal for commercial and non-commercial purposes. The data cannot have a charge attached to its access or use, its license must allow for modification, and its availability cannot be discriminatory against any person or group. Although ODH's definition has set the standard for what constitutes open data, the definition may continue to change as open data becomes more and more mainstream, and the requirements for open data become more stringent.

While the Open Data Handbook definition is thorough in its definition of openly-licensed data, the actual implementation of ODH's definition requires additional conditions. In order for government data to be useable, it must be machine-readable, meaning files are digital and can be processed easily by a computer,[5] available to download in bulk,[6] and is up-to-date and accurate. These additional terms become important in determining the quality of an open data portal because more sensitive data (e.g. datasets on police officers involved in shootings) may appear

[4] "What is Open Data?" *Open Data Handbook*, 2016, http://opendatahandbook.org/guide/en/what-is-open-data/. The Open Data Handbook was first issued in 2010 and is currently available in 20 languages. Its definition of open data is cited in open data national action plans from many countries around the world, and is a commonly used definition by open data NGOs. The document was created by Open Knowledge International (OKI), an international NGO that works with national governments to develop open data initiatives.

[5] Machine-readable file types include XLS, CSV, JSON, XML. Not machine-readable files include HTML, PDF, DOC, JIF, JPEG, PPT.
"Methodology," *Open Data Index,* 2016, http://index.okfn.org/methodology/.

[6] Data is available in bulk if the whole dataset can be downloaded easily (e.g. zip format). It is non-bulk if users are limited to getting parts of the dataset individually.
Ibid.

open due to their open-license agreement, but the lack of machine-readability or high charge for access may render them inoperable to the user. Corrupt government officials may also attempt to misdirect users away from questionable financial records by denying access or hiding them in badly scanned, non machine-readable formats such as JPEG files. These additional terms become essential in evaluating the open data portals of countries lacking ample financial resources or employees familiar with the necessary computer software.

An open data portal is a web-based platform designed to make it easier to find open datasets. The European Commission likens them to library catalogues, in that open data portals contain metadata records of datasets and house search functionalities to find datasets of interest and of specific qualities (e.g. format, size, content).[7] While private entities or municipal governments may create open data portals, this thesis addresses only the quality of national (public) open data portals and initiatives.

National open data portals are a single point of access for users to find federal government-collected data as well as data collected at the state and municipal level. Open data initiatives refer to broader government programs that include national open data portals, but also other initiatives such as encouraging specific ministries to publish their data, create stronger data-sharing initiatives between the private and public sectors, or fostering a startup culture around an open data portal.

1.4 Organizations Involved

Individuals and organizations are involved in the open data process at all levels: individuals, companies, and nonprofits who act as users of the data; open data advocate groups

[7] "Open Data Portals," *European Commission*, https://ec.europa.eu/digital-single-market/en/open-data-portals.

who monitor the process and support further development of open data initiatives; experts and researchers who examine the factors affecting open data adoption and its impacts; and NGOs/IGOs who, in addition to all of these levels, also work at the government level, creating new initiatives and setting guidelines for existing ones.

At the user level, open data can be incorporated in businesses' existing products and services to amplify the scope of businesses' available information. New companies and organizations often appear alongside open data portals, and the potential economic benefits of open data are a frequently touted figure in arguments for open data. Open data can also be leveraged by individuals or government-monitoring groups to reduce corruption, track spending and government contracts, etc. The Center for Open Data Enterprise, which records cases of open data use by both nonprofit and for-profit organizations, has logged over 1,766 organizations worldwide using open data.[8]

At the advocacy/expert level, prominent open data organizations include Open Knowledge International, founded in 2004.[9] OKI (previously Open Knowledge Foundation (OKF)) is the parent organization for the Global Open Data Index (GODI),[10] the Open Data Handbook,[11] and the Open Definition.[12] The World Wide Web Foundation has also sponsored a great deal of research in open data, and is responsible for producing the Open Data Barometer, an index that evaluates and scores the strength, readiness, implementation, and impact of national open data initiatives globally. There are many more organizations under the open data banner,

[8] Center for Open Data Enterprise, "Open Data Impact Map," 2016, http://www.opendataimpactmap.org/map.html.
[9] Rufus Pollack, *Open Knowledge International Blog,* "Open Knowledge Foundation Launched," May 24, 2004, https://blog.okfn.org/2004/05/24/open-knowledge-foundation-launched//.
[10] One of several indexes assessing and scoring the strength of open data initiatives globally.
[11] Set open data guidelines for a government body, whether it be at the local or national level, to successfully guide the public entity towards a successful portal.
[12] The commonly accepted definition for many open data organizations. Of the organizations already (and soon-to-be) mentioned, the Center for Open Data Enterprise, GODI, and the Open Data Barometer, all accept the "open definition" as the definition for openness.

such as the UK-based Open Data Institute, the Center for Open Data Enterprise, the Sunlight Foundation, and others. Many of these organizations have contributed to literature on open data, partnering with corporate sponsors like Deloitte or Mckinsey, with foundations like Omidyar, the Bill and Melinda Gates Foundation, or with government sponsors like USAID, UKAID, the Government of Canada. They have also consulted directly with governments who are considering starting open data initiatives; the Center for Open Data Enterprise, for example, worked with the Government of Jamaica on the launch of its first portal in June, 2016.[13]

[13] *The Caribbean Open Institute,* "Open Data for Business in Jamaica: Initial Findings and Recommendations," http://caribbeanopeninstitute.org/od4b_roundtable_jamaica.

Chapter 2: Findings of the Critical Review of Literature

2.1 Overview

In this section, I will first provide a brief summary of the history of open data. Second, I will examine the current state of open data around the world. Third, I will discuss factors promoting open data adoption. Existing literature on open data does not often explicitly address the factors involved in creating open data initiatives but rather focuses on policy implications and benefits of open data. Therefore, while this section functions as an overview of open data, it also is used to illustrate the common circular argument that if a democratic government's function is to foster transparency and innovation, then the primary motivating factor behind the creation of open data initiatives is this normative interpretation of the function of a state.

2.2 History of Open Data

The concept of open data existed long before President Obama's 2009 open government initiative. The ideological parent of open data is Freedom of Information (FOI) legislation, the first of which, the Freedom of the Press Act, was passed in 1766 in Sweden.[14] The idea that it is a fundamental right of a citizen to be guaranteed public access to documents from government agencies is at the crux of arguments advocating for open data. Although open data initiatives today, especially those from other regions than Western Europe and the United States, stem from other origins as well, such as a government's desire to appear accountable to its trade partners or foster rapid economic growth through the development of a city's startup culture.

[14] Lennart Weibull, "Freedom of the Press Act of 1766," *Encyclopaedia Britannica,* December 21, 2015, https://www.britannica.com/topic/Freedom-of-the-Press-Act-of-1766.

Early open data initiatives can also be seen in the medical and research communities. Since 1958, with the formation of the World Data Center,[15] these communities have worked to improve data sharing and accessibility and reduce data silos.[16] Openness in the medical research community is still a relevant topic today; medical research organizations have become increasingly pressured to share data with their peers, as proponents argue sharing has the potential to decrease error through greater peer-data review, as well as increase the possible solutions through greater information access for all in the medical field. In the U.S., former Vice President Joe Biden's Moonshot Initiative, an initiative announced by the Obama Administration in 2016 to find a cure for cancer, is a prime example of opening medical research data, as it takes a large step towards eliminating data silos managed by academic and other organizations (including those funded primarily by the public sector) that conduct cancer research.[17] The Initiative, for which funding of $1 billion dollars was requested, is meant to drive solutions through the cutting of red tape and by bringing researchers and their data together rather than larger amounts of funding.[18]

Also in the United States, both THOMAS.gov, a website created in 1995 by a Republican-controlled House that provided information to the public about pending legislation in Congress, and the 1994 memorandum from the Clinton Administration to executive branch

[15] The World Data Center (WDC) was created in 1958 "to archive and distribute data collected from the observational programs of the 1957-1958 International Geophysical Year." It has expanded past its original founders (US, Europe, Russia, and Japan) to encompass other countries. In 2008, the WDC was reformed into the World Data System.
"World Data Centers," *NOAA*, https://www.ncdc.noaa.gov/customer-support/world-data-centers.
[16] Data silos are databases in which data is 'trapped.' Usually, this means that the data is proprietary and held by a company or organization. Proponents of open data argue that exponential benefits come from opening silos and letting the public and new eyes view and leverage the data.
"Breaking Down Data Silos to Achieve Better Results in the Public Sector," *OpenDataSoft,* February 21, 2017, https://www.opendatasoft.com/2017/02/21/breaking-data-silos-achieve-better-results-public-sector/.
[17] Sarah Karlin-Smith et al, "Biden's Farewell Gift: Cancer Moonshot Helps Pass $6.3 Billion Research Bill," *Politico,* December 07, 2016, http://www.politico.com/story/2016/12/joe-biden-cancer-moonshot-bill-232342.
[18] Ibid.

agencies asserting the public's right to know through information technology are early examples of government open data initiatives akin to the web-based ones with we are familiar today. However, to the average citizen consumer globally, open data was first experienced in the form of weather reports, which use raw data from the U.S. National Weather Service, with the data initially disseminated in the interest of public safety, to warn of impending natural disasters.[19]

2.3 Applications of Open Data in the Private Sector

In the private sector, open data is integrated in companies' products and services across all sectors of business. It is used extensively in companies such as Zillow, which overlays its own customer-collected data with U.S. market and Census data to provide users with greater information when moving or purchasing a house. Other examples outside the United States include Farmerline, an SMS-based service that texts farmers in Ghana up-to-date agricultural and weather information. This allows farmers to yield larger and steadier harvests due to the agronomic advice, as well as up-to-date knowledge of competitive pricing.[20] Haezoom, a South Korean company, uses "building, geographical, solar radiation, air temperature,"[21] and other datasets from the South Korean government's national data portal to develop accurate estimates of how much a homeowner can save by installing solar panels. In the art world, the Michael Culture Association uses open data to bring together art pieces from over 100 public and private

[19] Joshua Tauberer, "History of the Movement," *Open Government Data: The Book (2nd Edition), Lulu Press, Inc.* October 6, 2014, https://opengovdata.io/2014/history-the-movement/.
[20] "Highlighting Use Cases From the Open Data Impact Map: Farmerline Empowering Farmers Through Mobile Technology and Information," *Center for Open Data Enterprise,* September 23, 2016, http://opendatacon.org/highlighting-use-cases-from-the-open-data-impact-map-farmerline-empowering-farmers-through-mobile-technology-and-information/.
[21] "Highlighting Use Cases From the Open Data Impact Map: Haezoom Accelerating the Adoption of Solar Energy in South Korea," *Center for Open Data Enterprise,* September 28, 2016, http://opendatacon.org/highlighting-use-cases-from-the-open-data-impact-map-farmerline-empowering-farmers-through-mobile-technology-and-information/.

organizations throughout Europe and host them for viewing under a single database. Through this network of organizations, the Michael Culture Association has also physically brought together individuals interested in digital preservation and digital cultural heritage in workshops and group meetings.[22] Another example are the visualizations Ben Wellington made of NY parking tickets. In 2016, Wellington analyzed and visualized parking ticket geospatial data available on the New York City open data portal. Through his visualization, he was able to identify millions of dollars in "erroneous parking tickets, issued by NYPD officers to vehicles that were in fact parked legally."[23] In the following weeks, the NYPD sent training messages to all officers on parking ticket rules and created steps to digitally monitor parking tickets to ensure they are being issued correctly.

[22] "About," *Michael Culture Association*, *http://www.michael-culture.eu/about*.
[23] Max Galka, "How an Open Data Blogger Proved the NYPD Issued Parking Tickets in Error," *The Guardian*, July 26, 2016, https://www.theguardian.com/cities/2016/jul/26/open-data-blogger-parking-tickets-new-york-nypd.

2.4 Current State of Open Data

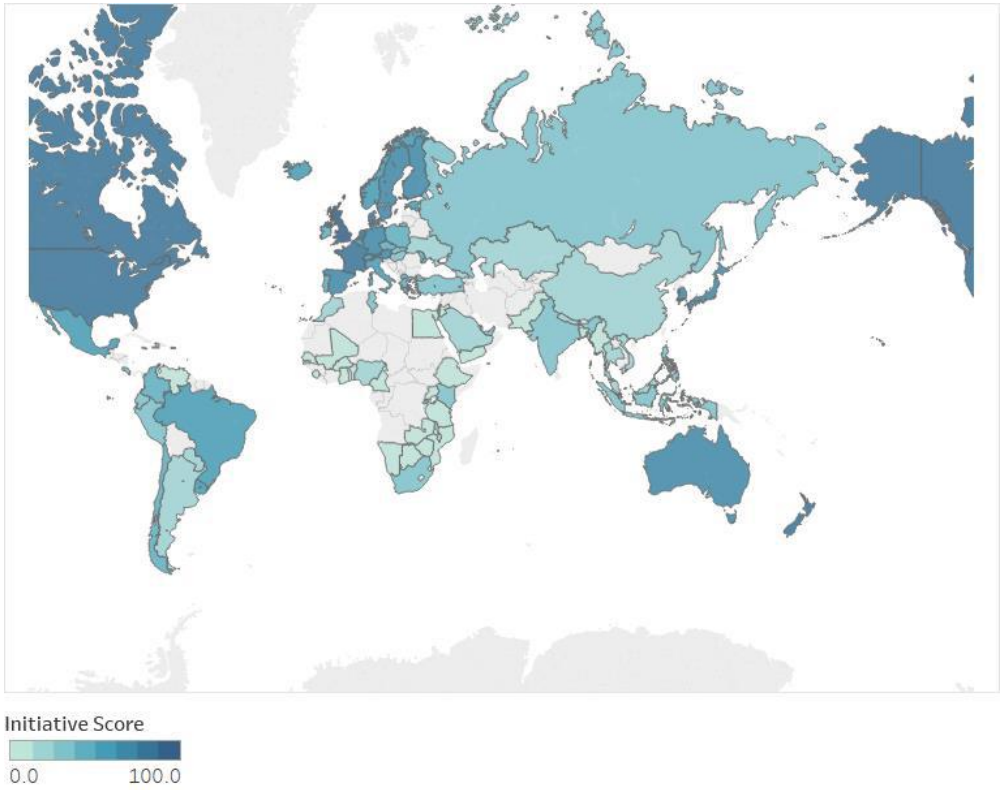


Figure 1: Open Data Around the World, 2015[24]

As researchers at Open Data Barometer commented, "Open data is entering the mainstream."[25] They are referring not only to the rapid proliferation of open data initiatives globally, but also to the fact that many open data initiatives have reached a stage of maturity such that the primary goal is no longer simply to support open data publishing, but to develop data-minded user communities around data portals. These communities are reflected in the innovative applications of open data emerging in the private sector, in the uptake of interest in open data meetups and hackathons, and in the growth of the open data user community online. In countries that have supported open data initiatives longer than most, open data has spread to the local

[24] Data source: *Open Data Barometer 3rd Edition,* "ODB-3rdEdition-Datasets-Scored.csv," http://opendatabarometer.org/3rdEdition/data/.
[25] *Open Data Barometer,* "Global Report," 2015, 7, http://opendatabarometer.org/doc/3rdEdition/ODB-3rdEdition-GlobalReport.pdf.

level, with open data communities built around city and state portals in addition to national portals.[26] At the same time, new open data initiatives are underway in a number of countries, including Ecuador, Nepal, Thailand, Ethiopia, Nigeria, and Uganda. Online, more and more websites are being built not only to host open data, but also to create a data-sharing culture. A prime example is data.world,[27] a startup founded in 2014, whose beta version was launched in late 2016. The platform, a cross between a data portal and social media site, allows users to upload datasets, collaborate online, track topics and trending datasets, and participate in public projects.

Yet, while all of these developments relating to open data globally hint at developments in depth as well as breadth, governments involved, users, and advocates all face challenges in ensuring that the implementation of open data policies is carried out to the fullest extent possible. The issue of "open washing," in which an organization or government creates a data initiative to amplify their credentials of openness and transparency, but while their public statements may articulate openness, their portal lacks data that fully meets the criteria for open data. Open Data Barometer (ODB) addresses the issue of "open washing" in a section of its 2015 global report entitled, "Civil rights and the role of citizens." ODB's writers argue that strong, accountable institutions, robust privacy laws and freedom of information legislation, are all foundational prerequisites of a successful open data initiative: "if these factors aren't in place, open data initiatives risk simply being window-dressing, or 'open washing.'"[28] ODB argues that by

[26] Jacob Perez, "Open Data in Mexico: The Search for Civic, Private, and Public Sector Innovators," *Medium*, May 18, 2016, https://medium.com/@canada2020/open-data-in-mexico-the-search-for-civic-private-and-public-sector-innovators-a974547d1c22.
Noelle Knell et al, "Open Data Policies in State and Local Government," *GovTech*, March 17, 2014, http://www.govtech.com/data/are-governments-committed-to-open-data-interactive-map.html.
Karolis Granickas, "Open City: Local Government & Open Data," *European Data Portal*, August, 2015, https://www.europeandataportal.eu/sites/default/files/2015_open_city_local_government_and_open_data.pdf.
[27] "Understanding data.world's features," Data.world, https://data.world/features-overview.
[28] "Global Report," *Open Data Barometer*, 2015, 14.

avoiding the issue of "open washing," the likelihood that media and citizens will seek out information and become active movers in an open data initiative increases.[29]

However, the development of open data initiatives in countries such as China, where "the chance of a citizen winning an Open Government Information (China's version of FOI)-related lawsuit in Beijing is 5%,"[30] show that there is more nuance to ODB's claim that accountable institutions with robust freedom of information legislation are a prerequisite to data initiatives. Certainly, China does not have a strong data initiative, but its city portals and Big Data Development Action Plan were all still created in the last three to four years.[31] It is clear that China does not plan to adapt its open data initiative to an open government initiative, and Article Eight of its OGI legislation limits information requests to discloses that do not endanger "the state, public, or economic security or stability" of China.[32] Yet China's initiative shows promise; over 3,000 participants attended Shanghai's Open Data Innovation 2016 Challenge and were given 1,000gb of municipal government data to use.[33]

In this example, China represents a new wave of open data adopters, countries whose motivations in developing an open data initiative do not adhere to conventional transparency-building, democratic frameworks of understanding where open data portals stem from. The Open Data Barometer identifies "new challengers" to the traditional leaders of open data as France, Canada, South Korea, Mexico, Uruguay, and the Philippines.[34]

---

[29] Ibid.
[30] "Right to Know," *The Economist,* May 3, 2014, http://www.economist.com/news/china/21601564-leaders-discover-some-transparency-can-help-make-society-more-stable-right-know.
[31] Ma, Jinxin Yolanda, *Global Investigative Journalism,* March 28, 2017, http://gijn.org/2017/03/28/why-open-data-is-good-for-china/.
[32] "Right to Know," *The Economist,* May 3, 2014.
[33] Ma, Jinxin Yolanda, *Global Investigative Journalism,* March 28, 2017.
[34] *Open Data Barometer,* "Global Report," 2015, 14.

Additionally, based on rank changes from 2014 to 2015, countries considered by Open Data Barometer (ODB) and Global Open Data Index (ODI to be traditional open data supporters as well as those who are considered a new generation of open data adopters are listed below, in Tables 1 and 2 respectively.

ODB Position	ODB Rank Change	GODI Position	GODI Rank Change (from 2014)	Country	Income	HDI
1	0	2	-1	UK	High Income	Very High
2	0	8	0	USA	High Income	Very High
5	+4	3	-1	Denmark	High Income	Very High
7	-1	8	+9	Netherlands	High Income	Very High
9	-6	27	-13	Sweden	High Income	Very High
10	0	5	0	Australia	High Income	Very High
11	+1	5	-1	Finland	High Income	Very High

Table 1: Traditional Open Data Supporters[35]

ODB Position	ODB Rank Change	GODI Position	GODI Rank Change (from 2014)	Country	Income	HDI
-	-	1	+10	Taiwan	High Income	-
2	+2	10	-7	France	High Income	Very High
4	+3	17	+5	Canada	High Income	Very High
8	+9	23	+5	South Korea	High Income	Very High
16	+8	13	+15	Mexico	Upper-Middle Income	High
19	+6	7	+6	Uruguay	High Income	High
28	+12	4	+8	Colombia	Upper-Middle Income	High
17	+4	12	+14	Brazil	Upper-Middle Income	High

Table 2: Next Generation of Open Data Adopters[36]

[35] Table is kept to around the top 10 for both indexes. I cross-listed to determine inclusion: New Zealand (ODB rank = 6, GODI rank = 123), Norway (ODB rank = 17, GODI rank = 10) were not included based on their ranking in one of the indexes.

[36] Table was cross-listed between both indexes to determine inclusion. Although mentioned as a country part of the next generation of open data adopters in ODB's 2015 global report, the Philippines was not included in the table due to its rankings (ODB rank = 36, GODI rank = 78).

These tables indicate that successful open data initiatives are most common in high income countries. ODB notes this in the executive findings section of their 2015 global report, stating that twenty-six of the top thirty countries in their rankings are high-income countries, and that the "the gap between data have and have-nots needs urgent attention."[37] ODB goes on to cite the UN Independent Expert Advisory group on a Data Revolution for Sustainable Development (IEAG), which further argues that gaps in these initiatives could 'create a whole new inequality frontier,' if they persist going forward.[38]

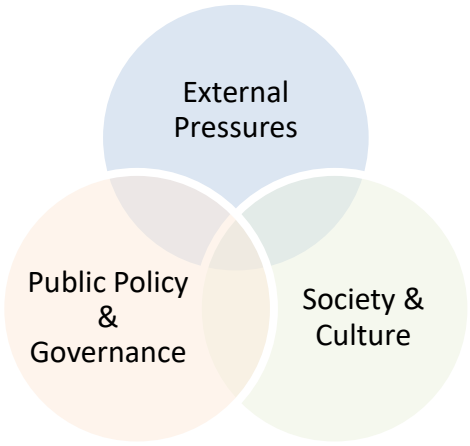
2.5 Path Towards a Portal

There is no single pathway to creating an open data initiative or portal. While literature on open data frequently cites transparency, anti-corruption, economic, and smart city motivations, there is a host of possible reasons that a president, prime minister, legislature, or government ministry may choose to create an open data initiative. Figure 2 presents general factors that may affect the choice to create an open data portal.

A number of factors may detract from the creation of an open data initiative, including several factors listed in Figure 2. For example, while anti-corruption initiatives may support open data adoption as well as foster a culture of transparency in a given country, corrupt officials most likely would not want to create open data initiatives and increase risk of exposure. Similarly, interest groups that benefit from closed doors to their data may pressure a government to sideline

[37] *Open Data Barometer,* "Global Report," 2015, 8.
[38] *Open Data Barometer,* "Global Report," 2015, 8.
UN Independent Expert Advisory group on a Data Revolution for Sustainable Development (IEAG), "A World That Counts: Mobilizing the Data Revolution for Sustainable Development," 2015, http://www.undatarevolution.org/report/.

open data initiatives. An example of this is the extractive industry, whose firms often lobby to keep their government filings private.[39]



External Pressures:
- Interest groups
- Companies
- International organizations
- Global view of country and international org membership
- City/state government self-created data initiatives

Public Policy & Governance
- Public approval
- E-government initiatives
- Transparency & anti-corruption initiatives
- Economic innovation initiatives
- Smart city initiatives

Society & Culture:
- Cultures of transparency and accountability
- Education attainment
- Perceived responsibilities of government (citizens, officials)
- Views of technology and innovation (citizens, officials)

Figure 2: Path Towards a Portal
Goetemann, 2017

Combatting this clampdown on information, open data organizations such as Open Oil,[40] a Berlin-based NGO whose platform allows users to search over 2 million public domain documents filed by oil, gas, and mining companies around the world, have specifically developed around the unwillingness of the extractive industries to share data. Other factors that may also detract from an open data initiative are cultures of secrecy, unwillingness of government officials at the subnational level to support the national portal and publish their own data, and disinterested or unaware citizens.

[39] "Open Data: the Extractive Industries Case-Study," *Sherpa*, December 8, 2016, https://www.asso-sherpa.org/open-data-the-extractive-industries-case-study.
[40] Open Oil, http://openoil.net/.

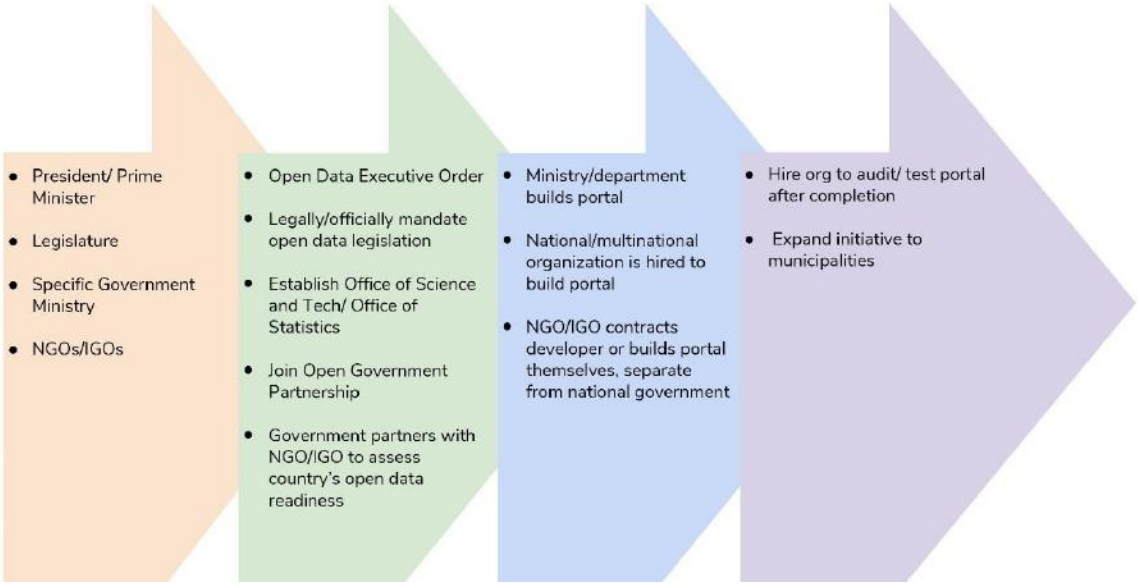


Figure 3: Path Towards a Portal cont.
Goetemann, 2017

Figure 3 shows a typical process in creating an open data portal. Early legal support, in the form of executive orders and legally mandating open data legislation may be factors in the eventual strength of a data portal; both require departments and ministries to recognize the initiative and support it by publishing department/ministry data. It is also worth noting the role that NGOs/IGOs play in the creation of a data portal. In each step, NGOs/IGOs can influence the path of a data portal. They can lobby for the creation of a data portal, partner with governments to create readiness reports before the actual creation of an initiative, as well as be the ones to actually create the portal. There are cases where an NGO/IGO builds a portal without the express support and collaboration of a government, such as portals created by the East-West Management Institute in Southeast Asia.[41] Cases like this, similar to national portals created without the support of a governments' ministries, raise the question of legitimacy in publishing data to these portals, as well as legitimacy's effect on the eventual strength of the portal.

[41] "Open Development – Vietnam," *Open Development Initiative,* https://vietnam.opendevelopmentmekong.net/.

Finally, it is worth noting that while there are many cultural and societal factors that affect the creation of an open data initiative, there are also economic and political factors. For example, the choice of who builds a data portal, the government, a private company or organization, or a multi-national organization, may impact the format of the website, as well as the breadth and depth of the datasets it hosts. While there is a lack of research on this topic, the choice of who creates the data portal is a topic to consider when evaluating data portals globally. For example, France's choice to hire a small company named Etalabs to build its portal, while much of Armenia's open data initiatives are driven by the World Bank's BOOST program may affect the eventual outcomes of each portal.[42]

2.6 Factors Influencing Open Data Adoption

Despite media enthusiasm over open data, the vast majority of national open data initiatives and national open data portals are less than a decade old, and little literature exists on the factors and motivations behind open data initiatives. In fact, the United States, one of the first adopters of open data for federal agencies, only recently conducted its first series of Open Data Roundtables[43] in 2016 to identify case studies and best practices in government open data.[44] Existing literature on open government data is minimal, and the majority of literature published on open data has been sector specific or conceptual papers.[45] Sector specific literature includes discussions on best practices for shared data in fields like agriculture, medicine, or insurance.

[42] "New Initiative Aimed at Increasing Trade and Investments in Armenia," *International Finance Corporation*, http://www.ifc.org/wps/wcm/connect/news_ext_content/ifc_external_corporate_site/news+and+events/news/new+initiative+aimed+at+increasing+trade+and+investments+in+armenia.
[43] A series of events hosted by the White House and the Center for Open Data Enterprise that brought open data users, experts, and advocates from across the private and public sectors together to discuss (1) protecting privacy (2) improving data quality (3) applying research data (4) public-private collaboration.
[44] "Open Data Roundtables," *Data.gov*, March 4, 2016, https://www.data.gov/meta/open-data-roundtables/.
[45] M. Janssen, Y. Charalabidis & A. Zuiderwijk, "Benefits, Adoption Barriers and Myths of Open Data and Open Government," *Information Systems Management (ISM)*, 2012, vol. 29, no.4, pp. 258-268, pg. 3, http://www.tandfonline.com/doi/abs/10.1080/10580530.2012.716740.

These discussions generally address intricacies such as shared information practices between private companies and research facilities and do not broaden out towards a comprehensive discussion of open data federally, much less globally. The major sources of literature on government open data are NGOs, private companies examining the potential economic benefits of open data, and the World Bank. The literature generally comprises descriptions of and justification for open data, policy recommendations, and/or theoretical economic models supportive of open data legislation. Thus, difficulty arises in separating causal factors as to why open data policies are being implemented globally from normative arguments that define government's function as a supporter of transparency and innovation.

For developed countries with robust representative democracies, commonly cited reasons for open data legislation revolve around the concepts of transparency and innovation. For example, when President Obama announced the expansion of the 2009 open data policy, he described America's spirit of innovation and history of government transparency.[46] Similarly, when a bipartisan group of legislators announced the OPEN Government Data Act in April, 2016, they described how the Act "will empower the government to be more effective, private sector to innovate, and public to participate."[47] These reasons are mirrored in the European Council's charter for the EU open data portal,[48] in northern European governments' stated open data initiatives,[49] as well as in other developed representative democracies like France or

[46] "Executive Order – Making Open and Machine Readable the New Default for Government Information," Office of the Press Secretary, *The White House*, May 09, 2013, https://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-.
[47] Christian Hoehner, "The OPEN Government Data Act: A Sweeping Open Data Mandate for All Federal Information," *Data Coalition*, April 18, 2016, http://www.datacoalition.org/the-open-government-data-act-a-sweeping-open-data-mandate-for-all-federal-information/.
[48] "Open Data," *Council of the European Union, http://www.consilium.europa.eu/en/general-secretariat/corporate-policies/transparency/open-data/.*
[49] "Norway," *Open Data Partnership,* http://www.opengovpartnership.org/country/norway.
"Finland," *Open Data Partnership*.

Taiwan.[50] For these countries, the open government movement generally stems from three distinct communities: open government/FOI advocates, open scholarly and medical data advocates, and open innovation entrepreneurs.[51] As information/communications technology and government data collection methods improved over the past decade, these three communities all played a part in lobbying for formalized open data policies. For example, the executive and legislative branches of government both have an incentive to foster civic participation and accountability, or at least appear to be actively supporting programs like open data initiatives. This can be due to maintaining popular support among constituents or keeping rival political parties in check when their members hold office. Therefore, for countries with robust representative democracies, the two sides of the circular factors/government functions argument actually overlap; it is fair to argue that some of the factors behind the creation of open data initiatives overlap with the normative argument that a government's role is to foster transparency and innovation.

However, in the case of weak democracies, in which government legitimacy depends on direct involvement by organized interests rather than ordinary citizens,[52] it is less clear why government officials would push forward with open data legislation. Countries with weak democracies are more susceptible to fragility or failure, and include countries like Guatemala or Kenya. Due to minimal accountability of officials, the more expected route for these countries to take would be that of Italy or Turkey. Both Italy and Turkey initially joined the Open

---

"France Becomes 64th Country to Join the Open Government Partnership," *Open Government Partnership,* April 24, 2014, http://www.opengovpartnership.org/blog/blog-editor/2014/04/24/france-becomes-64th-country-join-open-government-partnership-press.

[50] Ralph Jennings, "How Taiwan Fostered the World's Most Open Government," *Forbes,* December 15, 2015, http://www.forbes.com/sites/ralphjennings/2015/12/15/how-taiwan-rose-to-world-no-1-in-government-transparency/#54d5020a7179.

[51] Joshua Tauberer, "History of the Movement."

[52] Viktor Mayer-Schonberger and David Lazer, *Governance and Information Technology: From Electronic Government to Information Government,* MIT Press (2007), 137.

Government Partnership (OGP), a multilateral initiative comprised of 75 participating countries that have all endorsed a high-level Open Government Declaration and delivered a country open data action plan; these countries are eligible for grant funding and external aid. Italy took part for over a year, Turkey for over two. In 2013, the Italian Data Protection Authority took drastic steps away from introducing FOIA-like legislation, and prohibited indexing and spidering[53] by web search engines.[54] In September 2016, the Steering Committee of OGP rescinded Turkey's membership for failing to deliver a national open data action plan since 2014.[55] Yet both Guatemala and Sudan, along with Lebanon, Ukraine, and other DDs, weak democracies, and strong states all have government-created open data initiatives.[56] Thus, normative conceptions of government-funded innovation and transparency projects are less likely to be the factors leading to open data initiatives in these countries, and further study is necessary to determine corollary variables that explain these countries' participation in OGP and the open government data movement.

Open data initiatives in China— a tightly run, single party state— further confound the hypothesis that governments seeking to increase transparency foster open data policies. In May 2016, Premier Li Keqiang stated that the Chinese government should give the public and businesses greater access to data. This statement was in line with a speech he gave two years ago

[53] Running a script that crawls through a webpage. Oftentimes used for web indexing, spidering can also be used to systematically pull information from web pages and generate this information in machine-readable formats.
[54] Stefano Costa, "Italian Government Ditches Transparency and Open Data," *Open Knowledge International,* March 1, 2013, http://blog.okfn.org/2013/03/01/italian-government-ditches-transparency-and-open-data/.
[55] "Turkey Made Inactive in the Open Government Partnership," *Open Government Partnership,* September 21, 2016, http://www.opengovpartnership.org/blog/open-government-partnership/2016/09/21/turkey-made-inactive-open-government-partnership.
[56] "Place Overview," *Open Data Index,* http://index.okfn.org/place/.
Did you know, if type in "Ukraine" on Google, the top four predictive results are: "Ukraine," "Ukraine is weak," "Ukraine time," and "Ukraine corruption."

in which he announced his vision for a new, entrepreneurial and innovative China.[57] Over the past several years, the Shanghai municipal government has brought the Premier's plan into fruition, opening its own open data portal two years ago and offering reimbursement pledges-- up to 60% for investment losses in a seed stage startup-- to venture capitalists investing in Shanghai startups. Interestingly, the Premier's vision for China increasingly runs contrary to President Xi Jinping's recent crackdowns on Internet freedoms. This tension over releasing government functions to the public body— at the very center of the Party's consolidation of power— highlights one of the larger trends in the open government data movement: open data initiatives for economic growth.

Contrasting this narrative of economic growth, another framework to understand the development of open data initiatives is through the notion of public value. Drawing on Barry Bozeman's definition of "publicness" as "the principles on which governments and policies should be based," Carla Bonina redefines public values within the context of open data as "those basic building blocks of the public sector."[58] The link she draws between public value, an intangible yet fundamental building block of government, is more akin to traditional conceptions of open data's role as a facilitator of democracy rather than the neoliberal lens of economic growth that open data is also viewed through.[59]

---

[57] "President Xi Stresses Innovation to Bolster Economy," *Xinhua Net,* May 10, 2016, http://news.xinhuanet.com/english/2016-05/10/c_135348451.htm.

[58] Carla M Bonina, "New Business Models and the Value of Open Data: Definitions, Challenges, and Opportunities," pg. 11.
Barry Bozeman, *Public Values and Public Interest: Counterbalancing Economic Individualism,* Georgetown University Press, 2007.

[59] James Manyika et al., *McKinsey Global Institute,* "Open Data: Unlocking Innovation and Performance with Liquid Information," October, 2013, http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information.
Tim Cashman, *Socrata,* "The Economic Impact of Open Data," February 27, 2014, https://socrata.com/blog/economic-impact-open-data/.
Omidyar Network, "Open for Business: How Open Data Can Help Achieve the G20 Growth Target," June, 2014, http://www.omidyar.com/sites/default/files/file_archive/insights/ON%20Report_061114_FNL.pdf.

A 2006 study conducted by the Center for Technology in Government (CTG) and SAP AG examined public value from a similar standpoint, conducting research on the non-financial benefits of government investment in information technology.[60] Their new methodology, termed "Public Return on Investment (ROI)," demonstrated how "assessments of public value creation can extend" past the economic realm and show gains in governmental integrity and transparency for both constituents and politicians. While the study defined IT quite broadly, as anything from simple websites to government-wide information systems, its attempt to measure public value outside the economic realm sets an academic backdrop for the current shift in public values the open government data movement represents.

Janssen and colleagues also identified this shift from a closed to an open system of governance[61] and compared the benefits of open data with the barriers to its adoption. They argue that opening federal data leads to two assumptions about the government. First, public agencies are willing to begin a discourse with the public over feedback and opposing inputs. Second, governments are willing to give up control of the asymmetric information gap that previously existed between the government and the public. As did Bonina, Janssen and colleagues acknowledge that most research on open data "consists of conceptual papers, descriptions of the empirical uses of open data, or the design of technology and systems for harnessing the power of open data."[62]

A similar study by Noor Huijboom and Tijs Van den Broek examined the key motivations in implementing open data initiatives in the United States, the United Kingdom,

[60] "SAP Advances Effort to Help Public Sector Organizations Measure Political, Social and Operational Value of IT Projects," *SAP*, September 13, 2006, http://global.sap.com/news-reader/index.epx?pressid=6714.
[61] M. Janssen, Y. Charalabidis & A. Zuiderwijk, "Benefits, Adoption Barriers and Myths of Open Data and Open Government. Information Systems Management (ISM)," 2012, vol. 29, no.4, pp. 258, http://www.tandfonline.com/doi/full/10.1080/10580530.2012.716740.
[62] Ibid, 259.

Spain, Denmark, and Australia.[63] They then conducted a survey among policy-makers and experts within the five countries to determine the top barriers to open data policy implementation. Huijboom and Tijs Van den Broek's study follows a common pattern in open data literature and research: the circular argument fallacy of confounding the motivations behind the creation of open data initiatives with the normative claim the government should promote transparency and innovation. For example, the United Kingdom's key motivation in their study was, "Ultimately a more informed citizen is a more empowered citizen…Data can also be used in innovative ways… by releasing untapped enterprise and entrepreneurship."[64] While a bold, normative claim, this key motivation is both the result and the reason. Additionally, Huijboom and Tijs Van den Broek's study employs a qualitative approach to assessing open data barriers and motivations. This approach was most recently utilized by the United States government in the Open Data Roundtable discussions, which were essential in informing the recently published Open Data Transition Report for the next administration.[65] Huijboom and Tijs Van den Broek acknowledge the weaknesses of their methodology, stating that, "All in all, one has to conclude that evidence of economic, social and democratic impacts of open data policy is still immature or lacking."[66]

The open government data movement faces several other challenges in addition to the paucity of objective assessment data. For developed countries with robust democracies, government open data initiatives run the risk of being freely available, but only to the trained few who can navigate raw datasets and transform them into meaningful analysis. Michael Gurstein

[63] Noor Huijboom, & Tijs Van den Broek, "Open Data: An International Comparison of Strategies," *European Journal of ePractice,* April 2011, pg. 3, http://www.tandfonline.com/doi/full/10.1080/10580530.2012.716740.
[64] Ibid, 4.
[65] "The 2016 Open Data Roundtables," *Data.gov,* https://www.data.gov/meta/open-data-roundtables/.
Joel Gurin, & Laura Manley, "Open Data Transition Report: An Action Plan for the Next Administration," *Center for Open Data Enterprise,* October, 2016, http://opendataenterprise.org/transition-report.
[66] Noor Huijboom, & Tijs Van den Broek, "Open Data: An International Comparison of Strategies," pg. 10.

addresses this issue, discussing how the digitalization of land records in Bangalore was primarily being used by middle and upper-income people and corporations to "gain ownership of land from the marginalized and the poor."[67] Gurstein goes on to examine a mirroring problem in Nova Scotia, where individuals lose land on which they have lived for years with clear, documented title to the land. When the land becomes publicly "visible" via GIS, those who have an interest in its ownership (e.g., timber companies, extracting companies) can initiate claims of ownership.[68] Although these are extreme examples, they still do represent individuals' first exposure to open data in other places around the world, and demonstrate the harm that open data can cause if misused. Aside from land ownership, the initial knowledge barrier to working with data formats like JSON and software requirements to manipulate datasets limit general public access. Castells similarly argues that as societies move into an information age, the ability to participate in a network becomes essential for fair play. If open data is meant as a service for the public, corresponding efforts should be made to lower the knowledge barrier in leveraging data and navigating APIs.[69] Castells argument extends not only to frequent open data users, but to the average technology user as well, whether that be a journalist, individual involved in local government, or interested citizen.

For low and lower-middle income countries, another problem stems from external pressure from international organizations such as the World Bank to create an open data portal. While the pressure may force strong states and corrupt officials' hands in creating a national

[67] Michael Gurstein, "Open Data: Empowering the Empowered or Effective Data Use For Everyone?" *First Monday*, February 7, 2011, http://firstmonday.org/ojs/index.php/fm/article/view/3316/2764#p3.
[68] Ibid.
[69] Ironic in a 2016 literature review about open data; one of the primary critiques of Castells' network society was his belief that the world moving into the information age. Critics argued that the world instead was merely extending its time in the industrial age.
Manuel Castells, *The Rise of the Network Society: The Information Age: Economy, Society, and Culture Vol. 1*, Blackwell, 1996, pg. 165.

open data portal, with minimal maintenance and initial costs to create the portal, the time and financial resources invested in the portal would be better spent elsewhere.[70]

2.6.1 FOI and Cultures of Transparency and Accountability

FOI legislation has long been associated with efforts to strengthen democracy. FOI law gives citizens the right to access government documents without being "obliged to demonstrate any legal interest or 'standing.'"[71] In other words, it shifts government information from a need-to-know basis to a right-to-know basis.[72] Currently 114 countries have freedom of information laws or similar administrative regulations.[73] In many ways, open government data is the ideological follow-up to FOI legislation. A key distinction between the two is the intended result of each: FOI laws are primarily intended to increase intra-state government accountability and transparency. Open government data's purpose, on the other hand, is less clear and still evolving, but while it encompasses greater government accountability and transparency, it also has more direct economic incentives than FOI legislation, and a government with open data policies may benefit from higher levels of international investment or greater ability to join multinational organizations due to transparent financial datasets. An example of this, one which predates the existence of national open data portals, is the case of China joining the World Trade Organization (WTO) in 2001. In order to join, China had to agree to economic and trade-related transparency commitments, opening data that the Chinese government had kept to itself.[74]

---

[70] Manuel Stagars, *Open Data in Southeast Asia: Towards Economic Prosperity, Government Transparency, and Citizen Participation in the ASEAN,* (2015), 131.
[71] John M. Ackerman, & Irma E. Sandoval-Ballesteros, "The Global Explosion of Freedom of Information Laws," *Administrative Law Review,* vol. 58, no. 1, 2006, pg. 93, http://www.jstor.org/stable/40712005.
[72] Ibid.
[73] "114 FOI Regimes," *Freedom Info,* April 5, 2016, http://www.freedominfo.org/2016/04/105-foi-regimes-freedominfo-org-count-shows/.
[74] Ben Bernanke, "China's Transparency Challenges," *Brookings,* March 8, 2016, https://www.brookings.edu/blog/ben-bernanke/2016/03/08/chinas-transparency-challenges/.

As the administrative size of states and the authority of the executive continue to grow, citizens feel increasingly distant from the decision-making processes, and credibility and trust in the government fall.[75] FOI laws combat this credibility gap because, in the age of information with ever-increasing amounts of raw data to process, citizens are only as informed and able to participate in democratic processes as their ability to access and leverage government information. Mark Bovens characterizes citizens' right to information as the difference between open government as a "question of public hygiene," or as "an element of citizenship."[76] In the first case, FOI laws are simply a means to increase government transparency; in the latter case, FOI laws are a fundamental part of individual rights as a citizen.

In tracing the path a country has taken to reach its current open data policies, FOI legislation can serve as an ideological parent to open data, and lessons can be drawn from the history of FOI in different regions of the world. For example, up until President Obama's executive action mandating government data be machine-readable and available at data.gov, U.S. FOI law previously covered data from agencies only within the executive branch.[77] This may hint at a more complicated federal stance on the right to information than simply stating that the U.S. has upheld the FOIA since 1966. Similarly, before open government data legislation in the UK, UK agencies actually used to be able to charge individuals for all four steps of the FOI process.[78]

---

[75] Cal Clark, & Don-Terry Veal, *Advancing Excellence and Public Trust in Government,* Lexington Books, 2011, pg. 23.
Bruce Ladd, *Crisis in Credibility*, The New American Library, Inc., 1968, pg. 217.
[76] Mark Bovens, "Information Rights: Citizenship in the Information Society," *The Journal of Political Philosophy,* 10(3), 327, Georgetown University Library, http://dx.doi.org/10.1111/1467-9760.00155.
[77] John M. Ackerman, & Irma E. Sandoval-Ballesteros, "The Global Explosion of Freedom of Information Laws," pg. 100.
[78] Ibid, pg. 109.

However, while FOI legislation overlaps with open data in its goal of government transparency and citizen awareness, it does not always correspond to similar support for federal open data programs. For example, Pakistan ranked #61 out of all 122 countries with open data initiatives in 2015,[79] and the country hosted its first national hackathon earlier this year. However, it has a questionable history with FOI law with current right to information (RTI) bills stalled in revisions; the current law has blanket immunity for certain government institutions.[80] In cases such as Pakistan, alternative explanations are needed to shed light on the factors behind the country's path towards an open data program.

Existing research on FOI legislation shows little impact on improving public understanding, participation, and trust in government. Based on individual interviews with UK government officials and a representative survey of FOI citizen users, Ben Worthy found that while Britain's 2000 FOI Act achieved its core objectives of increasing transparency and, in some cases, accountability, it did not improve public understanding, participation, or trust in government.[81] Worthy attributes this failure to the overreach of the stated goals of Britain's FOI Act rather than a failure in implementing the law. He argues that politicians overstated what FOI could do; consequently, when the FOI Act did not create 'transformative' democratic changes in trust and political participation, Britain's FOI Act seemed a letdown in comparison to its expected impact. It is also worth noting that Worthy found that cultures of secrecy,[82] were not a barrier to FOI implementation in the British government. A common argument as to why

[79] Argentina, Portugal, and South Africa share #54 rank. "Place Overview," *Open Data Index,* http://index.okfn.org/place/.
[80] Ahmed Mehboob, "The Right to Information," February 20, 2016, *Dawn*, http://www.dawn.com/news/1240632.
[81] Ben Worthy, "More Open but Not More Trusted? The Effect of the Freedom of Information Act 2000 on the United Kingdom Central Government," *Governance*, vol. 23, issue 4, September 2010, pg. 573-576, http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0491.2010.01498.x/full.
[82] Culture of secrecy is when a government and its officials guard knowledge, keeping the public uninformed, as it creates an asymmetric power dynamic between the government and its citizens.

transparency initiatives fail is that cultures of secrecy within a government inhibits openness and dialogue between the government and its citizens.

A similar study was conducted by the Open Democracy Advice Center in South Africa. They found that only 30% of public bodies and 11% of private bodies were aware of the existence of FOI legislation and active in implementing it.[83] Suzanne Piotrowski and colleagues and Nicola White also focus on problems that arise in the implementation of FOI law. Piotrowski and colleagues examined the Open Government Information (OGI) Regulations that went into effect in 2008 in China.[84] Nicola White focused on FOI legislation in New Zealand.[85] While the two countries vary considerably in form of government, research findings suggest that the drafting of the FOI law or OGI regulations is directly tied to the implementation and outcomes of the legislation are remarkably similar. Both researchers also state that the political will from all levels of government ties directly to the success of implementation.

While FOI and open data overlap in purpose, open data is distinctly different than FOI in the scope of its purpose and function. A notable negative aspect to FOI legislation, particularly in the United States, are the FOIA "hurdles" that citizens must often jump through to access the information they seek. "Submitting a FOIA is a convoluted and often-obscure process…citizens often get caught up in lengthy delay processes and rejections…when information is successfully obtained, it often lacks in comprehensibility."[86] On the other hand, open data portals offer a new and streamlined way of accessing information. Datasets can be easily downloaded from

[83] Ian Cume, & Jonathan Klaaren, "An Update on Access to Information in South Africa: New Directions in Transparency," 107 *Freedom of Information Review,* 2003.
[84] Suzanne J Piotrowski, et al, "Key Issues for Implementation of Chinese Open Government Information Regulations," *Public Administration Review,* vol. 69, December 2009, pg. 129-135, http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6210.2009.02100.x/full.
[85] Nicola White, *Free and Frank – Making the Official Information Act 1982 Work Better,* Institute of Policy Studies (Victoria University), November 2007.
[86] Donald Gordon, *Transparent Government: What It Means and How You Can Make It Happen,* Prometheus Books, 2014, pg. 93.

government data platforms and can be searched for in a number of ways, including by file type, government sector, or alphabetically. While FOI legislation sought to create transparency and increase citizen participation in government, but was hindered by its unwieldy design, the fundamental openness of open data portals has eliminated inefficiencies in past FOI implementation.

2.6.2 Involvement of a President/Prime Minister

The executive branch, whether it is at the municipal, state, or federal level, is often involved in the creation of open data initiatives. This raises the question of whether a president/prime minister's support plays a substantial role behind the open government data movement. In robust representative democracies, the president is beholden to public input and opinion in a different way than are legislators. "Unlike individual legislators, voters hold the president responsible to and for the nation as a whole."[87] A president is more likely than a legislature to support projects that have a low social return but are politically valuable; in other words, while these projects may look good for the nation on the surface, they yield little or no net return.[88] At the subnational level, examples of this are the public statements issued by mayors and governors when discussing the launch of an open data initiative. When announcing these projects, mayors and governors' rhetoric can be seeped with democracy-related and community-building language. Take Mayor Garcetti's announcement of the Los Angeles open data portal, for example. He states, "Angelenos deserve transparency and accountability, and an open view of government... I (Mayor Garcetti) called up on this council to propose ideas… to help

[87] Jeffrey E. Cohen, *Presidential Leadership in Public Opinion,* Cambridge University Press, 2015, 137.
[88] Martin Ravallion, "Economy-wide Policies," Lecture, Georgetown University Economics Department, Washington D.C, November 17, 2016.
Katrin Voltmer, "Chapter 6: The Media, Government Accountability, and Citizen Engagement," *Public Sentinel: News Media & Governance Reform*, The World Bank, November, 2009, pg. 2.

government work better, to save taxpayers money, and to prepare our students, workforce, and economy for the 21st century."[89] Normative claims of transparency and accountability, preparation of the next generation and workforce, and government efficiency are condensed into a single press statement on open data.

Proposed projects from the executive branch may not always be motivated by self-preservation and re-election chances. Brandice Canes-Wrone acknowledges that a president's "involvement of the mass public does shift policy toward majority opinion," she finds that "a president will not endorse a popular policy he believes is contrary to the interest of society. In other words…the popular policies that the president takes to the airwaves are the ones that he believes will improve societal welfare."[90]

Again looking at the subnational level, regardless of the political motivations behind the creation of open data initiatives, these initiatives can have positive effects in generating community-sourced solutions for local infrastructure and governance. Donald Gordon points to Portland, Oregon and its support of CivicApps, a city open data portal and expanding community of developers and enthusiasts, as an example of an executive-supported open data initiative.[91] In India, Hyderabad's IT and open data initiatives serve as another example of cities' motivation, particularly for the executive branch, to create innovation policies that attracts young professionals and capital.[92] Regardless of whether these policies have a high or low return, the executive branch, at both the state and federal level, has a strong motivation to foster innovation

[89] *City of Los Angeles Website*, "Mayor Garcetti Launches Open Data Portal For City of Los Angeles at #TechLA Conference," May 31, 2014, https://www.lamayor.org/mayor-garcetti-launches-open-data-portal-city-los-angeles-techla-conference.
[90] Brandice Canes-Wrone, *Who Leads Whom? Presidents, Policy, and the Public,* The University of Chicago Press, 2006, pg. 5, 48, 182.
[91] Donald Gordon, *Transparent Government: What It Means and How You Can Make It Happen,* Prometheus Books, 2014, pg. 188.
[92] Sudheer Goutham, "KTR's Mission Hyderabad as India's Silicon Valley," *eGov*, October 04, 2016, http://egov.eletsonline.com/2016/10/ktrs-mission-hyderabad-as-indias-silicon-valley/.

and open data programs, as they are politically valuable in garnering public support. These municipal open data projects in turn feed into a growing national open data portal.

In weak democracies or strong states, a president/prime minister may also play an antagonistic role in open data projects. This can be due to the possibility that opening government data may undermine the authority of the executive. Many countries in sub-Saharan Africa face this issue, and the major proponents for open data initiatives are instead international organizations or lobby groups within the country. There can also be resistance in moving away from secretive cultures national governments in sub-Saharan Africa.[93] If an open data policy is adopted, a number of factors, ranging from executive pushback to a systemic lack of political can hinder any significant open data development in this region.[94]

2.6.3 Security, Privacy, and Anonymity

Mass surveillance by robust democracies such as the United States or Germany has continued to expand following 9/11.[95] In 2015, 74% of Americans believed that they should not give up privacy and freedom for the sake of safety.[96] As record lows in public trust in government continue today, the largest deterrent to open data policies are concerns over security,

[93] Rebecca Zausmer, *Global Partners & Associates*, "Towards Open and Transparent Government: International Experiences and Best Practice," December, 2011, 22, http://www.gp-digital.org/wp-content/uploads/pubs/Towards%20open%20and%20transparent%20government-%20International%20experiences%20and%20best%20practice.pdf.
Joshua Masinde, *Quartz,* "More African Governments are Enacting Open Data Policies But Still Aren't Willing to Share Information," September 14, 2016, https://qz.com/780582/african-governments-still-wont-let-their-citizens-get-data-to-support-accountability-and-fight-corruption/.
[94] Gilbert Ronald Sendugwa, "Legislating and Implementing Public Access to Information in Africa: What are the Incentives for Government and Civil Society Actors?" *Africa Freedom of Information Centre (AFIC),* 2010, https://spaa.newark.rutgers.edu/sites/default/files/files/Transparency_Research_Conference/Papers/Sendugwa_Gilbert.pdf.
[95] Ironically, immediately after 9/11, America saw a substantial increase in trust in government.
[96] George Gao, "What Americans Think About NSA Surveillance, National Security and Privacy," *Pew Research Center,* May 29, 2015, http://www.pewresearch.org/fact-tank/2015/05/29/what-americans-think-about-nsa-surveillance-national-security-and-privacy/.

privacy, and anonymity. Security is defined as a risk materializing (e.g. hacking).[97] Privacy refers to the privacy of individuals, and anonymity alludes to the mosaic effect and cleaning data before release.[98]

At the heart of this issue is the tradeoff between risk and reward. On one hand, the digital economy relies on the ability to move data around digitally, raising security risks.[99] Similarly, in order to derive benefits from open data initiatives, citizens must be willing to accept the risks involved in uploading massive amounts of anonymized granular data for public use. One of the most frequently cited concerns is the mosaic effect, in which individuals can piece anonymized datasets together to obtain personal information on thousands of users. For example, in 2014, researchers were able to identify every driver part of The New York City Taxi and Limousine Commission by using cryptographic hashing on a public dataset to de-identify medallion numbers.[100] While larger, more valuable anonymized open datasets have greater potential to spark innovation and create tangible benefits in citizens' lives, these datasets also pose greater risks in exposing personal and financial information about millions of citizens. Since the Snowden leaks in 2013, in which an ex-CIA contractor leaked details of US government surveillance of citizens to the media, citizen trust in the government handling personal data has continued to flounder. The term the "Snowden effect," which refers to the increase in public concern about government data collection and internet privacy, is part of a larger shift in the way citizens view intelligence agencies.[101]

---

[97] Hrushikesha Mohanty, Prachet Bhuyan, & Deepak Cehnthati. *Big Data: A Primer*. Springer, 2015, pg. 124.
[98] Ibid.
[99] Rob Thomas, & Patrick McSharry, *Big Data Revolution: What Farmers, Doctors, and Insurance Agents Teach Us About Discovering Big Data Patterns,* John Wiley & Sons Ltd, 2015, pg. 91.

[100] "Briefing Paper on Open Data and Privacy," *The Center for Open Data Enterprise,* 2016, http://reports.opendataenterprise.org/BriefingPaperonOpenDataandPrivacy.pdf.
[101] Rob Thomas, & Patrick McSharry, *Big Data Revolution: What Farmers, Doctors, and Insurance Agents Teach Us About Discovering Big Data Patterns,* pg. 95.

2.6.4 The Economics of Open Data

The potential economic impact of open data is one of the most prominent narratives in pro-open data rhetoric. The 2013 McKinsey study on open data estimated that open data could generate approximately $3.2 trillion annually across seven sectors, such as education ($890 – $1,180 trillion) or consumer products ($520 - $1,470 trillion).[102] Tied to these figures are smart city initiatives. Smart Cities are urban development projects to integrate information technology and Internet of Things (IoT) solutions into a "smarter" city. These solutions are meant to improve the quality of life for residents in smart cities as well as increase efficiency in municipal functioning.

Apart from the McKinsey study, other researchers have been hesitant to generate a numerical value for the economic benefits of open data. Huijboom and Van den Broek, for example, cautioned that evidence of economic impacts of open data policy is still immature.[103] Even in 2017, the open data White House Transition Report refrained from forming financial estimates on the potential of open data, and instead focused on the tangible ways open data can create opportunity, such as creating a National Hunger Heat Map to help distribute food supplies to areas of greatest need.[104]

Some argue that open data's economic potential rests in its ability for governments to demonstrate increased financial transparency, which in turn will increase levels of investment in those governments.[105] Relly and Sabharwal argue that because of the globalization movement, China and India have both adopted regulations and legislation to "demonstrate and implement

[102] James Manyika, et al. "Open Data: Unlocking Innovation and Performance with Liquid Information," pg. 9.
[103] Noor Huijboom, & Tijs Van den Broek, "Open Data: An International Comparison of Strategies," pg. 10.
[104] Joel Gurin, & Laura Manley, "Open Data Transition Report: An Action Plan for the Next Administration," pg. 27.
[105] Cal Clark & Don-Terry Veal, "Advancing Excellence and Public Trust in Government," pg. 35.

various levels of transparency."[106] Fostering strong international investment has become dependent on displaying transparent behavior, something the open government data has the ability to create.

However, Ackerman and Sandoval-Ballesteros warn against the idea that economic development is tied to increased transparency. They argue that "it would be a mistake to imagine that freedom of information is a natural outgrowth of economic development,"[107] citing Singapore as an example of an economic hub with draconian laws and an opaque government. Ackerman and Sandoval-Ballesteros further add that countries like Jamaica are much poorer than Singapore but have FOI legislation[108] and open data initiatives.

Nevertheless, in the case of low and lower-middle income countries, adopting open data initiatives may be an essential part of generating income in the form of external aid. Joining OGP and other transparency-fostering organizations may further expand the options for countries to receive donor assistance that is tied to open government and open data, so it may be in their interest to create open data programs at least for the financial rewards.

2.6.5 Corruption

Literature on corruption is oftentimes tied to discussions of freedom of information legislation and cultures of transparency and accountability. Particularly when discussing open data initiatives and e-government initiatives, anti-corruption legislation goes hand-in-hand with transparency and accountability. Much of the literature on corruption examines the issue as (1)

[106] Jeannine E. Relly & Meghna Sabharwal, "Perceptions of Transparency of Government Policymaking: A Cross-national Study," *Government Information Quarterly*, vol. 26, is 1, January 2009, pg. 148-157, http://dx.doi.org/10.1016/j.giq.2008.04.002.
[107] John M. Ackerman, & Irma E. Sandoval-Ballesteros, "The Global Explosion of Freedom of Information Laws," pg. 115.
[108] Ibid.

corruption's effect on government and society (2) tactics to mitigate corruption (3) the benefits of decreased corruption due to these tactics. The discussion of corruption in this paper, however, is to examine corruption's effect on open data initiatives, as well as current open data initiatives who work to reduce corruption.

Corruption can occur in many forms and on multiple levels. It occurs on the citizen-to-citizen and citizen-to-government official levels, often in the form of bribes. It occurs on the institutional level in the form of diverting public funds for personal benefit, nepotism and clientelism, and institutional capture, in which an entire state or public institution is taken over "by an elite cartel of political and business oligarchs."[109]

Many open data initiatives sponsored by international organizations are based on anti-corruption initiatives. These initiatives generally operate at the institutional level, and include the International Budget Partnership's Open Budget Survey initiative,[110] the Open Contracting Partnership,[111] and the Global Initiative for Fiscal Transparency;[112] all three examples are NGOs that work in multiple countries to promote transparency and anti-corruption initiatives. For these organizations, anti-corruption initiatives and budget transparency overlap greatly with open data initiatives, both of which have been tied to normative arguments advocating for greater democratic freedom and freedom of information.[113] In fact, democratization's effect on corruption is a well-established field of literature, particularly in terms of post-communist

[109] Rasma Karklins, *The System Made Me Do It: Corruption in Post-Communist Societies,* M.E. Sharpe (2005), 30.
[110] "Open Budget Survey," *International Budget Partnership,* http://www.internationalbudget.org/opening-budgets/open-budget-initiative/open-budget-survey/.
[111] *Open Contracting Partnership,* http://www.open-contracting.org/.
[112] *Global Initiative for Fiscal Transparency*, http://www.fiscaltransparency.net/.
[113] "Fighting Corruption in Eastern Europe and Central Asia: The Istanbul Anti-Corruption Action Plan," OECD (2008), 71.
Craig Fagan, "Why Open Data Can Stop Corruption," *Transparency International Blog,* March 24, 2016, http://blog.transparency.org/2016/03/24/why-open-data-can-stop-corruption/.

countries in Eastern Europe.[114] Kostadinova, who conducted regression analysis on EU association and the Corruptions Perceptions Index, found that EU integration from 1997 to 2008 had a pronounced impact on reducing corruption.[115] It is possible that, in regards to open data initiatives, the influence the EU exerts on its member states may also exhibit a similar positive relationship with the strength of open data initiatives. Kostadinova goes on to discuss how the technical capacity of a nation reduces corruption.[116] Citing Philippe Schmitter,[117] she argues that education and communications systems improve government service, and can be utilized to combat corruption. She argues that more competent, educated bureaucrats "are likely to be appointed who would perform their duties in a professional manner free from political influence."[118] Additionally, she argues that as quality of service from the government increases due to e-government and other advanced forms of administration, the opportunity for bribery decreases.

2.6.6 Influence of International Organizations

Open data initiatives do not develop and exist in a vacuum, separate from one another. The vast majority of initiatives are constantly guided by nongovernment organizations and are funded by international organizations and foundations. However, while the involvement of NGOs/IGOs is widespread in the field of national open data initiatives, little research has been done on the influence of these organizations' involvement. In fact, it is difficult to discern a single unified body of academic literature solely focused on NGOs that fits into mainstream

[114] Rasma Karklins, *The System Made Me Do It: Corruption in Post-Communist Societies*, M.E. Sharpe (2005). Tatiana Kostadinova, *Political Corruption in Eastern Europe*, Lynne Rienner Publishers, Inc (2012).
[115] Ibid, 56.
[116] Ibid, 136.
[117] Ibid, Philippe Schmitter, "Democratization and the State Capacity," 2005.
[118] Ibid.

theories in international relations or other fields. [119] In some cases, NGOs are considered a part of the domestic nonprofit sector, while in others, they are considered public interest groups, or fit in the framework of social movements.[120] This is especially true of NGO work in open data initiatives; open data NGO proponents come from transparency, e-government, anti-corruption, smart city, sustainable development, education, and many other fields.

There is no single definition of an NGO; indeed, scholars argue that a definition of NGOs is an inherently normative statement.[121] This is due to the tendency of NGOs themselves to inject moral beliefs, "universal norms such as justice or freedom,"[122] in their self-described functions. In doing so, they set themselves apart from other interest groups,[123] as well as assume a role in determining public good. Baur illustrates this point by examining the UN Department of Public Information's definition of NGOs: "a not-for-profit, voluntary citizens' group, which is organized on a local, national or international level to address issues in support of the public good."[124] Baur argues that by writing this definition with the inclusion of a normative claim to the public good, the UN hints that it only deals with legitimate NGOs. This issue of legitimacy within an international arena dominated by nation-states is a topic that is returned to frequently in NGO literature.

For the purposes of this paper, an NGO refers to a "non-profit organization that operates independently of any government, typically one whose purpose is to address a social or political

[119] Samima Ahmed and David M. Potter, *NGOs in International Politics*, Kumarian Press (2006), 9.
[120] Ibid.
[121] Ibid, 8.
Dorothea Baur, *NGOs as Legitimate Partners of Corporations: A Political Conceptualization,* Springer (2011), 3.
[122] Ibid.
[123] Ibid.
[124] Ibid, 5.

issue."[125] An IGO refers to "an entity created by treaty, involving two or more nations, to work in good faith, on issues of common interest."[126]

NGOs and IGOs hold a unique place in the power dynamics between state and non-state actors. They have the potential to be influential due to their bargaining assets: "expertise, closeness to a target group, domestic political constituencies, access to the media, and alliance building."[127] Additionally, NGOs have the ability to bring issues to attention of states as well as shape how these issues are discussed.[128] They operate within multiple sectors of international politics, such as charity and disaster relief, social and economic development, political and advocacy, public education, etc.

In open data, NGOs and IGOs work at all levels of open data initiatives. They work with citizens, provide training and resources for open data advocacy and use. Open Knowledge International's global network is an example of this. ODI provides citizens with the knowledge and tools to create a local chapter of ODI, and its network connects individuals with other open data advocates around the world. The value of this network cannot be understated; open data is still a nascent field and best practices are still evolving; connecting open data advocates allows for them to share experience and tried-and-tested routes. Conversely, in creating such a large network, all of whose members are dependent on ODI's continued support and network, ODI is

[125] "NGO," *Oxford Dictionary*, https://en.oxforddictionaries.com/definition/NGO.
The choice of the Oxford Dictionary was deliberate; it defines the constraints of an NGO (1) non-profit (2) independent of any government, as well as gives a general context for the usage of the word. Additional criteria outlined by the UN definition of NGOs include (3) cannot advocate the use of violence (4) cannot be a school, university, or political party.
[126] "Intergovernmental Organizations (IGOs)," *Harvard Law School*, http://hls.harvard.edu/dept/opia/what-is-public-interest-law/public-international-law/intergovernmental-organizations-igos/.
[127] Charlotte Dany, *Global Governance and NGO Participation: Shaping the Information Society in the United Nations,"* Routledge (2013), 23.
[128] Ibid.

able to promote its agenda on a multinational level, with country-specific experts available at the NGO's disposal.

Open data NGOs also operate on the municipal, state, and national levels. At these levels, NGOs may work as consultants to countries in conducting the Open Data Readiness Assessment, as developers who build a data portal, and as initiators, who reach out to governments to partner or lead open data projects. Many of the open data projects sponsored by international organizations and foundations at the national level go through the World Bank. These initiatives range from creating open budgets portals in collaborating countries (BOOST initiative), to the Extractive Industries Transparency Initiative (EITI). Although the World Bank and many other IGOs/NGOs have made attempts to open budget data, tracking open data-specific program budgets is difficult, as programs vary in classification (e.g., transparency, e-government, and sustainable development). Therefore, while it is possible to track membership of countries to open data programs funded by IGOs/NGOs, it is difficult to ascertain actual aggregate annual funding for these programs. However, this information is necessary in evaluating the role NGOs/IGOs play in national open data initiatives, as the strength of a given initiative may depend on continued external funding from the NGO/IGO.

NGOs generally influence open data initiatives within countries in three ways (1) work with a government to develop an open data initiative, perhaps conducting the Open Data Readiness Assessment, or developing the portal for a country (2) work with citizens in a country to develop and advocate for an open data initiative (3) create a data portal for a country without the involvement of the country in question.

Just as the strength of an open data initiative financially sponsored by an NGO is dependent on that NGO, the strength of open data initiatives in which an NGO builds a portal for

a country, either with or without the involvement of that country, is dependent on the legitimacy of the NGO. Baur separates the challenges NGOs face in legitimacy into the structural, substantive, and procedural dimensions.[129] Structurally, NGOs face a legitimacy deficit due to the fact that they work with a population that has not elected them. Furthermore, they "cannot be held accountable for their actions on a transnational level."[130] Scholars identify this legitimacy deficit as the lack of formal democratic legitimacy. In the substantive dimension, Baur argues that while the purpose of NGOs are usually to serve the public good, who decides that public good, and what criteria is used to determine this good is unclear and not explicitly established.[131] On the procedural level, Baur raises the question, "What kinds of behavior are legitimate for NGOs in order to put forward their claims?"[132] In order to promote their agenda, NGOs may create campaigns containing leading or misinformation. These campaigns may cause a loss of legitimacy, and also bring into question the normative claims that NGO goals align with the public good.

The issue of legitimacy is especially important when working alongside a national government on an open data initiative. Successful open data initiatives are not only reliant on the active government officials involved in the program, but also on the various departments and/or ministries of a government, who will eventually be the ones to generate and upload their data in machine-readable, openly available formats. Therefore, if an NGO/IGO is spearheading an open data initiative, the legitimacy of that NGO/IGO in the eyes of these ministry officials is essential for the success of the portal.

[129] Dorothea Baur, *NGOs as Legitimate Partners of Corporations: A Political Conceptualization,* 7.
[130] Ibid. Cited in Habermas (2001); Zurn (2004); Habermas (2004), 174.
[131] Ibid, 8.
[132] Ibid.

NGOs may also bypass the national government, and work with citizens directly on open data initiatives. Keck and Sikkink discuss the "boomerang pattern," a form of lobbying that aligns with this kind of citizen-NGO relationship.[133] The "boomerang pattern" describes a path in which an NGO who finds their "claims of rights blocked by their own government," bypass their governments by locating and reaching out to other NGOs and IGOs internationally, to pressure the original government on behalf of the local NGO.[134] Ahmed and Potter provide an example of the "boomerang pattern," describing how Indonesian NGOs, in collaboration with residents against the construction of the Koto Panjang Hydroelectric Dam, leveraged media press over their lawsuit to garner the support of the Foundation for Human Rights in Asia, a Japanese NGO. The NGOs then sought a public statement from the Japanese government decrying the construction of the dam.

In an ever more-connected world, the boomerang pattern, as well as the more general pattern of advocates of any cause finding each other online and sharing best practices, is observable in open data. Examples of this include ODI's chapters around the world, as well as a Google Sheets document created by the World Bank in June, 2014, consolidating the professional information of open government data experts.[135] Additionally, the number of in-person meetings and conferences has risen over the past few years. The International Open Data Conference in 2016 hosted close to 1,700 attendees.[136]

[133] Samima Ahmed and David M. Potter, *NGOs in International Politics*, 45.
[134] Ibid.
[135] "World Bank Open Government Data Community of Expert Consultants," *World Bank*, August 12, 2016, https://docs.google.com/spreadsheets/d/1JN1BeXagQNLnyRs3YRcIdpDbmdIH9yeYtm1hXQZm400/edit#gid=0.
[136] "International Open Data Roadmap," *OD4D*, 2016, 5, http://od4d.com/roadmap/assets/files/report-iodc-2016-web.pdf.

Chapter 3: Research Methods for Data Analysis

3.1 Chapter Overview

The political, cultural, and societal factors influencing the strength of a national data portal are assessed using secondary data from a variety of intergovernmental and nongovernment organizations, including the World Wide Web Foundation, the World Bank, and others. The dependent variable for these tests is the strength of a national data portal, while independent variables are split into four categories: cultures of transparency and accountability, government involvement and support for technology, education attainment and technology access and adoption, and the involvement of international organizations promoting open data initiatives. Data for the strength of national data portals is taken from the 3rd edition of Open Data Barometer's (ODB) rankings and scores of national open data initiatives around the world.

3.2 Statistical Analysis

This paper employs Spearman's rank correlation coefficient and Mann-Whitney U tests, where the dependent variable for both tests is the strength of open data initiatives. Both Spearman's rho and the Mann-Whitney U test are used because in comparison to a standard linear regression or t-test, Spearman's rho and the Mann-Whitney U test do not require the assumption of normal distributions. Given the nonlinear nature of the data, both Spearman's rho and Mann-Whitney U are able to test the relationship between various independent variables and the strength of open data initiatives. The Mann-Whitney U test is used with OGP's data on what government organizations were involved in the development of the open data initiative, which are coded as either 0 (not involved), or 1 (involved). Spearman's rho is used for all other independent variables, as the rest are all continuous data.

Each independent variable is tested against ODB's strength of open data initiatives scores. The first test examines the relationship between the independent variable in question and the overall open data country score as well as each of the ODB scores for the 15 kinds of datasets (e.g. Map, Land, Census, Budget). The same test is then conducted on a regional level, and in some cases, income level is used as an additional filter following region grouping. Statistically significant variables within the regional grouping are recorded in the table.

3.3 Dependent Variable: Open Data Barometer

The primary dependent variable in this thesis is the strength of national data portals; data are taken from ODB's 2015 dataset scoring open data initiatives globally. Open Data Barometer is produced by the World Wide Web Foundation as a collaborative work of the Open Data for Development (OD4D) network, with the support of the Omidyar Network.[137] It aims to assess "the true prevalence and impact of open data initiatives around the world."[138] ODB has released two previous reports on the state of open data globally in 2013 and 2014.

ODB's current 2015 edition encompasses 92 countries around the world, and in addition to assessing the strength of a country's data portal, the report also ranks nations on readiness and implementation of open data initiatives, as well as the impact of open data on business, politics, and civil society. The third edition also has a stronger regional analysis in ODB's assessment of their data.

The data itself is based on peer-reviewed expert surveys, government self-assessment surveys, and secondary data from the World Economic Forum, World Bank, United Nations e-

[137] Open Data Barometer, "The Open Data Barometer," http://opendatabarometer.org/barometer/.
[138] Ibid.

Government Survey, and Freedom House. For the expert surveys, trained country specialists are asked questions about open data contexts, policy, implementation, and impacts of each country. Experts are given detailed scoring guidance and respond to each question on a 0 to 10 scale. These responses are then normalized using z-scores for each question. As this thesis examines factors driving open data portal initiatives rather than the resulting implementation and impact of open data portals, I do not include this expert survey data in my dataset.

Using the same team of researchers, ODB also conducts dataset assessments split into 15 kinds of data within each country. The data are scored by the following criteria:[139]

Question	Weight	Chaining logic	Qualitative data collected
A Does the data exist?	5	IF a = No THEN 0 ELSE 5	Description of data; Agency responsible; Reasons for non-collection
B Is it available online from government in any form?	10	IF a = No THEN 0 ELSE (IF b = Yes THEN 10 ELSE 0)	URL; Limits on data published; Policies preventing publication
C Is the dataset provided in machine-readable formats?	15	IF b = No THEN 0 ELSE (IF c = Yes THEN 15 ELSE 0)	URL; File formats;
D Is the machine readable data available in bulk?	15	IF c = No THEN 0 ELSE (IF d = Yes THEN 15 ELSE 0)	URL
E Is the dataset available free of charge?	15	IF c = No THEN 0 ELSE (IF e = Yes THEN 15 ELSE 0)	Details of charging regimes
F Is the data openly licensed?	15	IF c = No THEN 0 ELSE (IF e = Yes THEN 15 ELSE 0)	URL; License details
G Is the dataset up to date? Logic: lose 5 points if machine readable data is the data is outdated. Gain 10 points if it is timely.	10	IF c= No THEN 0 ELSE (IF g = No THEN -5) ELSE (IF c = Yes AND g = Yes THEN 10)	Last update date; Frequency of updates
H Is the publication of the dataset sustainable?	5	IF c = No THEN 0 ELSE (IF h = Yes THEN 5 ELSE 0)	Evidence of sustainability
I Was it easy to find information about this dataset?	5	IF c = No THEN 0 ELSE (IF i = Yes THEN 5 ELSE 0)	Notes on discoverability
J Are (linked) data URIs provided for key elements of the data?	5	IF c = No THEN 0 ELSE (IF j = Yes THEN 5 ELSE 0)	URL of linked data publication

[139] Open Data Barometer, "ODB Methodology – v1.0 28th April 2015," April, 2015, http://opendatabarometer.org/doc/3rdEdition/ODB-3rdEdition-Methodology.pdf.

Responses are then peer-reviewed and weighed to emphasize questions C, D, E, and F, which relate to key aspects of Open Knowledge International's definition of "openness."[140] The overall score of the kind of dataset is then calculated on a 0-100 score. The 15 kinds of datasets are:[141]

Variable Name	Short Name	Long Name	Description
D1	Map	Mapping data	A detailed digital map of the country provided by a national mapping agency and kept updated with key features such as official administrative borders, roads and other important infrastructure. Please look for maps of at least a scale of 1:250,000 or better (1cm = 2.5km).
D2	Land	Land ownership data	A dataset that provides national level information on land ownership. This will usually be held by a land registration agency, and usually relies on the existence of a national land registration database.
D4	Stats	National statistics	Key national statistics such as demographic and economic indicators (GDP, unemployment, population, etc), often provided by a National Statistics Agency. Aggregate data (e.g. GDP for whole country at a quarterly level, or population at an annual level) is considered acceptable for this category
D5	Budget	Detailed budget data	National government budget at a high level (e.g. spending by sector, department etc). Budgets are government plans for expenditure, (not details of actual expenditure in the past which is covered in the spend category).
D6	Spend	Government spend data	Records of actual (past) national government spending at a detailed transactional level; at the level of month to month government expenditure on specific items (usually this means individual records of spending amounts under $1m or even under $100k). Note: A database of contracts awarded or similar is not sufficient for this category, which refers to detailed ongoing data on actual expenditure
D7	Company	Company registration data	A list of registered (limited liability) companies in the country including name, unique identifier and additional information such as address, registered activities. The data in this category does not need to include detailed financial data such as balance sheet etc.
D8	Legislation	Legislation data	The constitution and laws of a country.
D9	Transport	Public transport timetable data	Details of when and where public transport services such as buses and rail services are expected to run. Please provide details for both bus and rail services if applicable. If no national data is available, please check and provide details related to the capital city.
D10	Trade	International trade data	Details of the import and export of specific commodities and/or balance of trade data against other countries.
D11	Health	Health sector performance data	Statistics generated from administrative data that could be used to indicate performance of specific services, or the healthcare system as a whole. The performance of health services in a country has a significant impact on the welfare of citizens. Look for ongoing statistics generated from administrative data that could be used to indicate performance of specific services, or the healthcare system as a whole. Health performance data might include: Levels of vaccination; Levels of access to health care; Health care outcomes for particular groups; Patient satisfaction with health services.

[140] Open Knowledge International, "The Open Definition," http://opendefinition.org/.
[141] Open Data Barometer, "ODB Methodology – v1.0 28th April 2015."

D12	Education	Primary and secondary education performance data	The performance of education services in a country has a significant impact on the welfare of citizens. Look for ongoing statistics generated from administrative data that could be used to indicate performance of specific services, or the education system as a whole. Performance data might include: Test scores for pupils in national examinations; School attendance rates; Teacher attendance rates. Simple lists of schools do not qualify as education performance data.
D13	Crime	Crime statistics data	Annual returns on levels of crime and/or detailed crime reports. Crime statistics can be provided at a variety of levels of granularity, from annual returns on levels of crime, to detailed realtime crime by crime reports published online and geolocated, allowing the creation of crime maps.
D14	Environment	National environmental statistics data	Data on one or more of: carbon emissions, emission of pollutants (e.g. carbon monoxides, nitrogen oxides, particulate matter etc.), and deforestation. Please provide links to sources for each if available.
D15	Elections	National election results data	Results by constituency / district for the most all national electoral contests over the last ten years.
D16	Contracting	Public contracting data	Details of the contracts issued by the national government.

I use ODB's calculated scores from these expert surveys as well as the overall calculated score of a country's open data initiative as the dependent variables in this paper. I do, however, exclude variables D2 – Land, D6 – Spend, D7 – Company, and D8 – Legislation. ODB identifies these types of datasets as "unlikely to be found;"[142] D2 – Land Ownership, is only available as fully open data (meaning it meets all requirements of The Open Definition)[143] in Australia, Canada, Estonia, the UK, and Uruguay. D6 – Government Spending, is only available in Brazil and the UK. D7- Companies Registers, is only available in Australia, with significant limitations. D8 – Legislation, is only available in Brazil, France, South Korea, and the UK.[144]

The final component to ODB's data is the secondary data from the World Bank and other IGOs, which are used alongside ODB's expert survey data to calculate countries' readiness scores. ODB evaluates readiness through three components: (1) Government; (2) Citizens and

[142] Open Data Barometer, "Global Report," 18.
[143] OKFN, "The Open Definition," http://opendefinition.org/.
[144] Open Data Barometer, "Global Report," 18.

Civil Society; and (3) Entrepreneurs and Business.[145] This distinction is further expanded on in ODB's detailed methodology, in which it cites Tim Berners-Lee, the inventor of the World Wide Web, stating open data 'has to start at the top, it has to start in the middle and it has to start at the bottom.'[146]

I do not use ODB's readiness scores, as their scores are calculated using data from the UN e-Government survey and Freedom House's political and civil liberties datasets, which I include as independent variables in my own dataset. These readiness scores reflect a similar process as my own thesis: using secondary data to assess what a country's resulting open data portal initiative will look like. However, my research questions are testing the association between variables like civil liberties and the strength of an open data portal; ODB includes them as assumed influential factors in the readiness of a country for an open data initiative, based on qualitative research. Nevertheless, the framework ODB provides to evaluate readiness: citizens and civil society, entrepreneurs and business, and government, will be useful in evaluating how different factors associated with the strength of open data portals fit in terms of what role they played in shaping an open data portal initiative as well as what policy implications can be drawn from their roles going forward.

3.3.1 Other Open Data Portal Ranking Initiatives

In addition to Open Data Barometer, Open Data Index (ODI), whose parent organization is Open Knowledge International, and Open Data Watch (ODW) both evaluate the strength of open data initiatives globally. While ODW's data is navigable and has a strong regional comparison visualizing tool, its raw data is not as accessible as GODI or ODB's. Between GODI

[145] Open Data Barometer, "ODB Methodology – v1.0 28th April 2015.
[146] Ibid.

and ODB, I had initially planned on using GODI's dataset, which closely mirrors ODB's 15 kinds of datasets dataset, with calculated country total score as well. GODI's dataset is more comprehensive in its breadth of open data initiatives, as it includes Taiwan as well as dependency countries such as Hong Kong and the Cayman Islands. Comparatively, GODI covers 122 countries, 25 new countries from its 2014 report; ODB covers 92 countries, 6 new countries from its 2014 report. This breadth is important, especially as GODI ranks Taiwan, which is not included in ODB's dataset, as the number one country in its 2015 index, moving up from #11 in 2014.

At the same time, there has been hesitation to take GODI's country rankings at face value. Hatem Ben Yacoub, an e-government specialist from Jeddah, Saudi Arabia, points to the over performance of several Arab countries' in GODI's 2015 report, and the dynamic change in some countries' rankings from 2014 to 2015 (eg. Oman's score raised +27 from 2014 to 2015 while Saudi Arabia fell -29).[147] This may be due to GODI's decision to not "bring forward" data submitted from the previous year with only basic review. Instead, the 2015 report was created solely with datasets submitted in 2015.[148] Unless drastic open data policy changes are in place between 2014 and 2015 across the board, it's possible that GODI's drastic rank changes are more reflective of methodology changes rather than open data initiative changes globally. I also found that many of the countries that appear in GODI's index but do not appear in ODB's index do not have a national open data portal. Rather, these countries may have an office of statistics or ministry of weather that releases its data online, but the country itself lacks a single national open data portal. For example, the Islamic Republic of Iran, which is not listed in ODB's dataset, is

[147] Hatem Ben Yacoub, "Why OKF Global Open Data Index 2015 is a Failure," December 10, 2015, http://www.hbyconsultancy.com/blog/why-okf-global-open-data-index-2015-is-a-failure.html.
[148] Global Open Data Index, "Methodology," 2015, http://index.okfn.org/methodology/.

ranked as #117 in GODI's index. Its listing is due to having weather forecasts and the country's laws available online, from its Meteorological Organization and its Laws and Regulations Portal, respectively.[149]

These indexing differences demonstrate that while IGOs, governments, and corporations may be touting their open data research and initiatives, open data is still very much in its early stages of development, and the development of a comprehensive index with universally accepted definitions of what constitutes a national open data initiative is still in the air.

3.4 Independent Variables

The independent variables used in this thesis are divided into four groups: cultures of transparency and accountability, government involvement and support for technology, education attainment and technology access and adoption, and the involvement of international organizations promoting open data initiatives.

In assembling the independent variables for this paper, I avoided including multiple secondary data sources for a single test, as well as avoid including indices whose sub components overlap. This was a prominent issue in choosing which dataset to address FOI legislation. The Centre for Law and Democracy's Right to Information (RTI) dataset is based on the actual text of FOI legislation in a country, rather than its implementation. I chose to only use the Centre for Law and Democracy's RTI dataset rather than include an additional organization's dataset on the implementation of FOI legislation globally, as this paper is not meant to evaluate organizations' choices in methodology, data collection, and analysis. Similarly, instead of

[149] Global Open Data Index, "Iran," 2015, http://index.okfn.org/place/iran/.

including the HDI as its own independent variable, I instead use components of the HDI, such as the mean years of schooling variable, as independent variables.

As the countries of which the data is collected on varies from dataset to dataset, the majority of tests will include a list of countries who are included in ODB's index (dependent variable), but who are missing from the index of the independent variable.

Within cultures of transparency and accountability, I use data provided by the Centre for Law and Democracy for FOI legislation scores globally,[150] Freedom House's 2015 political rights and civil liberties dataset,[151] and the total scores by country from Transparency International's 2015 Corruptions Perceptions Index.[152] All three of these variables are continuous data, and I use Spearman's rho to test for the strength and direction of their association to the strength of open data portals.

Within government involvement and support for technology, I use data provided by the Open Government Partnership (OGP), one of the most prominent international open data organizations, to determine whether the president/prime minister of a country was involved in the development of his/her country's open data initiative.[153] This data is categorical, coding the involvement of the president/prime minister in an open data initiative as a 0 (not involved) or 1 (involved). I also use data provided by OGP on whether the open data initiative was officially/legally mandated,[154] and UN Development Programme's Online Service Index variable within its 2014 e-government survey. This variable assesses the web content accessibility of a

[150] "Global Right to Information Rating," Centre for Law and Democracy, http://www.rti-rating.org/country-data/.
[151] Freedom House, "FIW 2016 Data," https://freedomhouse.org/report/freedom-world-2016/download-fiw-2016-data.
[152] Transparency International, "Corruptions Perceptions Index 2015," http://www.transparency.org/news/feature/corruption_perceptions_index_2016.
[153] "Public Access_IRM SU Commitments Database March 2017," *Open Government Partnership,* https://docs.google.com/spreadsheets/d/1MEF5zN41_4tW1pK5Ptp0vqK0HvA0ofwMN9XaXK-QVNU/edit?usp=sharing.
[154] "Public Access IRM SU Commitments Database March 2017," *Open Government Partnership.*

country's national website as well as several of that country's ministries' websites, which I use as a proxy for the familiarity of governments with web technology.

Within education attainment and technology access and adoption, I use data from the UN's e-government 2014 survey for e-government development scores,[155] its aggregate EGDI score, telecommunication score, and human capital index score.

Within the involvement of international organizations promoting open data initiatives, I use data I collected for a total tally of government involvement and membership in international organizations that promote open data within a country. The primary criteria for an organization or program to be included in this variable is whether or not the organization or program specifically creates and/or promotes an open data initiative within a given country. For example, while organizations and programs that work on anti-corruption or democracy-building initiatives, may include promoting openness in their stated goals, without a specific program developing government open data initiatives, these organizations and programs do not meet the set criteria.

I do not include variables regarding security, privacy, and anonymity, nor do I include variables relating to the economics of open data. While both these topics were covered in the literature review, data that directly reflects these topics are difficult to obtain. The World Values Survey features a question in its 6th wave (2010-2014) on whether respondents believe the government is wire-tapping or reading their mail or email. Apart from the WVS, OECD has a report on trust in government released March 2017, but while the report itself is available for viewing, the dataset is not.

Similarly, obtaining data on how much a country spends on open data initiatives, which can be categorized under transparency, e-governance, technology innovation, or a number of

[155] United Nations, "UN E-Government Survey 2014," https://publicadministration.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2014.

other government initiatives and programs, is difficult. While NGOs and IGOs may be transparent on where funding goes, it again is not clear under which field open data initiatives are categorized, and the level of detail of these expenditure datasets usually does not extend such that it is clear that money used for transparency initiatives went to open data rather than another transparency-fostering program.

3.5 Data Groupings

After initial test for association between a given independent variable and the strength of open data initiatives, I use region, income level, and the presence of municipal/state data portals as additional filters in the subsequent tests.

The region grouping is from GODI's dataset. Regions are: Americas, Asia, Caribbean, Europe, Middle East and North Africa, Oceania, and Sub-Saharan Africa. Comparatively, ODB's dataset, whose open data initiative scores I use as dependent variables, groups regions as: East Asia & Pacific, Europe and Central Asia, Latin America and Caribbean, Middle East and North Africa, North America, and Sub-Saharan Africa. Neither organization's groupings were ideal; GODI combines East, South, and Central Asia to Asia. ODB isolates North America but combines Latin America with the Caribbean. For the purposes of testing association, fewer groups of two or three countries (e.g. North America, Oceania) is better. I also wanted to test Europe by a grouping that more strongly reflected the European Union, and did not want to group Latin American open data initiatives with Caribbean initiatives.

In tests using Spearman's rho, Oceania and the Caribbean are excluded from the findings. This is due to their small sample size: Oceania has two countries in ODB's index and the

Caribbean has three. In tests grouping countries by region, groupings with less than four countries are similarly excluded from findings and analysis.

Income level is also taken from GODI's dataset, but both GODI and ODB use the World Bank's categorizations for their income level variables. Income level groupings are: low, lower-middle, upper-middle, and high. Low-income economies are defined as those with a GNI per capita of $1,025 or less in 2015; lower-middle income economies are between $1,026 and $4,035; upper-middle income economies are between $4,036 and $12,475; and high-income economies are $12,476 or more.[156] Using GNI as a grouping variable has its weaknesses; GNI does not reflect inequalities in income distribution, does not account for differences in domestic price levels, and may underestimate lower-income economies with more informal, subsistence activities.[157] A more holistic alternative I considered was the Human Development Index (HDI) from the UN Development Programme. However, this index covers a number of factors that overlap with other independent variables in my paper. Instead, I chose to test components of the HDI individually, such as mean number of years of education, or percent of internet users in a country.

[156] World Bank Data Team, "New Country Classifications by Income Level," July 1, 2016, https://blogs.worldbank.org/opendata/new-country-classifications-2016.
[157] "Why Use GNI per Capita to Classify Economies into Income Groupings?" World Bank, https://datahelpdesk.worldbank.org/knowledgebase/articles/378831-why-use-gni-per-capita-to-classify-economies-into.

Chapter 4: Results

4.1 Overview

The purposes of this analysis were: (1) to examine the relationship between cultures of transparency and accountability, government involvement and support for technology, education attainment and technology access and adoption, involvement of international organizations, and the strength of a country's open data initiative; and (2) to explore how these relationships vary depending on region and income level.

Researchers and open data organizations are still in the process of determining best practices in evaluating open data globally. The 3rd edition of the ODB made great strides in its regional analyses, particularly in its analyses of open data readiness. The following analyses are meant to complement ODB's open data readiness assessments, as well as provide a quantitative assessment on which factors influence a country's path towards open data initiatives, the strength and direction of their relationships to the strength of a country's open data initiative, and where around the world they have an influence.

The results are presented in two broad groupings: (1) overall descriptive statistics; and (2) by research aim. Analysis and discussion of these findings are included in Chapter 5.

4.2 Overall Descriptive Results

The strength of data initiatives acts as the dependent variable in this paper. The strength of data initiatives variable uses data from ODB's quantitative datasets.[158]

[158] *Open Data Barometer*, "Get the Data: Quantitative Datasets," http://opendatabarometer.org/3rdEdition/data/.

The Open Data Barometer includes 92 countries in its dataset. Grouped by region, Figure 4 shows all countries in ODB's dataset grouped by region, with their placement on the y-axis determined by their country's open data initiative's overall score.

Figure 4: ODB Scores Scaled Box and Whisker Plot

Although not included in this paper's findings, Oceania and the Caribbean are both worth noting. Both New Zealand and Australia support robust open data initiatives.[159] The New Zealand government has created open data fellowships for native developers to work with its open data, receiving stipends and mentoring by the Code For All network.[160] Australia, also

[159] As well as some of the most interesting open datasets on the planet. Australian datasets range from most popular baby names to records of the first criminals sent to Australia.
[160] Sara Barker, "Govt Calls Forth Kiwi Developers & Data Scientists for Open Data Fellowships," *FutureFiveNZ,* August 5, 2016, https://futurefive.co.nz/story/govt-calls-forth-kiwi-developers-data-scientists-open-data-fellowships/.

committed to its open data initiative, joined the Open Data Charter, a multilateral initiative in which members pledge for increased government open data initiatives, in 2016.

In ODB's index, the Caribbean countries include Jamaica, Saint Lucia, and Haiti, all of which began their data initiatives between 2014 to 2016. Saint Lucia was the first country in the Caribbean to conduct the Open Data Budget Readiness Assessment, which was conducted in partnership with the UK Department of International Development and the World Bank's program of assistance for the Caribbean region.[161] Jamaica conducted its own readiness assessment with the Center for Open Data Enterprise in 2016. Haiti, who holds the lowest rank (#92) in ODB's index, does not seem to have a national data initiative, but rather has a Department of Statistics that publishes some data online.[162] Similar to the case of Haiti, GODI further lists Bermuda, Dominica, Grenada, Saint Kitts and Nevis, St. Vincent & the Grenadines, Trinidad and Tobago, and Turks and Caicos Islands as having open data initiatives. These countries do not currently have a national portal, but rather have their respective departments of statistics publish certain datasets online.

Figure 5 presents descriptive statistics for the ODB's overall score for each country's open data initiative.

	ODB Country Overall Score
Mean	32.956
Standard Error	2.525
Median	27.305
Mode	15.99
Standard Deviation	24.222
Sample Variance	586.738
Kurtosis	-0.586
Skewness	0.644

Count = 92 Countries

Figure 5: ODB Overall Country Scores Descriptive Statistics

[161] "Open Data Readiness Assessment," *St Lucia Open Data Portal*, http://data.govt.lc/story/open-data-readiness-assessment.

[162] "Haiti," *Open Data Watch*, http://odin.opendatawatch.com/Report/countryProfile/HTI?appConfigId=1.

Figures 6 and 7 are descriptive statistics on each of the dataset types that are scored by ODB for each country. Values for kurtosis and skewness imply that the data, both the overall scores as well as each of the dataset kind scores, do not reflect normal distributions. The overall country score has a positive skewness, suggesting that the right-hand tail is longer than the left. This is the case for all of the dataset types except for D4 – Census, and D10 – Trade. Both these dataset types also have medians of 65, higher than the majority of dataset kinds whose medians are 15. Standard deviation is higher among these dataset kinds than the overall country score (SD = 24.222). Standard error is also lower for the overall country score (SE = 2.525) than compared to the 11 types of datasets.

	D1 - Map Score	D4 – Census Score	D5 – Budget Score	D9 – Transport Score	D10 – Trade Score
Mean	36.847	56.358	44.565	29.130	53.913
Standard Error	3.383	3.258	3.590	3.583	3.177
Median	15	65	15	15	65
Mode	15	15	15	15	15
Standard Deviation	32.452	31.256	34.439	34.367	30.479
Sample Variance	1053.141	976.9798	1186.073	1181.104	929.025
Kurtosis	-1.163	-1.478	-1.677	-0.730	-1.457
Skewness	0.532	-0.343	0.356	0.995	-0.304

Count = 92 Countries

Figure 6: ODB Dataset Kind Scores Descriptive Statistics

	D11 – Health Score	D12 – Education Score	D13- Crime Score	D14 – Environment Score	D15 – Elections Score	D16 – Contracts Score
Mean	44.728	43.097	39.456	41.032	43.695	27.554
Standard Error	3.223	3.268	3.423	3.587	3.564	2.975
Median	52.5	47.5	15	40	15	15
Mode	15	15	15	15	15	15
Standard Deviation	30.914	31.350	32.837	34.406	34.184	28.540
Sample Variance	955.694	982.880	1078.273	1183.812	1168.61	814.557
Kurtosis	-1.522	-1.504	-1.426	-1.596	-1.733	0.189
Skewness	0.040	0.203	0.402	0.167	0.217	1.367

Count = 92 Countries

Figure 7: ODB Dataset Kind Scores Descriptive Statistics cont.

4.3 Results by Research Aim

This section presents results of the tests of association and group differences for countries with open data initiatives. Tests are grouped into four topics: cultures of transparency and accountability, government involvement and support for technology, education attainment and technology access and adoption, and the involvement of international organizations. Each section will first present an overview of the dataset and its source, second, provide descriptive statistics on the dataset of the independent variable, and third, examine findings from the tests.

4.3.1 Cultures of Transparency and Accountability

This section addresses Research Aim 1, exploring the association between the strength of democracy and the strength of open data initiatives. The dependent variable, the strength of open data initiatives, uses expert-survey data provided by Open Data Barometer. The independent variables, FOI legislation, democratic freedoms, and corruption, use data provided by the Centre for Law and Democracy, Freedom House, and Transparency International.

Research Aim 1: Explore the association between the strength of democracy in a country and the strength of its national open data initiative.

- Specifically examining fulfillment levels of FOI requests, political rights and civil liberties, and level of corruption.

4.3.1.1 Freedom of Information Legislation (Right to Information)

4.3.1.1.1 Dataset Overview

This test contains two independent variables: the year in which right to information (RTI) legislation was passed in a country, and the 2015 score of its legislation. The data is provided by the Centre for Law and Democracy and Access Info Europe, who created a numerical assessment for the overall legal framework for the right to information in a country. The Centre for Law and Democracy (CLD) is a Canadian-based NGO whose work deals with freedom of expression, right to information, and digital rights.[163]

The Global RTI Ratings index is scored from 0 to 150 points; this total score is developed from seven indicators: right of access (6), scope (30), requesting procedures (30), exceptions and refusals (30), appeals (30), sanctions and protections (8), and promotional measures (16).

Their Global RTI Ratings dataset measures the legal framework of a country's RTI legislation, but not the actual implementation of the legislation. This means that Russia, who ranks 180 out of 199 countries for press freedom by Freedom House, and where at least 34 journalists have been murdered since 2000,[164] ranks 35th in the Global RTI Ratings index, just below the UK (34th).[165] However the strength of language in RTI legislation, as well as the existence of RTI legislation are both factors to consider in examining the relationship between RTI legislation and the strength of open data initiatives.

While ODB includes the following countries in its 2015 index as having open data initiatives, they do not have RTI legislation and are therefore excluded from tests of association: Bahrain, Benin, Botswana, Cameroon, Costa Rica, Egypt, Ghana, Malawi, Malaysia, Mali,

[163] "What We Do," *Centre for Law and Democracy*, http://www.law-democracy.org/live/about-us/what-we-do/.
[164] Linda Qiu, "Does Vladimir Putin Kill Journalists?" *PunditFact*, January 4, 2016, http://www.politifact.com/punditfact/article/2016/jan/04/does-vladimir-putin-kill-journalists/.
[165] *Global Right to Information Rating*, "Country Data," http://www.rti-rating.org/country-data/.

Mauritius, Morocco, Myanmar, Namibia, Qatar, Saint Lucia, Saudi Arabia, Senegal, Singapore, Tanzania, UAE, Venezuela, and Zambia.

4.3.1.1.2 Descriptive Statistics

Figure 8: RTI Date, Grouped by Region

Figure 8 displays countries who are included in both ODB's data initiative dataset and the RTI ratings dataset. The figure is grouped by region, with the date when a country's first RTI legislation was passed on the y axis. Oceania includes New Zealand and Australia; both have RTI legislation first passed in 1982, and subsequently, since both countries occupy the same space in the figure, only one country, New Zealand, is displayed. Additionally, Sweden, the first country ever to introduce RTI legislation in 1766, is included under a break in the y axis.

Figure 9: RTI Score, Grouped by Region

Figure 9 displays the score received by RTI legislation in each country on the y axis, and regional groupings along the x axis. Comparing Figures 10 and 11, the date of initial RTI legislation does not necessarily correspond to higher scores.

Table 3 lists the countries that have the oldest RTI legislation, as well as those that received the highest scores from the RTI ratings index but do not currently have open data initiatives.

	RTI Date	RTI Score
Mean	1998.701	86.721
Standard Error	2.693506	2.2804
Median	2003	87
Mode	2003	64
Standard Deviation	26.527	22.459
Sample Variance	703.732	504.432
Kurtosis	62.768	-0.600
Skewness	-7.265	0.143
Range	250	104
Minimum	1766	32
Maximum	2016	136

Count = 97 Countries

Table 4: RTI Descriptive Statistics

Countries with oldest RTI legislation without open data initiative (year)	Countries with highest RTI scores without open data initiative (score)
Lithuania (1996)	Serbia (135)
Uzbekistan (1997)	Slovenia (129)
Latvia (1998)	Albania (127)
Albania (1999)	El Salvador (122)
Trinidad and Tobago (1999)	South Sudan (120)

Table 3: RTI and ODB Country Overlap

Of these countries, Albania, Lithuania, Uzbekistan, and El Salvador are taking steps to assess open data readiness or create open data initiatives. Albania is currently a part of the World Bank's BOOST initiative for budget transparency. Lithuania is an interesting case. It signed a commitment for open government and sent a letter of intent to join OGP in 2011.[166] However, it never ended up creating a national portal. Nevertheless, Vilnius, the capital of Lithuania, has been working on an open data portal since 2015, and the city hosted its first open data hackathon in 2017.[167] Table 4 displays descriptive statistics on both the variables RTI date and RTI score.

4.3.1.1.3 Findings

I first test to see if there is a significant difference between the strength of open data initiatives in countries without RTI legislation compared to countries with RTI legislation. I test

[166] "Lithuania," *Open Government Partnership,* https://www.opengovpartnership.org/country/lithuania.

[167] Ramanauskaite, Egle Marija, "Technarium Hackerspace Celebrates Open Data Day in Vilnius, Lithuania," *Open Knowledge Foundation Blog*, April 27, 2017, https://blog.okfn.org/2017/04/27/technarium-hackerspace-celebrates-open-data-day-in-vilnius-lithuania/.

if these two groups are significantly different from one another using a Mann-Whitney U test. I assign 0 to countries without RTI legislation, and 1 to countries with RTI legislation.

	Mann-Whitney U	Z	Asymp. Sig. (2-tailed)
Country Overall Score	282.000**	-4.748	0.000
D1 - Map Score	452.000**	-3.284	0.001
D4 - Census Score	468.500**	-3.163	0.002
D5 - Budget Score	364.000**	-4.330	0.000
D9 - Transport Score	379.000**	-4.025	0.000
D10 - Trade Score	562.000*	-2.301	0.021
D11 - Health Score	551.500*	-2.391	0.017
D12 - Education Score	490.500**	-2.953	0.003
D13 - Crime Score	319.000**	-4.508	0.000
D14 - Environment Score	399.000**	-3.736	0.000
D15 - Elections Score	324.000**	-4.544	0.000
D16 - Contracts Score	562.000*	-2.538	0.011

0 (no RTI legislation) = 24, 1 (has RTI legislation) = 68, Total = 92
** = $p \leq 0.01$, * = $p \leq 0.05$
Table 5: Existence of RTI Legislation and the Strength of Open Data Initiatives

All countries in ODB's dataset were assessed, 24 countries did not have RTI legislation while 68 did. Table 5 displays the results of the nonparametric test. Both the country overall open data initiative score, as well as every dataset kind score showed statistically a different group difference between countries with and without RTI legislation.

	RTI Date	RTI Total Score
Country Overall Score	-.603** (0.000)	-.108 (0.380)
D1 - Map Score	-.577** (0.000)	-.160 (0.193)
D4 - Census Score	-.498** (0.000)	-.049 (0.690)
D5 - Budget Score	-.209 (0.087)	-.083 (0.499)
D9 - Transport Score	-.400** (0.001)	-.045 (0.716)
D10 - Trade Score	-.554** (0.000)	-.130 (0.290)
D11 - Health Score	-.343** (0.004)	.179 (0.144)
D12 - Education Score	-.178 (0.147)	.020 (0.868)
D13 - Crime Score	-.561** (0.000)	.038 (0.757)
D14 - Environment Score	-.478** (0.000)	-.097 (0.431)
D15 - Elections Score	-.481** (0.000)	-.107 (0.385)
D16 - Contracts Score	-.140 (0.256)	.037 (0.764)

68 Countries
** = p≤0.01, * = p≤0.05
Table 6: RTI Start Date and Total Score Globally

I next examine the association between the start date of RTI legislation as well as the RTI score a country's RTI legislation received and the strength of that country's open data initiative. I use Spearman's rho to test association. Table 6 displays the results of the test. Globally, the start date of a country's RTI legislation was significantly associated with the strength of that country's open data initiative (-0.603, p = 0.000). Additionally, mapping, census, public transport timetable, international trade health sector performance, crime statistics, national environmental statistics, and national election results data displayed statistically significant relationships with the start date of RTI legislation.

RTI total score however, did not hint at an association between the rating of a country's RTI legislation and the strength of its open data initiative.

Region	N	ODB Dataset Score	RTI Date	RTI Total Score
Americas	11	Country Overall Score	-.584 (0.059)	-.260 (0.440)
Asia	14	Country Overall Score	-.106 (0.719)	.191 (0.513)
Europe	27	Country Overall Score	-.361 (0.065)	.301 (0.127)
Sub Saharan Africa	10	Country Overall Score	-.293 (0.412)	-.170 (0.638)
Asia	14	D11 – Health Score	-.619* (0.018)	-.049 (0.868)
Americas	11	D13 - Crime Score	-.782** (0.004)	.466 (0.148)

Excluding Caribbean (N=1), MENA (N=3), Oceania (N=2)
** = p≤0.01, * = p≤0.05
Table 7: RTI Start Date and Total Score by Region

I conducted the same test as in Table 6, testing for association between the start date of RTI legislation, the score given to RTI legislation from the RTI ratings index, and the strength of a country's open data initiative, this time using region as a grouping variable. The results of this test are shown in Table 7. Although I tested across all datasets (D1, D4, D5, D9, D10, D11, D12, D13, D14, D15, and D16), I only include dataset types by region if they exhibit a statistically significant relationship. None of the regions exhibited a statistically significant relationship between both independent variables and the strength of a country's open data initiative.

4.3.1.2 Democratic Freedoms

4.3.1.2.1 Dataset Overview

This test contains two independent variables: political rights and civil liberties. The data are provided by Freedom House, a US-based watchdog NGO that analyzes levels of freedom

around the world. I use data from Freedom House's 2015 annual dataset on the level of democratic freedoms around the world.[168]

Freedom House's annual report scores political rights and civil liberties from 0-40 points and 0-60 points, respectively. The political rights variable is broken down into three subcategories: electoral process, political pluralism and participation, and functioning of government. Likewise, the civil liberties variable is broken down into: freedom of expression and belief, associational and organizational rights, rule of law, and personal autonomy and individual rights. Freedom House does not typically change a score given to a country from the previous year unless "there has been a real-world development during the year that warrants a decline or improvement."[169]

The scores themselves are produced by Freedom House's in-house team and external analysts and advisers, who come from academic, think tank, and human rights backgrounds. Analysts score countries based on events during the year; these scores are discussed and reviewed by Freedom House staff and expert advisers.

In order for a country to garner a high score in political rights, it must "enjoy a wide range of political rights, including free and fair elections;"[170] during these elections, the opposition must play an important role in, and interests of minority groups are both present and well represented in government. Scores lower due to factors such as political corruption, foreign and military influence on politics, restricted political rights, government oppression, or civil war.

[168] *Freedom House*, "Freedom in the World 2016," https://freedomhouse.org/report/freedom-world-2016/methodology.
[169] Ibid.
[170] Ibid.

Countries with high civil liberties scores are ones in which citizens enjoy "freedoms of expression, assembly, association, education, and religion."[171] There is a fair legal system that ensures rule of law and protects the equality of opportunity for its citizens. Lower scores correspond to restrictions in media, restrictions in trade union activities, and discrimination against minority groups and women.

4.3.1.2.2 Descriptive Statistics

Figure 10: Political Rights, Grouped by Region

Figure 10 displays a box and whisker plot of Freedom House's political rights variable on the y axis, and regional groupings on the x axis. The point labeled New Zealand again is both New Zealand as well as Australia; in both Figures 10 and 11, these two countries occupy the

[171] Ibid.

same point in the plots because they both received a 39 for political rights and a 58 for civil liberties in Freedom House's 2015 index.

Figure 11: Civil Liberties, Grouped by Region

	Political Rights	Civil Liberties
Mean	26.815	39.434
Standard Error	1.247	1.624
Median	29.5	39
Mode	39	58
Standard Deviation	11.966	15.581
Sample Variance	143.185	242.797
Kurtosis	-1.008	-1.185
Skewness	-0.603	-0.304
Range	38	53
Minimum	2	7
Maximum	40	60

Count = 92 Countries

Table 9: Democratic Freedoms Descriptive Statistics

Political Rights	Civil Liberties
Belgium (40)	Finland (60)
Denmark (40)	Norway (60)
Sweden (40)	Iceland (60)
United Kingdom (40)	Sweden (59)
Uruguay (40)	Canada (59)
…	…
Kazakhstan (6)	China (15)
Bahrain (5)	UAE (13)
Vietnam (3)	Ethiopia (11)
Saudi Arabia (3)	Bahrain (10)
China (2)	Saudi Arabia (7)

Table 8: Highest and Lowest Scoring Countries with Open Data Initiatives

Table 8 lists the top and bottom performing countries in Freedom House's 2015 dataset. Out of the bottom five performers, the United Arab Emirates ranks the highest in ODB's open data initiative dataset at #47. Ethiopia ranks the worst at #78. The other bottom performing countries range in the 50s. The top performers for both political rights and civil liberties all score high on the strength of their open data portals. Table 9 displays descriptive statistics on Freedom House's two variables. Both have a moderate left skew. Both distributions have means and medians close to one another (PR: mean = 26.815, median = 29.5, CL: mean = 39.434, median = 39).

4.3.1.2.3 Findings

Using Spearman's rho, I test for association between both political rights and civil liberties with the strength of open data initiatives globally. Table 10 displays the results.

	Political Rights	Civil Liberties
Country Overall Score	.708** (0.000)	.720** (0.000)
D1 - Map Score	.612** (0.000)	.619** (0.000)
D4 - Census Score	.531** (0.000)	.529** (0.000)
D5 - Budget Score	.513** (0.000)	.541** (0.000)
D9 - Transport Score	.444** (0.000)	.469** (0.000)
D10 - Trade Score	.580** (0.000)	.582** (0.000)
D11 - Health Score	.438** (0.000)	.442** (0.000)
D12 - Education Score	.459** (0.000)	.461** (0.000)
D13 - Crime Score	.591** (0.000)	.594** (0.000)
D14 - Environment Score	.558** (0.000)	.578** (0.000)
D15 - Elections Score	.651** (0.000)	.660** (0.000)
D16 - Contracts Score	.305** (0.003)	.305** (0.003)

92 Countries
** = p≤0.01, * = p≤0.05
Table 10: Democratic Freedoms Globally

Globally, both the overall score of a country's open data initiative as well as every dataset kind exhibited strong relationships with the strength of a country's open data initiative. These associations are all positive, meaning that as democratic freedoms increase, one can expect stronger open data initiatives from a country.

Table 11 displays democratic freedoms' association with the strength of open data initiatives separated by region. Again, although I tested across all datasets (D1, D4, D5, D9, D10, D11, D12, D13, D14, D15, and D16), I only include dataset kinds by region if they exhibit a statistically significant relationship. Out of the regions, Europe exhibits a strong relationship between both political rights (0.779, $p = 0.000$) and civil liberties (0.707, $p = 0.000$) and the overall country score for open data initiatives. Asia also exhibits a relationship between both political rights (0.530, $p = 0.029$) and civil liberties (0.685, $p = 0.002$) with the overall country score for open data initiatives. The Americas, MENA, and Sub Saharan Africa do not exhibit statistically significant associations.

Region	N	ODB Dataset Score	Political Rights	Civil Liberties
Americas	13	Country Overall Score	.518 (0.070)	.456 (0.117)
Asia	17	Country Overall Score	.530* (0.029)	.685** (0.002)
Europe	27	Country Overall Score	.779** (0.000)	.707** (0.000)
MENA	9	Country Overall Score	0.025 (0.949)	0.084 (0.830)
Sub Saharan Africa	21	Country Overall Score	.236 (0.304)	.239 (0.297)
Asia	17	D1 – Map Score	.712** (0.001)	.801** (0.000)
Europe	27	D1 – Map Score	.596** (0.001)	.645** (0.000)
Europe	27	D4 – Census Score	.596** (0.001)	.500** (0.008)
Europe	27	D5 – Budget Score	.433* (0.024)	.340 (0.083)
Americas	13	D9 - Transport Score	.650* (0.016)	.600* (0.030)
Asia	17	D9 – Transport Score	.519* (0.033)	.723** (0.001)
Europe	27	D10 – Trade Score	.600** (0.001)	.621** (0.001)
Asia	17	D11 – Health Score	.290 (0.259)	.507* (0.038)
Europe	27	D11 – Health Score	.519** (0.006)	.446* (0.020)
Europe	27	D12 – Education Score	.414* (0.032)	.391* (0.044)
MENA	9	D12- Education Score	.667* (0.050)	.693* (0.038)
Europe	27	D13 – Crime Score	.628** (0.000)	.625** (0.000)
Asia	17	D14 – Environment Score	.349 (0.169)	.535* (0.027)
Europe	27	D14 – Environment Score	.566** (0.002)	.478* (0.012)
Asia	17	D15 – Elections Score	.615** (0.009)	.587* (0.013)
Europe	27	D15 – Elections Score	.510** (0.007)	.523** (0.005)
Asia	17	D16 – Contracts Score	.521* (0.032)	.570* (0.017)
Americas	13	D16 – Contracts Score	.596* (0.032)	.557* (0.048)

Excluding Caribbean (N=3), Oceania (N=2)
** = p≤0.01, * = p≤0.05
Table 11: Democratic Freedoms by Region

4.3.1.3 Corruption

4.3.1.3.1 Dataset Overview

This test contains one independent variable: the level of perceived corruption. The data is from the Corruptions Perceptions Index (CPI), an annual index published since 1996 by the international NGO Transparency International.

The Corruptions Perceptions Index scores and ranks countries around the world based on "how corrupt a country's public sector is perceived to be"[172] on a scale of 0 to 100, 0 being the most corrupt score possible and 100 being the least.

The CPI is a composite index, meaning that it is a combination of polls and surveys from different institutions. These organizations include the Africa Development Bank, the Asia Development Bank, Political and Economic Risk Consultancy, the World Economic Forum, and others.[173] Transparency International also collects business people opinion surveys and country performance assessment surveys. Information from all these sources are rescaled and standardized before calculating the final CPI score.

The 2015 CPI index includes 91 out of the 92 countries in ODB's dataset. Saint Lucia is not included in CPI's index, and is therefore excluded in the following tests.

[172] *Transparency International*, "Corruptions Perceptions Index: In Detail," http://www.transparency.org/cpi2013/in_detail.
[173] "Corruptions Perceptions Index 2010: Long Methodological Brief," *Transparency International*, http://transparency.ee/cm/files/cpi2010_long_methodology_en.pdf.

4.3.1.3.2 Descriptive Statistics

Figure 12: Corruptions Perceptions Index, Grouped by Region

Figure 12 displays countries grouped by region, with the 2015 CPI score along the y axis and regions along the x axis.

	CPI Score
Mean	49.593
Standard Error	2.212
Median	44
Mode	38
Standard Deviation	21.105
Sample Variance	445.421
Kurtosis	-1.01
Skewness	0.471
Range	74
Minimum	17
Maximum	91

Count = 91 Countries

Table 13: CPI Descriptive Statistics

CPI Score
Denmark (91)
Finland (90)
Sweden (89)
New Zealand (88)
Netherlands (87)
…
Myanmar (22)
Zimbabwe (21)
Yemen (18)
Haiti (17)
Venezuela (17)

Table 12: Highest and Lowest Scoring Countries in CPI

Table 12 displays the top and bottom five countries in the CPI. While the top scores are occupied by northern European countries, New Zealand scores an 88, Singapore scores an 85, and Canada scores an 83 in the 2015 CPI. Out of the bottom five countries, Venezuela ranks the highest in terms of the strength of its open data initiative (ranked #78). The other four countries at the bottom of the CPI correspond to rankings from 88 to 92. Kenya, who ranks just above Myanmar in the CPI, bucks this trend, ranking #42 in ODB's dataset.

4.3.1.3.3 Findings

I first test the association between the CPI and the strength of open data initiatives using Spearman's rho. Table 14 displays the results. Both the overall country open data initiative score as well as all of the types of datasets exhibit a strong positive relationship with levels of perceived corruption. Public contracting data (0.257, p = 0.014) exhibits a weaker relationship

	Corruptions Perceptions Index Score
Country Overall Score	.743** (0.000)
D1 - Map Score	.622** (0.000)
D4 - Census Score	.596** (0.000)
D5 - Budget Score	.436** (0.000)
D9 - Transport Score	.560** (0.000)
D10 - Trade Score	.597** (0.000)
D11 - Health Score	.544** (0.000)
D12 - Education Score	.468** (0.000)
D13 - Crime Score	.587** (0.000)
D14 - Environment Score	.690** (0.000)
D15 - Elections Score	.477** (0.000)
D16 - Contracts Score	.257* (0.014)

91 Countries
** = p≤0.01, * = p≤0.05
Table 14: Corruption Globally

than other types of datasets.

I next test this relationship using region to group results.[174] The results are shown in Table 15. The Americas (0.757, p = 0.003), Asia (0.741, p = 0.001), and Europe (0.806, p = 0.000) all exhibit a positive relationship with the CPI. This suggests that as levels of perceived corruption decrease, one can expect stronger open data initiatives from a country. MENA (0.469, p = 0.203) and Sub Saharan Africa (0.170, p = 0.462) do not exhibit a relationship with the CPI.

[174] Similar to previous tests with results grouped by region, although I tested across all dataset types (D1, D4, D5, D9, D10, D11, D12, D13, D14, D15, and D16), dataset types are included only if they exhibit a statistically significant relationship.

Region	N	ODB Dataset Score	CPI Score
Americas	13	Country Overall Score	.757* (0.003)
Asia	17	Country Overall Score	.741* (0.001)
Europe	27	Country Overall Score	.806** (0.000)
MENA	9	Country Overall Score	.469 (0.203)
Sub Saharan Africa	21	Country Overall Score	.170 (0.462)
Americas	13	D1 – Map Score	.622* (0.023)
Europe	27	D1 – Map Score	.626** (0.000)
Americas	13	D4 – Census Score	.644* (0.018)
Europe	27	D4 – Census Score	.493** (0.009)
Americas	13	D5 – Budget Score	.643* (0.018)
Americas	13	D9 - Transport Score	.616* (0.025)
Asia	17	D9 – Transport Score	.648** (0.005)
Europe	27	D9 – Transport Score	.420* (0.029)
Americas	13	D10 – Trade Score	.678* (0.011)
Asia	17	D10 – Trade Score	.513* (0.035)
Europe	27	D10 – Trade Score	.680** (0.000)
Americas	13	D11 – Health Score	.629* (0.021)
Asia	17	D11 – Health Score	.728** (0.001)
Europe	27	D11 – Health Score	.583** (0.001)
Europe	27	D12 – Education Score	.466* (0.014)
Americas	13	D13 – Crime Score	.561* (0.046)
Asia	17	D13 – Crime Score	.638** (0.006)
Europe	27	D13 – Crime Score	.663** (0.000)
Americas	13	D14 – Environment Score	.601* (0.030)
Asia	17	D14 – Environment Score	.741** (0.001)
Europe	27	D14 – Environment Score	.660** (0.000)
Europe	27	D15 – Elections Score	.522** (0.005)
Americas	13	D16 – Contracts Score	.686** (0.010)

Excluding Caribbean (N=2), Oceania (N=2)
** = $p \leq 0.01$, * = $p \leq 0.05$
Table 15: Corruption by Region

4.3.2 Government Involvement and Support for Technology

This section addresses Research Aim 2, exploring the association between government involvement and support for technology and the strength of open data initiatives. The dependent variable, the strength of open data initiatives, uses expert-survey data provided by Open Data Barometer. The independent variables, OGP membership data, Open Data Charter membership data, and online service index scores, use data provided by the Open Government Partnership, the United Nations e-Government survey, and the Open Data Charter.

Research Aim 2: Explore the association between the level of support an open data initiative receives from its federal government and the strength of the country's national open data initiative.

- Specifically examining open data initiatives that involved the president/prime minister's office and the legal support provided to initiatives.

4.3.2.1 Involvement of President/Prime Minister, Other Agencies

4.3.2.1.1 Dataset Overview

The OGP is a multinational initiative that works to promote transparency, fight corruption, and strengthen governance. The OGP was launched September 2011 with eight founding members: Brazil, Indonesia, Mexico, Norway, Philippines, South Africa, the United Kingdom, and the USA.[175] Its "Open Government Declaration," a pledge for greater openness

[175] *Open Government Partnership,* "Open Government Declaration," https://www.opengovpartnership.org/about/open-government-declaration.

and transparency in government, which countries must commit to uphold in order to join the OGP, now has 75 participating countries.

While the OGP promotes open data initiatives, submitting an initial letter of intent to the OGP is not synonymous with starting an open data initiative. That being said, 42 of the 92 countries in ODB's index are participating nations in the OGP.

The OGP has taken on criticism over the past few years. Some argue that the rapid expansion of OGP is "a whitewash, giving governments a façade of openness behind which they write laws that restrict access to executive emails, forbid foreign funding of journalism, empower surveillance, and worse."[176] Countries are able to send in a letter of intent, garner media attention, and subsequently sidestep actual policy implementation and follow-through.

Nevertheless, the OGP is one of the most well-known multilateral initiative in the world of open data, and its advocates argue there are positive effects in government transparency and open data initiatives due to membership.

This test uses variables from the Open Government Partnership (OGP) "Public Access IRM SU Commitments Database," last updated March 2017.[177] Variables in the dataset are recorded as 0, 1, NA (not available), and NR (not reporting). I exclude entries that are recorded as NA and NR, and conduct Mann-Whitney U tests using the entries coded 0 and 1.

[176] Steven Adler, "Why the Open Government Partnership Needs a Reboot," *Global Investigative Journalism*, December 12, 2015, http://gijn.org/2015/12/12/why-the-open-government-partnership-needs-a-reboot/.
[177] Open Government Partnership, "Public Access_IRM SU Commitments Database March 2017," https://docs.google.com/spreadsheets/d/1MEF5zN41_4tW1pK5Ptp0vqK0HvA0ofwMN9XaXK-QVNU/edit?usp=sharing.

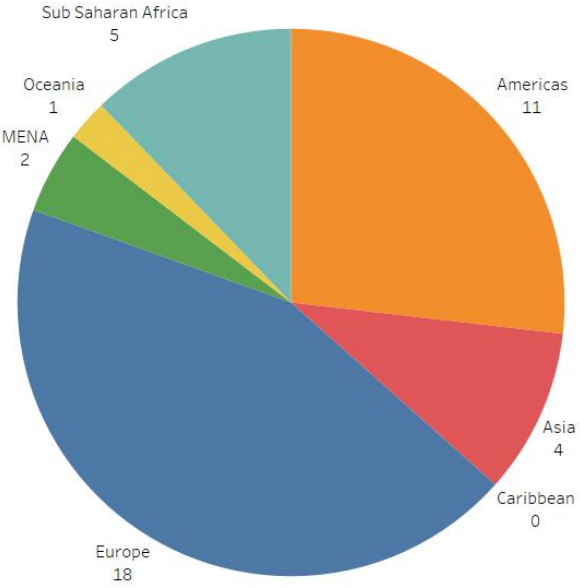
Independent variables are:

Variable	Description	0	1
President/Prime Minister	This variable indicates whether the President/Prime Minister, a committee of the President/Prime Minister, or a staff member within the President/Prime Minister's office is directly involved with OGP.	No	Yes
Change in Exec	This variable indicates whether the executive leader of a government changed during the duration of the OGP action plan development and implementation phase.	No	Yes
Exec Lead Agency	This variable indicates whether the country's Executive or its Executive branch (or equivalent) is the leading organization in organizing the country's OGP initiatives	No	Yes
Executive	This variable indicates whether the executive branch was involved in drafting/implementing the OGP action plan.	No	Yes
Implementing Agency	This variable indicates whether the government agency that developed the action plan is also the agency that implemented the action plan.	No	Yes
Legally Mandated	This variable indicates whether the government's commitment to OGP is established through a legally binding mandate.	No	Yes
Official Mandate	This variable indicates whether the government's commitment to OGP is established through an official, publically released mandate.	No	Yes
Ministry for Affair	This variable indicates whether the Department of State/MOFA is one of the agencies involved in contributing to the country's OGP development.	No	Yes
MOFA Lead Agency	This variable indicates whether the country's Ministry of Foreign Affairs (or equivalent) is the leading organization organizing the country's OGP initiatives.	No	Yes
Multiple Agencies	This variable tracks the number of government agencies involved in and engaging in the OGP process.	Only one government institution or agency is involved in the process.	Multiple government agencies or a working group is involved.
Multiple Arrangements	This variable indicates whether there was a change in the organization(s) leading or involved with the OGP initiatives during the development and implementation of the action plan.	No	Yes
Only One Branch	This variable indicates whether there was only one branch of government involved with the development and implementation of the action plan.	No	Yes

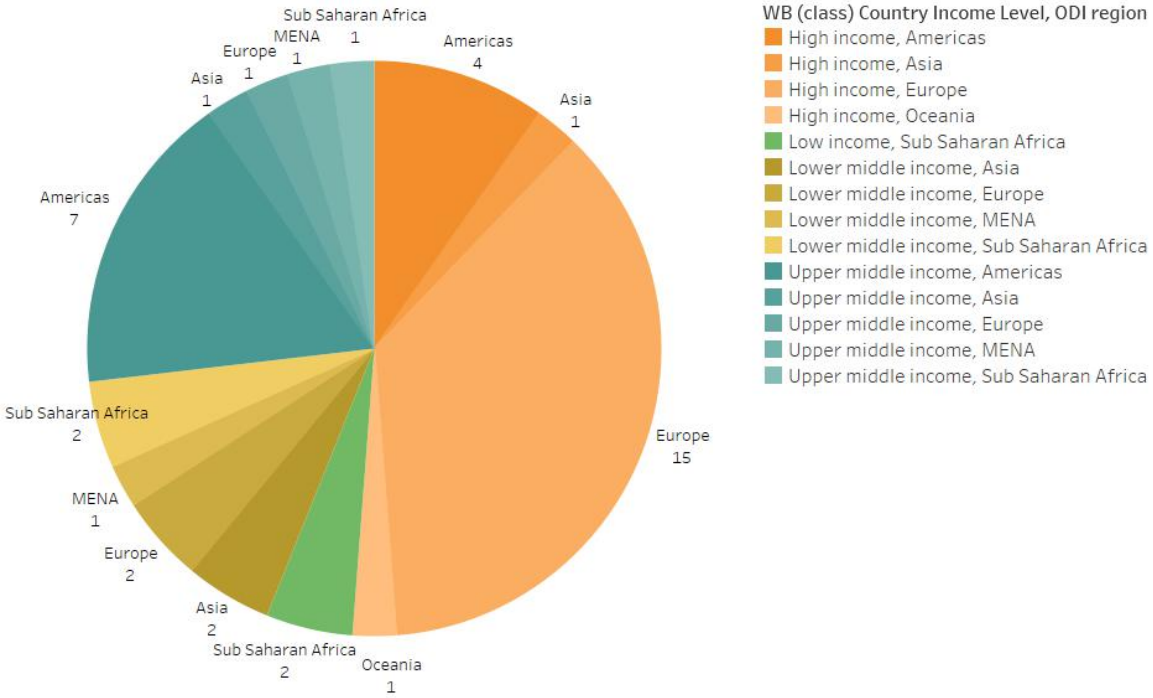
Table 16: OGP Independent Variables[178]

[178] Ibid.

4.3.2.1.2 Descriptive Statistics



Total Count = 41 Countries
Figure 13: Regional Breakdown of OGP



Total Count = 41 Countries
Figure 14: Income & Regional Breakdown of OGP

Income Level	Number of Participating Countries
High Income	21
Upper Middle Income	11
Lower Middle Income	7
Low Income	2

Total Count = 41 Countries
Table 17: Income Level of OGP Members

Income Level	Number of Countries
High Income	36
Upper Middle Income	24
Lower Middle Income	18
Low Income	14

Total Count = 92 Countries
Table 18: Income Level of Countries with Open Data Initiatives

Figures 13 and 14 display membership of the OGP based on region and income level. Table 19 displays income level of OGP member countries. Although the ratio of high income countries with open data initiatives to low income countries with open data initiatives in the world is ~2.57:1, participating high and low income countries in the OGP have a ratio of 10.5:1. Similarly, while the ratio of high income to lower middle income countries with open data portals in the world is 2:1 according to ODB's index, within OGP member countries, the ratio is 3:1.

4.3.2.1.3 Findings

I first test to see if there is a significant difference in the strength of open data initiatives between countries participating in the OGP and countries who do not. I run Mann-Whitney U tests for each OGP variable against the dependent variable measuring the strength of open data initiatives; Table 19 displays the results. While variables regarding the involvement of the President/Prime Minister do not return a significant difference in between initiatives involving the President/Prime Minister and initiatives that do not, the implementing agency variable returns a statistically significant difference between the two groups (MW = 39.000, Asymp Sig =

0.021, Exact Sig = 0.020). Additionally, the difference between OGP initiatives that are supported by legally binding legislation compared to initiatives that are not approaches statistical significance (MW = 18.000, Asymp Sig = 0.064, Exact Sig = 0.069).

OGP Member Data	0	1	Mann-Whitney U	Z	Asymp. Sig. (2-tailed)	Exact Sig. [2*(1-tailed Sig.)]
President/Prime Minister	17	9	66.000	-.566	0.571	.597
Change in Exec	16	10	75.000	-0.264	0.792	.816
Exec Lead Agency	14	7	47.000	-0.149	0.881	.913
Executive	1	25	3.000	-1.268	0.205	.308
Implementing Agency	12	14	39.000*	-2.316	0.021	.020
Legally Mandated	22	4	18.000	-1.849	0.064	.069

39 Countries
** = p≤0.01, * = p≤0.05
Table 19: OGP Participant Country Variables[179]

I next test to see if there is a statistically significant association between the year a country submits its first action plan to OGP, in which a government commits to specific transparency and governance pledges over a given period of time, and the strength of a country's open data initiative. Table 20 displays the results. Globally, the dataset type "Public transport timetable data" exhibited a negative relationship (-0.320, p = 0.042) with the strength of an open data portal. This suggests that the earlier an action plan was submitted to OGP, the stronger public transport timetable data is likely to be.

[179] A full table testing for difference between groups using all variables provided by OGP is available in the Appendix.

	OGP Year AP Submitted (earliest action plan submission)
Country Overall Score	-.154 (0.336)
D1 - Map Score	-.103 (0.522)
D4 - Census Score	.010 (0.949)
D5 - Budget Score	-.147 (0.359)
D9 - Transport Score	-.320* (0.042)
D10 - Trade Score	.091 (0.573)
D11 - Health Score	.022 (0.894)
D12 - Education Score	-.006 (0.968)
D13 - Crime Score	.027 (0.866)
D14 - Environment Score	-.101 (0.530)
D15 - Elections Score	-.051 (0.750)
D16 - Contracts Score	-.217 (0.173)

41 Countries
** = p≤0.01, * = p≤0.05
Table 20: OGP Year AP Submitted

4.3.2.2 Open Data Charter

Another multilateral initiative that has received attention from the open data community is the Open Data Charter. The charter was initially adopted by ten countries, five cities, and two states (Morelos, Mexico and Xalapa, Mexico) in 2015.[180] Unlike the OGP, which also deals with non-open government data initiatives, the Open Data Charter focuses exclusively on open government data. Its signatories commit themselves to releasing data that complies with specific openness principles.

[180] *Open Data Charter*, "Adopted By," http://opendatacharter.net/adopted-by-countries-and-cities/.

Income Level	Number of Participating Countries
High Income	7
Upper Middle Income	6
Lower Middle Income	3
Low Income	1

Count = 17 Countries
Table 21: Open Data Charter Income Level of Members

Participating countries are: Australia, Argentina, Chile, Colombia, Costa Rica, France, Guatemala, Italy, Mexico, Panama, Paraguay, Philippines, Sierra Leone, South Korea, Ukraine, the United Kingdom, and Uruguay. Open Data Barometer does not include Guatemala or Panama in its index.

I test to see if there is a significant difference between the strength of open data initiatives of countries who have adopted the Open Data Charter and countries who have not. Countries who are not a member are coded as 0 while countries who are members are coded as 1. Table 22 displays the results. There is no significant difference between the two groups.

	Mann-Whitney U	Z	Asymp. Sig. (2-tailed)
Country Overall Score	363.500	-2.262	0.024

0 (not a member) = 77, 1 (member) = 15, Total = 92
** = p≤0.01, * = p≤0.05
Table 22: Open Data Charter

4.3.2.3 Online Service Index

4.3.2.3.1 Dataset Overview

The Online Service Index is part of the United Nations' Public Administration Network's bi-annual e-government survey. This survey ranks countries around the world based on the "capacity and willingness of countries to use e-government for information and communication technology (ICT)-led development."[181] The index includes all UN member states.

One of the three components in creating the e-government readiness index (EGDI) is the online service index. This index measures a country's national website in terms of web content accessibility according to the World Wide Web Consortium's Web Content Accessibility Guidelines. When visiting countries' national websites, researchers follow a "citizen-centric approach" in assessing a government's online services; this means among other things, "putting themselves in the place of the average user. Thus responses [are] generally based on whether the relevant feature(s) could be found and accessed easily," not just whether or not a feature exists.[182] Final scores are calculated by subtracting the lowest scoring country, then dividing the value by the range of scores for all countries in the survey. The minimum score a country can receive is 20, and the maximum is 403. By dividing a country's score by the range of scores for all countries, the final value results in a decimal between 0 and 1.

[181] "United Nations e-Government Survey 2010," *United Nations Public Administration Network,* 124, http://unpan1.un.org/intradoc/groups/public/documents/un/unpan038851.pdf.
[182] Ibid.

4.3.2.3.2 Descriptive Statistics

Figure 15: E-gov Online Service Component, Grouped by Region

	Online Service Index
Mean	0.545
Standard Error	0.026
Median	0.555
Mode	0.307
Standard Deviation	0.254
Sample Variance	0.064
Kurtosis	-0.909
Skewness	-0.059
Range	0.976
Minimum	0.0236
Maximum	1

Count = 92 Countries

Table 24: E-gov Online Service Index Descriptive Statistics

Online Service Index
France (1)
Singapore (0.992)
South Korea (0.976)
Japan (0.945)
Spain (0.945)
…
Mali (0.134)
Haiti (0.110)
Benin (0.110)
Sierra Leone (0.047)
Myanmar (0.024)

Table 23: E-gov Highest and Lowest Scoring Countries

Figure 15 displays a box and whisker plot of countries, grouped by region on the x axis, with e-government online service index scores on the y axis. Tables 23 and 24 display descriptive statistics on the online service index and the top and bottom scoring countries in the

world with open data initiatives. The mean and median are 0.010 off from one another, the median being higher (0.555). The skewness suggests a slight left skew in the data.

4.3.2.3.3 Findings

I test for association between a country's score in its online service component and the strength of its open data portal. Table 25 displays the results. Both the country's overall strength of open data initiative score (0.783, p = 0.000) as well as all of the dataset kind scores exhibit statistically significant relationships with its online service score.

	Online Service Index Score
Country Overall Score	.783** (0.000)
D1 - Map Score	.603** (0.000)
D4 - Census Score	.628** (0.000)
D5 - Budget Score	.517** (0.000)
D9 - Transport Score	.685** (0.00)
D10 - Trade Score	.580** (0.000)
D11 - Health Score	.577** (0.000)
D12 - Education Score	.440** (0.000)
D13 - Crime Score	.567** (0.000)
D14 - Environment Score	.663** (0.000)
D15 - Elections Score	.474** (0.000)
D16 - Contracts Score	.359** (0.000)

92 Countries
** = p≤0.01, * = p≤0.05
Table 25: Online Service Index

I next test the relationship between a country's online service index score and the strength of its open data initiative. Table 26 displays the results.[183] The Americas (0.781, p = 0.002), Asia (0.757, p = 0.000), Europe (0.664, p = 0.000), MENA (0.700, p = 0.036), and Sub Saharan Africa (0.564, p = 0.008) all exhibit statistically significant positive relationships between online service index scores and data initiative scores. Across the 11 types of datasets, the Americas has seven kinds of datasets that also have p values ≤0.05, Asia has five, Europe has three, MENA has one, and Sub Saharan Africa has seven.

[183] Dataset types (D1, D4, D5, D9, D10, D11, D12, D13, D14, D15, and D16) are included only if they exhibit a statistically significant relationship.

Region	N	ODB Dataset Score	Online Service Index Score
Americas	13	Country Overall Score	.781** (0.002)
Asia	17	Country Overall Score	.757** (0.000)
Europe	27	Country Overall Score	.664** (0.000)
MENA	9	Country Overall Score	0.700* (0.036)
Sub Saharan Africa	21	Country Overall Score	.564** (0.008)
Americas	13	D1 – Map Score	.557* (0.048)
Asia	17	D1 – Map Score	.483* (0.049)
Sub Saharan Africa	21	D1 – Map Score	.494* (0.023)
Americas	13	D4 – Census Score	.651* (0.016)
Europe	27	D5 – Budget Score	.439* (0.022)
Sub Saharan Africa	21	D5 – Budget Score	.457* (0.037)
Americas	13	D9 – Transport Score	.786** (0.001)
Asia	17	D9 – Transport Score	.567* (0.018)
Sub Saharan Africa	21	D9 – Transport Score	.575** (0.006)
Americas	13	D10 – Trade Score	.571* (0.041)
Europe	27	D10 – Trade Score	.454* (0.017)
Americas	13	D11 – Health Score	.579* (0.038)
Asia	17	D11 – Health Score	.732** (0.001)
MENA	9	D11- Health Score	.684* (0.042)
Sub Saharan Africa	21	D12 – Education Score	.472* (0.031)
Americas	13	D13 – Crime Score	.657* (0.015)
Asia	17	D13 – Crime Score	.672** (0.003)
Asia	17	D14 – Environment Score	.751** (0.001)
Europe	27	D14 – Environment Score	.387* (0.046)
Sub Saharan Africa	21	D14 – Environment Score	.458* (0.037)
Sub Saharan Africa	21	D15 – Elections Score	.560** (0.008)
Americas	13	D16 – Contracts Score	.567* (0.043)
Sub Saharan Africa	21	D16 – Contracts Score	.454* (0.039)

Excluding Caribbean (N=3), Oceania (N=2) ** = $p \leq 0.01$, * = $p \leq 0.05$

Table 26: Online Service Index, by Region

4.3.3 Education Attainment and Technology Access and Adoption

This section addresses Research Aim 3, exploring the association between education attainment and technology access and adoption, and the strength of open data initiatives. The dependent variable, the strength of open data initiatives, uses expert-survey data provided by Open Data Barometer. The independent variables, EGDI, telecommunications infrastructure scores, human capital scores, and citizen views on science and technology, use data provided by the United Nations Public Administration Network.

Research Aim 3: Explore the association between the level of education and e-government familiarity and the strength of a national open data initiative.

- Specifically examining mean years of schooling, internet usage, views on science and technology, and e-government familiarity.

4.3.3.1 Dataset Overview

As discussed in 4.3.2.3, the E-Government Readiness Index (EGDI) measure "the capacity and willingness of countries to use e-government for ICT-led development."[184] The index is generated by the United Nations Public Administration Programme (UNPAP) since 2003, and its scope covers all members state of the United Nations.[185] In addition to its online services index, the EDGI also looks at telecommunication connectivity and human capacity in terms of e-government.[186]

[184] "United Nations e-Government Survey 2010," *United Nations Public Administration Network,* 124.
[185] E-Government Readiness Index," *Wikiprogress,* "http://wikiprogress.org/articles/governance-human-rights/e-government-readiness-index/#Online_Service_Index.
[186] Ibid.

The telecommunication infrastructure index is created with five indicators:[187]

- Number of personal computers per 100 persons
- Number of internet users per 100 persons
- Number of telephone lines per 100 persons
- Number of mobile cellular subscriptions per 100 persons
- Number of fixed broadband subscribers per 100 persons

All of this data is provided to the UNPAP from the International Telecommunication Union, a UN specialized information and communication technologies agency.[188] Final telecommunication infrastructure scores are calculated by adding these five indicators for a country together, then normalizing this value by subtracting the lowest scoring country in the telecommunication infrastructure survey and dividing by the range of scores.

The third component of calculating the EGDI is the human capital index. This score is a composite of two indicators: adult literacy rate and the combined primary, secondary, and tertiary gross enrollment ratio. This secondary data was both provided by the United Nations Educational, Scientific, and Cultural Organization. Just as the other two sub-indices are normalized, the lowest value from any country in the survey is subtracted from the value resulting from these two indicators. The resulting value is then divided by the range of scores in the human capital index.

The EGDI is then calculated using the online service, telecommunication, and human capital sub-indices. Each of these sub-indices is multiplied by one third and the resulting values are added together.

[187] Ibid.
[188] *International telecommunication Union,* "About," http://www.itu.int/en/about/Pages/default.aspx.

4.3.3.2 Descriptive Statistics

Figure 16: EGDI

Figure 17: Telecommunications Infrastructure

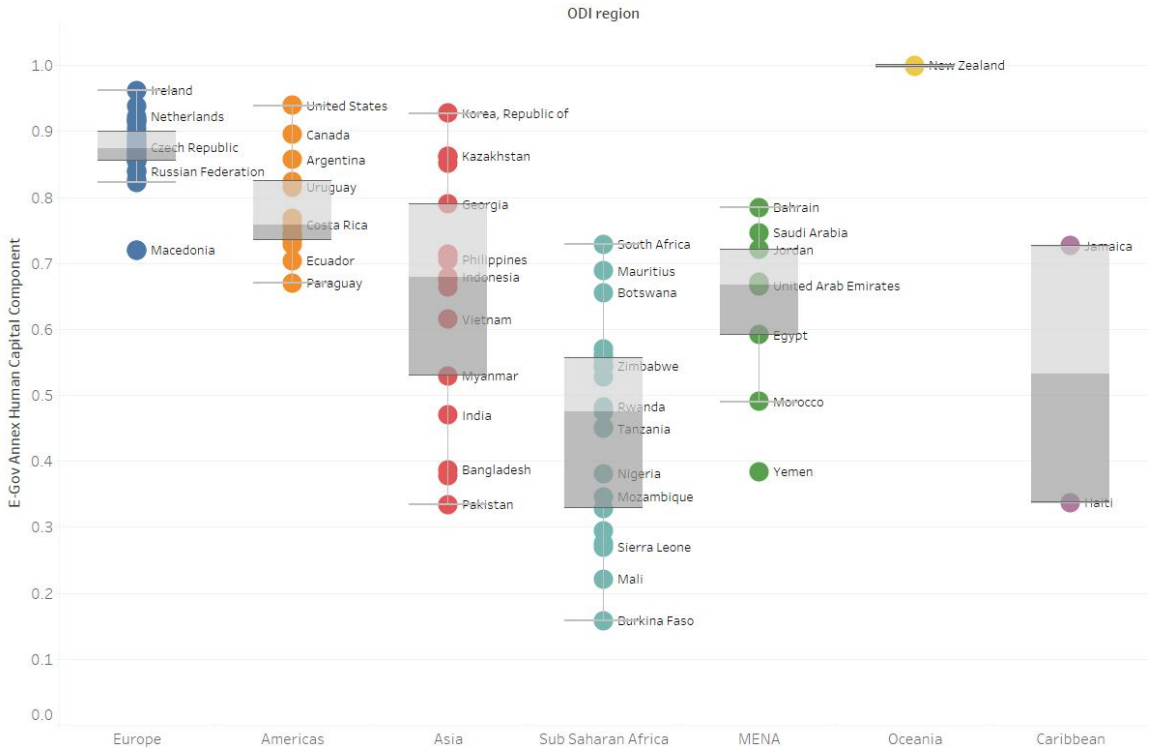


Figure 18: Human Capital

Figures 16, 17, and 18 are box and whisker plots of EGDI, telecommunications infrastructure, and human capital scores, respectively. The rankings of countries within its region does not change much regardless of the e-government sub index in question.

Table 27 displays the descriptive statistics associated with these three variables. Of the three, only the variable telecommunications infrastructure has a right skew. Standard error ranged from 0.021 to 0.027 between the three variables. Tables 28 displays the top and bottom five countries in their EGDI, telecommunications infrastructure, and human capital scores.

	EGDI	Telecomm Infrastructure	Human Capital
Mean	0.560	0.442	0.693
Standard Error	0.023	0.027	0.021
Median	0.556	0.411	0.727
Mode	0.816	#N/A	0.713
Standard Deviation	0.226	0.266	0.204
Sample Variance	0.051	0.071	0.0420
Kurtosis	-1.121	-1.219	-0.391
Skewness	-0.153	0.166	-0.739
Range	0.813	0.926	0.842
Minimum	0.132	0.008	0.157
Maximum	0.946	0.935	1

Count = 92 Countries

Table 27: E-Government Descriptive Statistics

EGDI	Telecomm Infrastructure	Human Capital
South Korea (0.946)	South Korea (0.935)	New Zealand (1)
Australia (0.910)	Sweden (0.886)	Australia (0.997)
Singapore (0.907)	Singapore (0.879)	Ireland (0.961)
France (0.893)	Denmark (0.874)	United States (0.939)
Netherlands (0.889)	Finland (0.859)	Norway (0.938)
…	…	…
Haiti (0.180)	Tanzania (0.080)	Ethiopia (0.293)
Burkina Faso (0.180)	Mozambique (0.054)	Benin (0.275)
Benin (0.168)	Malawi (0.048)	Sierra Leone (0.269)
Mali (0.163)	Ethiopia (0.026)	Mali (0.221)
Sierra Leone (0.132)	Myanmar (0.008)	Burkina Faso (0.157)

Count = 92 Countries

Table 28: E-Government Top and Bottom Scoring Countries

4.3.3.3 Findings

I first test these three E-Government indices globally, using the strength of open data initiatives as the dependent variable. The results are displayed in Table 29, and are similar to the tests of association with the online service index; all three variables display statistically

significant, positive relationships with a country's open data initiative. The country overall score test showed the strongest directional relationship when compared to the dataset kind association tests; EGDI (0.865, p = 0.000), telecommunication infrastructure (0.829, p = 0.000), and human capital (0.838, p = 0.000).

	E-Gov Annex EGDI	E-Gov Annex Telecomm Infrastructure Component	E-Gov Annex Human Capital Component
Country Overall Score	.865** (0.000)	.829** (0.000)	.838** (0.000)
D1 - Map Score	.676** (0.000)	.650** (0.000)	.691** (0.000)
D4 - Census Score	.689** (0.000)	.648** (0.000)	.662** (0.000)
D5 - Budget Score	.436** (0.000)	.436** (0.000)	.436** (0.000)
D9 - Transport Score	.697** (0.000)	.654** (0.000)	.648** (0.000)
D10 - Trade Score	.661** (0.000)	.660** (0.000)	.641** (0.000)
D11 - Health Score	.611** (0.000)	.582** (0.000)	.537** (0.000)
D12 - Education Score	.496** (0.000)	.482** (0.000)	.456** (0.000)
D13 - Crime Score	.693** (0.000)	.682** (0.000)	.731** (0.000)
D14 - Environment Score	.755** (0.000)	.748** (0.000)	.719** (0.000)
D15 - Elections Score	.577** (0.000)	.564** (0.000)	.620** (0.000)
D16 - Contracts Score	.334** (0.001)	.233* (0.025)	.302** (0.003)

92 Countries
*** = p≤0.001, ** = p≤0.01, * = p≤0.05
Table 29: EGDI, Telecomm Infrastructure, and Human Capital:

Next, I test these same e-government variables against the strength of open data initiatives using region as a grouping variable. Table 30 displays the results. Against the overall country open data initiative score, the Americas (0.615, p = 0.025), Asia (0.725, p = 0.001), Europe (0.843, p = 0.000, MENA (0.767, p = 0.016), and Sub Saharan Africa (0.458, p = 0.037) all displayed associations with p values ≤0.05 in EGDI tests. In telecommunications infrastructure tests and human capital tests, Asia (0.672, p = 0.003) and Europe (0.847, p =

0.000) both exhibit statistically significant relationships with p values less than both ≤0.05 and ≤0.01; the Americas, MENA, and Sub Saharan African do not display a significant relationship with these independent variables.

Region	N	ODB Dataset Score	E-Gov Annex EGDI	E-Gov Annex Telecomm Infrastructure Component	E-Gov Annex Human Capital Component
Americas	13	Country Overall Score	.615* (0.025)	.511 (0.074)	.401 (0.174)
Asia	17	Country Overall Score	.725** (0.001)	.672** (0.003)	.713** (0.001)
Europe	27	Country Overall Score	.843** (0.000)	.847** (0.000)	.554** (0.003)
MENA	9	Country Overall Score	.767* (0.016)	.450 (0.224)	.550 (0.125)
Sub Saharan Africa	21	Country Overall Score	.458* (0.037)	.157 (0.496)	.264 (0.248)
Americas	13	D1 – Map Score	.584* (0.036)	.576* (0.039)	.383 (0.197)
Asia	17	D1 – Map Score	.449 (0.071)	.483* (0.049)	.332 (0.193)
Europe	27	D1 – Map Score	.533** (0.004)	.614** (0.001)	.434* (0.024)
Americas	13	D4 – Census Score	.636* (0.019)	.600* (0.030)	.463* (0.111)
Europe	27	D4 – Census Score	.466* (0.014)	.429* (0.025)	.441* (0.021)
MENA	9	D4 – Census Score	.749* (0.020)	.402 (0.284)	.694* (0.038)
Europe	27	D5 – Budget Score	.518** (0.006)	.354 (0.070)	.541** (0.004)
Americas	13	D9 – Transport Score	.633* (0.020)	.469 (0.106)	.531 (0.062)
Asia	17	D9 – Transport Score	.525* (0.031)	.444* (0.074)	.493* (0.044)
Europe	27	D9 – Transport Score	.496** (0.008)	.575** (0.002)	.325 (0.098)
Americas	13	D10 – Trade Score	.596* (0.032)	.657* (0.015)	.503 (0.080)
Europe	27	D10 – Trade Score	.575** (0.002)	.632** (0.000)	0.379 (0.051)
MENA	9	D10 – Trade Score	.694* (0.038)	.347 (0.360)	.420 (0.260)
Asia	17	D11 – Health Score	.675** (0.003)	.654** (0.004)	.740** (0.001)

Europe	27	D11 - Health Score	.510** (0.007)	.676** (0.000)	0.298 (0.132)
Europe	27	D12 - Education Score	.368 (0.059)	.423* (0.028)	.214 (0.283)
Asia	17	D13 – Crime Score	.594* (0.012)	.564* (0.018)	.611** (0.009)
Europe	27	D13 - Crime Score	.547** (0.003)	.634** (0.000)	.300 (0.128)
Sub Saharan Africa	21	D13 – Crime Score	.589** (0.005)	.434* (0.049)	.560** (0.008)
Asia	17	D14 – Environment Score	.748** (0.001)	.756** (0.000)	.729** (0.001)
Europe	27	D14 - Environment Score	.499** (0.008)	.624** (0.001)	0.300 (0.128)
Sub Saharan Africa	21	D14 – Environment Score	.473* (0.030)	0.203 (0.379)	.389 (0.082)
Europe	27	D15 - Elections Score	.439* (0.022)	.533** (0.004)	.163 (0.416)
Sub Saharan Africa	21	D15 – Elections Score	.631** (0.002)	.360 (0.109)	.639** (0.002)
Sub Saharan Africa	21	D16 – Contracts Score	.476* (0.029)	.240 (0.294)	.350 (0.120)

Excluding Caribbean (N=3), Oceania (N=2)
** = p≤0.01, * = p≤0.05
Table 30: E-Government, by Region[189]

4.3.4 Involvement of International Organizations

This section addresses Research Aim 4, exploring the association between the involvement of international organizations in a country's open data initiative and the strength of open data initiative. The dependent variable, the strength of open data initiatives, uses expert-survey data provided by Open Data Barometer. The independent variables: Number of NGOs involved in open data initiatives, number of open data programs sponsored by NGOs/IGOs the government is involved in, and number of open data programs sponsored by NGOs/IGOs citizens and local NGOs are involved in, use data I collected for inclusion of an open data program in the dataset.

[189] Dataset types (D1, D4, D5, D9, D10, D11, D12, D13, D14, D15, and D16) are included only if they exhibit a statistically significant relationship.

Research Aim 4: Explore the association between the involvement of international organizations in a country's open data initiative, the participation of a government in open data organizations and programs, and the strength of its open data initiative.

4.3.4.1 Dataset Overview

Although NGOs play a large role in developing national open data initiatives by providing expert advice, best practices, financial resources, and access to their own open data community, there is little research that quantitatively evaluates the relationship between NGOs and the strength of an open data initiative.

In order to test for association between NGOs promoting open data initiatives and the strength of an open data portal, I first needed to choose a metric in order to evaluate NGOs' presence in a country in terms of open data.

While budgetary and spending data from NGOs may have provided deeper insights into their roles in fostering open data initiatives, finding budgetary and spending data at a precise-enough level to determine that "x" amount of dollars was allocated to open data initiatives is near impossible. While NGO budgetary and spending data is often open in 2017,[190] their data does not reach the level of granularity in order to determine how much spending was allocated to open data. Additionally, as open data initiatives are nested within multiple sectors of NGO work, such as transparency or smart city initiatives, creating a comprehensive dataset taking this approach (using relevant sectors to locate open data initiatives) may yielded a dataset that is stronger in its depth than its breadth.

[190] "FY16 World Bank Budget," *World Bank*, September 25, 2015, http://documents.worldbank.org/curated/en/525671468188047741/pdf/100629-BR-R2015-0106-1-Box393233B-PUBLIC.pdf.

Instead, I chose a simple count metric, and set strict criteria for a program's inclusion as a component of the NGO indicator variable.

Criteria for inclusion are:

1. Organization must be currently working on or have completed a project that promotes or develops open data within a country
    a. Project cannot be a one-off
2. On either the organization's "goals" or "about" page (or equivalent), organization must mention open data as its own goal, not listed in passing as one of many tactics to achieve some other goal (eg. Transparency, anti-corruption, etc.)

This criterion is meant to focus the organizations included in the NGO indicator variable, and to minimize noise and avoid creating a variable with so many accessory component organizations that the NGO indicator variable will exhibit a relationship with the dependent variable regardless of if the test is actually linking a third variable (eg. transparency initiatives) and the strength of an open data initiative.

In order to identify these NGOs, I began with two resources from the World Bank: (1) its technical assistance and funding page, which lists World Bank initiatives and programs that "use open data to address specific challenges."[191] (2) the World Bank's Open Budgets Portal's "Tools and Resources" page, which lists international initiatives the World Bank sponsors or partners with that work with open data and fiscal transparency and accountability.[192] While evaluating these organizations, I used suggested links on each website to gather new potential NGOs to evaluate. After these lists and related links were exhausted, I evaluated several other

[191] "Technical Assistance and Funding," *The World Bank,* http://opendatatoolkit.worldbank.org/en/technical-assistance.html.
[192] "Open Budgets Portal: Tools and Resources," *The World Bank,* http://wbi.worldbank.org/boost/tools-resources/intl-initiatives.

organizations with which I had prior knowledge of. In order to assess the stated goals of each NGO, I accessed each NGO webpage through Google searches.

The variables generated from this research are based upon four ways I observe NGOs who meet the stated criteria interact with countries and open data initiatives:

1. The government is a member/participant/adopter of an NGO's program
2. Individual citizens/local NGOs within a given country are members of an international NGO
3. A country has an open data project occurring in their country, but the government is not part of it
4. Both local NGOs and the government are members of an international NGO/multilateral open data initiative

To calculate scores for each country, each of these sub variables are calculated, then added to create a total tally of NGO involvement in a country.

4.3.4.2 Descriptive Statistics

Table 31 displays the organizations I use in creating the variable "Total Count," as well as the variables "Government is member," and "Individuals/NGOs Within Country are Member," along with their individual categorization.[193] Table 32 displays the representation breakdown regionally of countries included in these tests, and Figure 19 and Table 33 describe the median number of organizations a country in a specific region is engaged with.

[193] See Appendix for list of organizations.

Description	Count in Dataset
The government is a member/participant/adopter of an NGO's program	9
Individual citizens/local NGOs within a given country are members of an international NGO	5
A country has an open data project occurring in their country, but the government is not part of it	3
Both local NGOs and the government are members of an international NGO/multilateral open data initiative	1

Table 31: Organization Count by Government/NGO Interaction Types

Region	Number of Countries Represented	Income Level	Number of Countries Represented, Grouped by Income
Americas	13	High Income	4
		Upper Middle Income	9
Asia	17	High Income	3
		Upper Middle Income	6
		Lower Middle Income	7
		Low Income	1
Caribbean	3	Low Income	1
		Upper Middle Income	2
Europe	27	High Income	23
		Upper Middle Income	2
		Lower Middle Income	2
MENA	9	High Income	4
		Upper Middle Income	2
		Lower Middle Income	2
Oceania	2	High Income	2
Sub Saharan Africa	21	Upper Middle Income	4
		Lower Middle Income	5
		Low Income	12

Table 32: NGO Indicator Regional and Income Distribution

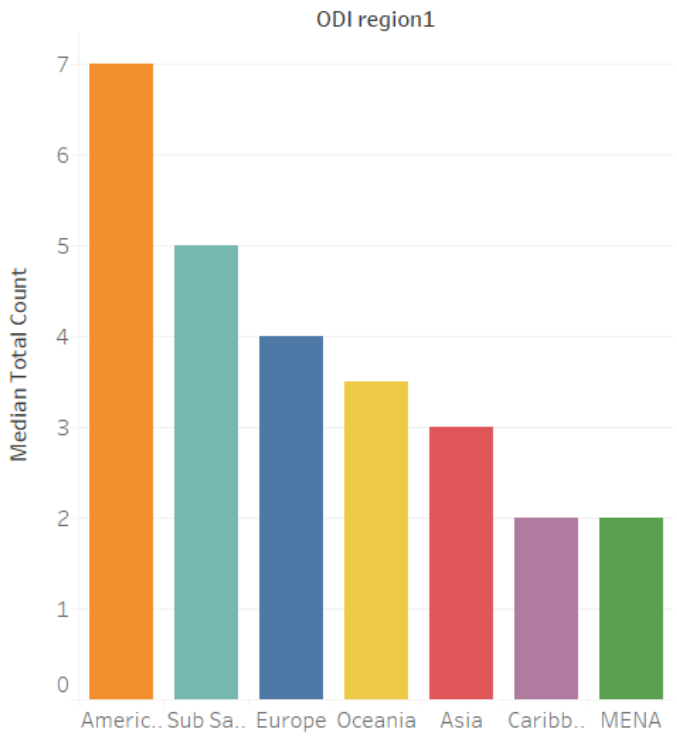


Figure 19: Median Total Count, by Region

Region	Median Gov is member Count	Median Individuals/NGOs within country are member Count	Median Total Count
Americas	2	3	7
Sub Saharan Africa	1	2	5
Europe	1	3	4
Oceania	1	2.5	3.5
Asia	0	2	3
Caribbean	1	1	2
MENA	1	1	2

Table 33: Median Number of NGOs per Country, by Region

4.3.4.3 Findings

I first test these three variables globally, using the strength of open data initiatives as the dependent variable. The results are displayed in Table 34. All three variables display statistically significant, positive relationships with the country's open data initiative.

For Total count, D14 – Environment Scores (.755, p = 0.000), exhibits the strongest directional relationship in comparison to the other dependent variables. For the other two independent variables, D13 – Crime Score, shows the strongest directional relationship.

	Total Count	Gov is Member Count	Individuals/NGOs Within Country are Member Count
Country Overall Score	.305** (0.003)	.302** (0.003)	.460** (0.000)
D1 - Map Score	.676** (0.000)	.650** (0.000)	.691** (0.000)
D4 - Census Score	.689** (0.000)	.648** (0.000)	.662** (0.000)
D5 - Budget Score	.436** (0.000)	.436** (0.000)	.436** (0.000)
D9 - Transport Score	.697** (0.000)	.654** (0.000)	.648** (0.000)
D10 - Trade Score	.661** (0.000)	.660** (0.000)	.641** (0.000)
D11 - Health Score	.611** (0.000)	.582** (0.000)	.537** (0.000)
D12 - Education Score	.496** (0.000)	.482** (0.000)	.456** (0.000)
D13 - Crime Score	.693** (0.000)	.682** (0.000)	.731** (0.000)
D14 - Environment Score	.755** (0.000)	.748** (0.000)	.719** (0.000)
D15 - Elections Score	.577** (0.000)	.564** (0.000)	.620** (0.000)
D16 - Contracts Score	.334** (0.001)	.233* (0.025)	.302** (0.003)

Excluding Caribbean (N=3), Oceania (N=2)
** = p≤0.01, * = p≤0.05

Table 34: NGO Involvement Globally

When grouping the results by region, the Americas (Total Count: .558, p = 0.048, Individ: .583, p = 0.037), Asia (Individ: .487, p = 0.048), and Sub Saharan Africa (Individ: .488, p = 0.025) all exhibit statistically significant relationships ≤0.05 with the overall country open data initiative score.

Region	N	ODB Dataset Score	Total Count	Gov is Member Count	Individuals/NGOs Within Country are Member Count
Americas	13	Country Overall Score	.558* (0.048)	.073 (0.813)	.583* (0.037)
Asia	17	Country Overall Score	.172 (0.510)	.390 (0.122)	.487* (0.048)
Europe	27	Country Overall Score	.315 (0.104)	.249 (0.222)	.322 (0.094)
MENA	9	Country Overall Score	.289 (0.451)	.381 (0.311)	.375 (0.321)
Sub Saharan Africa	21	Country Overall Score	.368 (0.100)	.001 (0.995)	.488* (0.025)
Americas	13	D1 – Map Score	.346 (0.247)	-.236 (0.438)	.629* (0.021)
Americas	13	D4 – Census Score	.504 (0.079)	-.083 (0.789)	.629* (0.021)
Europe	27	D5 – Budget Score	.459* (0.016)	.251 (0.207)	.193 (0.335)
Asia	17	D9 – Transport Score	.163 (0.532)	.193 (.457)	.555* (0.021)
Europe	27	D11 - Health Score	.192 (0.338)	.450* (0.019)	.172 (0.390)
MENA	9	D12 – Education Score	.842** (0.004)	.780* (0.013)	.626 (0.072)
MENA	9	D13 – Crime Score	.206 (0.594)	.741* (0.022)	.057 (0.885)
Americas	13	D16 – Contracts Score	.522 (0.067)	.659* (0.014)	.595* (0.032)
Asia	17	D16 – Contract Score	.566* (0.018)	.441 (0.076)	.628** (0.007)

Excluding Caribbean (N=3), Oceania (N=2)
** = p≤0.01, * = p≤0.05
Table 35: NGO Involvement, by Region[194]

[194] Dataset types (D1, D4, D5, D9, D10, D11, D12, D13, D14, D15, and D16) are included only if they exhibit a statistically significant relationship.

Chapter 5: Discussion of Findings

This chapter presents analysis and discussion based on the findings presented in Chapter 4. The order of the chapter follows the four angles from which I evaluate the factors influencing the path of a national open data initiative: (1) cultures of transparency and accountability; (2) government involvement and support for technology; (3) education attainment and technology access and adoption; and (4) involvement of international organizations.

5.1 Cultures of Transparency and Accountability

As one of the most frequently cited factors shaping the culture of open data in a country, cultures of transparency and accountability is possibly the most well-studied topic in relation to open data. Some open data advocates argue that FOI and cultures of transparency and accountability are not only a factor in the development of a portal, but a necessary condition that must be in place in order for (1) the initiative to be created in the first place and (2) for the initiative to be successful as well as a "true SDG accelerator."[195]

One organizational proponent of this argument is the Open Data Barometer. In its 2015 report, ODB calculated open data initiative scores, readiness scores, and implementation and impact scores. In calculating readiness scores, ODB leveraged its own expert surveys along with secondary data from Freedom House, the World Bank, the UN e-government survey, and the World Economic Forum. The selection of secondary data was selected based on the Web Foundation's qualitative Open Data Readiness assessment from 2010 to 2013, recommendations

[195] Open Data Barometer, "Global Report," 8.
Molly Schwartz, "Democracy and Open Data: Are the Two Linked?" May 22, 2014, http://congressionaldata.org/democracy-and-open-data-are-the-two-linked/.
"Open Data for Sustainable Development," *World Bank Group*, August 2015, 8, http://pubdocs.worldbank.org/en/999161440616941994/Open-Data-for-Sustainable-Development.pdf.
"Open Data, Transparency and Accountability: Topic Guide," *GSDRC*, September. 2016, 6, http://www.gsdrc.org/wp-content/uploads/2016/09/OpenDataTA_GSDRC.pdf.

from the International Open Data Charter, and the Open Data in Developing Countries research project.[196] ODB states that these readiness scores are not meant to measure readiness to start an open data initiative, but to measure readiness to secure positive outcomes from the initiative.[197] Embedded in this claim are the assumptions that the variables use to calculate its readiness scores, firm-level technology absorption (WEFGC), number of internet users per 100 people (World Bank), and democratic freedoms (Freedom House), are positively related to the strength of an open data portal. Moreover, ODB claims that its readiness scores, calculated from these variables, will measure the readiness of a country to secure positive outcomes in its open data initiatives. In other words, ODB believes these variables get at the underlying currents that determine the eventual success of an open data initiative.

Analysis of countries with open data initiatives indicates that there is a statistically significant relationship between the date FOI legislation is first introduced and the strength of a country's open data initiative, democratic freedoms and the strength of a country's open data initiative, and levels of corruption and the strength of a country's open data initiative.[198] ODB's choice of Freedom House's freedoms index as a component to calculating readiness scores is supported by my findings.

However, tests across all countries with open data initiatives compared to tests grouping these countries by region reveal different associations between regions and the strength of open data initiatives. For the introduction date of RTI legislation, both the country overall score as well as eight out of the eleven kinds of dataset examined exhibited statistically significant

[196] *Open Data Barometer*, "ODB Methodology – v1.0 28th April 2015," 4.
[197] Ibid.
[198] See Tables 6, 10, and 14.

negative relationships with the strength of open data initiatives.[199] The Americas (-.584, p = 0.059) and Europe (-.361, p = 0.065) both approached significance, but Asia and Sub Saharan Africa both showed little association with the date of RTI legislation introduction. MENA has three countries with open data initiatives and RTI legislation and was not included in this test.

This analysis at the regional level suggests caution in evaluating various democratic factors, such as the year of RTI legislation introduction, against the strength of open data initiatives. While this relationship may well be true in certain places, it should not be assumed that all regions share the same history with RTI legislation, freedoms of access, or even democracy in general. In assuming identical relationships with RTI histories across regions, other factors that shape an open data initiative may be lost. As a result, policy recommendations and future open data programs based upon the global association test between the date of RTI legislation introduction and the strength of an open data initiative (-.603, p = 0.000) may miss the regional nuance and unique relationship between these two variables.

In addressing FOI legislation, I also tested for an association between the RTI ratings' overall score of a country's RTI legislation. RTI ratings are not based on implementation of the actual legislation, but rather the robustness of the text of the legislation. Analysis does not show statistically significant relationships between the RTI overall score and the overall open data initiative score, nor any of the dataset kind scores.[200] The Centre for Law and Democracy themselves have acknowledged the weaknesses of creating an index about RTI legislation in which Serbia ranks #1 in the world.[201] Nevertheless, the fact that both global and regional tests

[199] D5- Detailed budget data, D12 – Primary and secondary education performance data, D16 – Public contracting data all did not have p values lower than 0.05, although D5 – Budget (-.209, p = 0.087), approached significance
[200] See Table 6.
[201] "RTI Rating Data Analysis Series: Overview of Results and Trends," Centre for Law and Democracy, September 28, 2013, http://www.law-democracy.org/live/wp-content/uploads/2013/09/Report-1.13.09.Overview-of-RTI-Rating.pdf.

did not show relationships that neared significance reinforces the idea that the robustness of written legislation does not imply implementation of the legislation.

Analysis of political rights and civil liberties for all countries with open data initiatives suggests strong positive relationships between the level of democratic freedoms and the strength of an open data initiative. This is the case not only for democratic freedoms' relationship with the overall country score for an open data initiative, but also for every dataset kind. This analysis supports ODB's choice in including democratic freedoms as a subcomponent of their readiness scores.

Region	Dataset Kind Count
Europe	9
Asia	6
Americas	2
MENA	1
Sub Saharan Africa	0

Table 36: Democratic Freedoms Regional Count[202]

Dataset Kind	Count
D1 – Map	2
D9 - Transport	2
D11 – Health	2
D12 – Education	2
D14 – Environment	2
D15 – Elections	2
D16 – Contracts	2
D4 – Census	1
D5 - Budget	1
D13 – Crime	1

Table 37: Democratic Freedoms Dataset Kind Count[203]

Table 36 displays the results of another method of analyzing these findings: the number of instances in which a region exhibited statistically significant relationships with at least one of the two variables measuring democratic freedoms. Table 37 displays the number of instances a dataset kind exhibited a statistically significant relationship with a region. Europe has the most types of datasets exhibiting statistically significant relationships with democratic freedoms. This suggests that its Ministries, which are the ones that actually choose to upload and publish the

[202] Region counted as one instance if one of its dataset kinds displays a statistically significant relationship with one or more variables (in this case, political rights and civil liberties).
[203] Dataset kind counted as one instance for each time it exhibits a statistically significant relationship with one or more of the independent variables, regardless of region.

kinds of data found in these dataset types (e.g. mapping data, health sector performance data, national environmental statistics data), have a positive relationship in publishing better quality data as democratic freedoms increase. Although Asia has six types of datasets that display statistically significant relationships with democratic freedoms, apart from mapping and public transport timetable data, Asian data exhibit significance levels between 0.01 and 0.05. Europe, on the other hand, exhibits significance levels below 0.01 in both political rights and civil liberties in its mapping, census, trade, crime, and elections data kinds.

Similar to democratic freedoms, the level of perceived corruption also exhibits a strong positive relationship with the strength of an open data portal.[204] When testing across all countries with open data initiatives, with the exception of public contracting data, every other dependent variable: the overall country score and the 11 kinds of datasets, all reported p values ≤ 0.001.

Region	Dataset Kind Count
Americas	9
Europe	9
Asia	5
MENA	0
Sub Saharan Africa	0

Table 38: CPI Regional Count

The most frequently occurring types of datasets exhibiting association were Transport, Trade, Health, Crime, and Environment, each occurring three times. Other than tests for association with EGDI readiness scores, levels of perceived corruption exhibit the second-most robust relationship across the 11 types of datasets.

[204] See Table 14.

Analysis supports existing literature that ties open data initiatives to corruption. A limitation of this test is that it does not determine which variable, levels of perceived corruption, or the strength of an open data portal, is the influencing factor on the other. Given open data's relatively brief existence, analysis suggests that lower levels of perceived corruption correspond to stronger open data initiatives in the Americas, Europe, and Asia. Another limitation of tests of association are that they (1) require a larger sample size than Oceania (N = 2) to run and (2) if one of the variables contains data that lacks variation, the reported values in tests of association do not sufficiently explain the relationship between the two variables.

Figure 20 displays a scatterplot in which CPI, the independent variable, runs along the x axis, while the strength of open data initiatives, the dependent variable, runs along the y axis. While Europe, the Americas, and Asia have varied values for their x coordinate, MENA, with the exception of Tunisia, does not score higher than a 30 for the strength of their data initiatives. Similarly, countries in Sub Saharan Africa do not have open data initiative scores above 30, and the data instead clusters slightly under a score of 10 for the strength of their data initiatives and ~30 for CPI scores. This does not mean that both regions' open data initiatives are not receptive to anti-corruption programs; there are many barriers to creating a strong open data portal, and corruption is just one of them. In fact, if both these regions did not face other challenges in creating strong open data initiatives, their plot in Figure 20 may reflect a positive relationship akin to the other regions. Possible explanations as to why MENA and Sub Saharan Africa both

stagnate in the strength of their open data initiatives are a lack of machine readable data[205] and strong cultures of secrecy within many of their governments.[206]

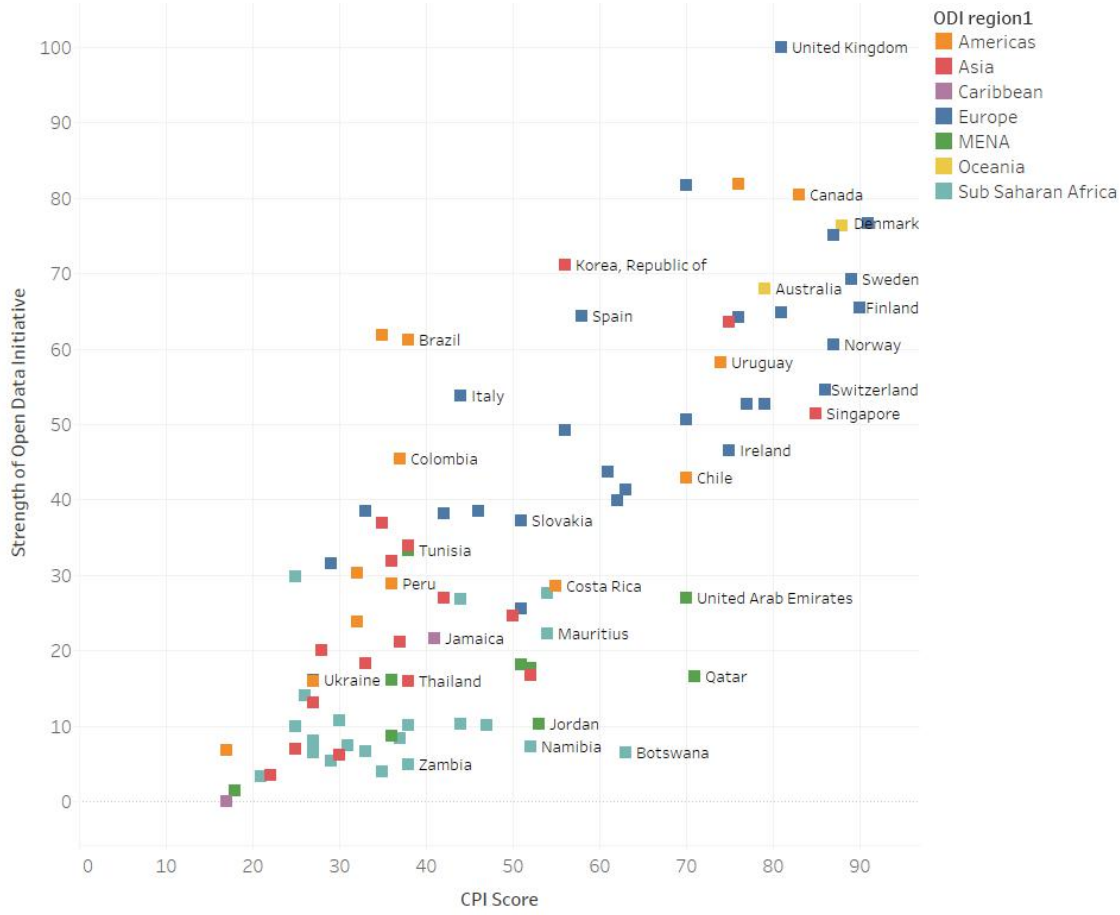


Figure 20: CPI Scatterplot

[205] Valentina Romei, *FT Data,* "Experts Raise Alarm Over Rise in "Open Data Inequality," April 21, 2016, http://blogs.ft.com/ftdata/2016/04/21/experts-raise-alarm-over-rise-in-open-data-inequality/.

[206] Advancing Access to Information in MNA: Supporting Coalitions & Networks," *World Bank Institute and Middle East and North Africa Social Development Department,* "May, 2012, http://documents.worldbank.org/curated/en/945061468299331196/pdf/NonAsciiFileName0.pdf. Mamadou Dia, *World Bank, A Governance Approach to Civil Service Reform in Sub-Saharan Africa,* International Bank for Reconstruction (1993).

5.2 Government Involvement and Support for Technology

As media attention over smart city and IoT initiatives continues to grow, it is likely that legislators, local government officials, and especially Presidents/Prime Ministers, whose time in office is dependent on perceived achievements by their constituents, will join and tout the benefits of open data initiatives.

Media attention is important, but follow-through on open data initiatives requires financial and human capital investment. A truly successful portal cannot be generated due only to one department's enthusiasm, nor can the quality of the data uploaded on the portal be outsourced to NGOs or private companies. Instead, successful portals require consultation with experts, support from multiple ministries or departments in government, and extensive training of government workers on practices of creating open, machine readable, and when necessary, anonymized data. They also require public support for expenditure of national resources for this endeavor.

Analysis of the Mann-Whitney U tests of OGP data indicate that the involvement of a President/Prime Minister is not associated with the quality of an open data initiative. In fact, OGP was meticulous in recording who from the executive office was involved in the development of the OGP action plan. Testing OGP member countries whose President/Prime Minster was not involved (group 0) against OGP member countries whose President/Prime Minister was involved (group 1), did show a relationship between involvement and the score of the open data initiative (MW = 66.000, Asymp Sig = 0.571, Exact Sig = 0.597). Additionally, whether or not the executive branch was the leading organization in joining OGP, whether or not there was a change of the leader in the executive branch during OGP membership, and whether or not the executive branch themselves were involved in drafting and implementing the OGP

action plan, all did not exhibit statistically significant differences between these groups of countries in the strength of their open data initiative.

Two variables that did exhibit a significant difference between two groups of countries were: (1) whether the government agency that developed the action plan is also the agency that implemented the plan (MW = 39.000, Asymp Sig = 0.021, Exact Sig = 0.020) and (2) whether governments' commitment to OGP is established through a legally binding mandate (MW = 18.000, Asymp Sig = 0.064, Exact Sig = 0.69). This finding suggests that while press noise over joining the Open Government Partnership can be created by a President or Prime Minister, their actual involvement is not associated with the eventual quality of the open data initiative. Unless the President/Prime Minister happens to have experience with data science, actionable input on developing a country's OGP action plan will come from the specialized team or agency that will implement the plan.

I also tested if there is a statistically significant difference between member states of the Open Data Charter and nations with open data initiatives who are not members. Members of this charter, while agreeing to stronger commitments to releasing data according to the principles of openness, already had open data initiatives. Therefore, this test is more to highlight a multilateral open data initiative other than OGP that has received praise, rather than to see if membership in the Open Data Charter makes a marked difference after one year.

The Online Service Index, part of EGDI's index, scores all UN member states on the ease-of-use and features of their national website. While these scores are representative of a number of factors, ranging from climate change denial to cultures of secrecy, this index is one of the closest proxies to evaluating the tools a national government has for web development initiatives.

Analysis shows that out of all countries with open data initiatives, the online service index exhibited a strongly positive relationship with the country overall score as well as with every dataset kind score (for all global tests, $p \leq 0.001$).

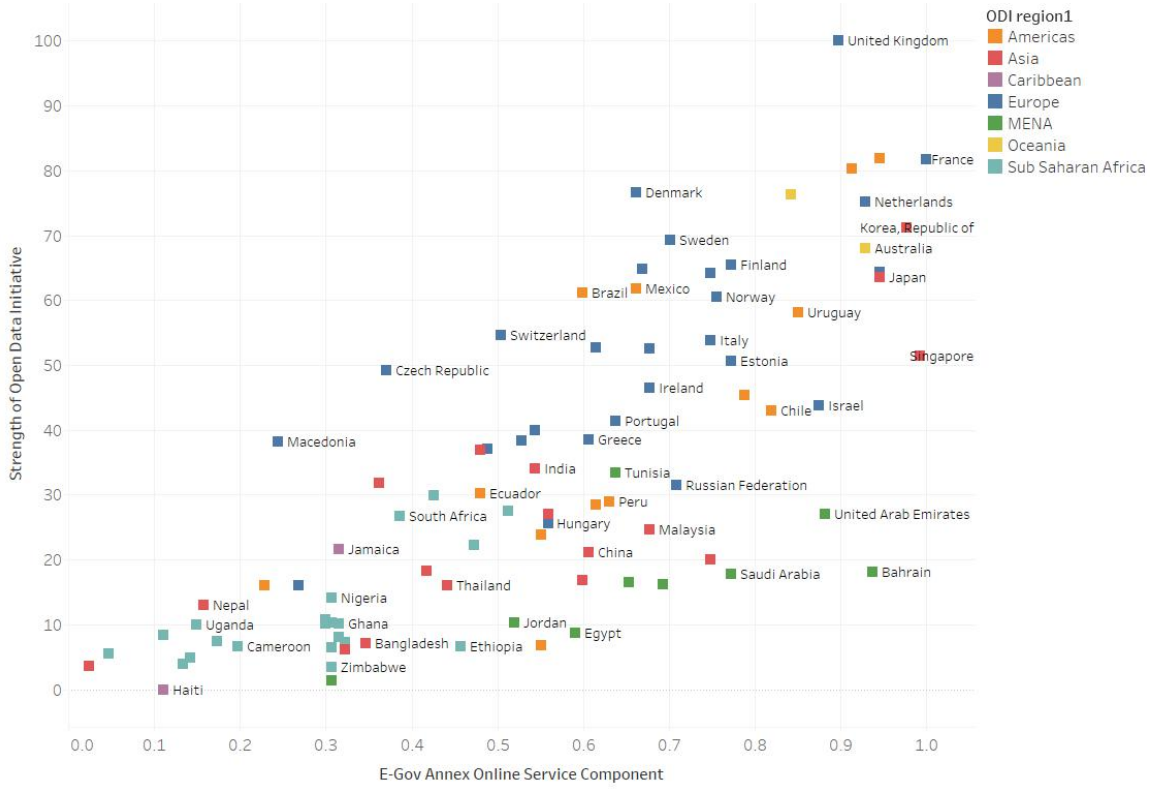


Figure 21: Online Service Component Scatterplot

The most interesting finding in the regional tests of association of the online service index and the strength of open data initiatives is the frequency with which Sub Saharan Africa appeared with a positive relationship of its country overall score variable (0.564, $p = 0.008$), as well as seven of its dataset types also exhibiting significance levels ≤ 0.05. This finding corroborates what has been a focus of many in the open data community: open data inequality. Open data inequality refers to the fact that nearly half of the open datasets used in ODB's global

report came from OECD countries, and "almost none in African countries."[207] Open data inequality is one of the reasons to look at what factors shape a country's path towards its open data initiative. A common question that arises in the topic of open data inequality is, "Why is it happening? What can we do to level the open data playing field?" Open data experts acknowledge one of the largest barriers to successful open data initiatives in Sub Saharan Africa is technology familiarity[208] in programs like Excel for government officials responsible for the cleaning and upload of data.

Analysis also indicates a weaker relationship between the Online Service Index and the strength of open data initiatives across the 11 types of datasets. Europe's only dataset types that exhibit a statistically significant relationship with the online services index are D5 - budget, D10 - Trade, and D14 - Environment. Comparatively, Sub Saharan Africa exhibits statistically significant relationships with the online services index in D1 – Mapping, D5- Budget, D9 – Transport, D12, Education, D14 – Environment, D15 – Elections, and D16 – Contracts data. The number of dataset types that show positive relationships with the online services index suggest that across the board, improving government knowledge of web technologies would improve the quality of data being published from multiple Ministries. Conversely, improving knowledge and access to web technology building tools would not have as large of an impact in the quality of data Europe would publish.

[207] *World Wide Web Foundation*, "Open Data Barometer: Data Poverty the Next Frontier Widening Inequality," April 20, 2016, http://webfoundation.org/2016/04/open-data-barometer/.
[208] Valentina Romei, *FT Data,* "Experts Raise Alarm Over Rise in "Open Data Inequality," April 21, 2016.

5.3 Education Attainment and Technology Access and Adoption

Analysis of countries with open data initiatives globally indicates that, like its subcomponent, the EGDI, the Telecommunications Infrastructure subcomponent, and the Human Capital subcomponent, all exhibit highly significant relationships with the strength of an open data initiative. While previous tests such as the start date of FOI, democratic freedoms and corruption, or the involvement of the President in open data initiatives examined factors influencing the path of an open data initiative through political and societal lenses, education attainment and technology access and adoption measures the capability of countries to create successful open data initiatives. Analysis suggests that, within countries with open data initiatives, as the telecommunications capacity and levels of education increase, the strength of an open data initiatives rises as well.

Regionally, Sub Saharan Africa again has several of its dataset kinds exhibit statistically significant associations with EGDI and its subcomponents. This further supports those in the open data community who have been examining the technology barrier for successful open data initiatives in Sub Saharan Africa.[209]

Unlike the Online Service Index, Europe appears with ten of the eleven dataset types holding strong associations to EGDI and telecommunications infrastructure scores. It is possible that as the Online Service index solely focuses on government websites, the variable did not capture the full breadth of the variance of Europe's technology infrastructure between nations, which the EGDI and telecommunications infrastructure indices focus on. Across all regions, the

[209] "Data for Africa, by Africa: Rebooting Open Data and the High Level Conference on the Data Revolution," *World Wide Web Foundation,* April 2, 2015, http://webfoundation.org/2015/04/data-for-africa-by-africa-rebooting-open-data-and-the-high-level-conference-on-the-data-revolution/.
"East Africa's Open Data Revolution," *Global Partnership For Sustainable Development Data,* September 8, 2016, http://www.data4sdgs.org/master-blog/2016/9/7/east-africas-open-data-revolution.

EGDI index showed more statistically significant associations with the different kinds of datasets (24 associations with p values ≤ 0.05) than any other test in this paper; the next highest number of associations is the CPI, with 23 associations to the 11 kinds of datasets with p values ≤ 0.05. While these summed numbers are not particularly useful as they re-reduce regional analyses into a less detailed presentation of the data, it is worth noting that these two variables show association across multiple types of datasets.

Various ministries in the government are responsible for creating and uploading these datasets. If the strength of these datasets is a proxy for evaluating the willingness and technical familiarity and ability of these ministries to provide specific datasets relevant to their sector, the high number of dataset types that are significantly associated with telecommunications infrastructure suggests that reducing improving telecommunications infrastructure could improve an open data initiative across multiple government ministries/departments.

Analysis of regional groupings indicated that Asia (0.713, $p = 0.001$) and Europe (0.554, $p = 0.003$) both exhibit statistically significant associations between the overall country scores and EGDI human capital index. However, while Europe exhibited association with a p value of only ≤ 0.05 with the budget dataset score, Asia exhibited statistically significant association with its transport, health, crime, and environment datasets, the latter three having p values ≤ 0.01. Although MENA and Sub Saharan Africa did not exhibit association between human capital and the overall country open data initiative score, MENA's census dataset score (0.694, $p = 0.038$) and the crime (0.560, $p = 0.008$) and elections (0.639, $p = 0.002$) datasets in Sub Saharan Africa all exhibited statistically significant relationships. The strength and positive directionality of the associations in Sub Saharan Africa suggest that as education levels increase, the quality and availability of crime statistics data and national elections results data both increase. Unlike other

kinds of datasets, such as public transportation routes or education datasets, both crime and national elections results data are strongly related to the accountability and transparency of a government. These associations between the level of education in a country and its relationship with the strength of specific kinds of datasets would benefit greatly from in depth qualitative research.

5.4 Involvement of International Organizations

Although NGOs and IGOs operate at the citizen, local, and national level with open data initiatives, there is a lack of research on the actual influence NGOs/IGOs have on the path of open data portals. Analysis indicates a strong positive association between the number of open data programs sponsored by NGOs that governments are involved in and the strength of their open data initiatives, as well as a strong positive association between the number of open data programs and networks citizens and local NGOs are involved in and the strength of their country's open data initiative.

When grouping results regionally, the Americas is the only region that exhibits a statistically significant relationship between the total count of open data NGO programs operating within a country and the overall score of that country's open data initiative. Asia and Sub Saharan African exhibit positive associations with the number of international NGOs citizens and local NGOs are involved in and the strength of the national open data portal. These findings run alongside current concerns in the open data community of self-sustainability of open data initiatives.[210] Without external pressure from multilateral institutions, some open data initiatives that have been spearheaded by NGOs in countries like Sierra Leone may not receive

[210] *Global Partnership for Sustainable Development Data*, "East Africa's Open Data Revolution," September 8, 2016.

the support they require from government ministries to produce updated, machine-readable data.[211] Nevertheless, the fact that the variable measuring the number of open data NGOs individuals and local NGOs are a part of, while the variable measuring the total number of NGO projects a national government is a part of, hints at the citizen-driven undercurrent in open data initiatives in Sub Saharan African and Asia.

5.6 Limitations

While this paper examines open data from four angles: cultures of transparency and accountability, government involvement and support for technology, education attainment and technology access and adoption, and the involvement of international organizations promoting open data initiatives, there are other fields of study in open data that I did not address. Several of these fields examine the threats open data could potentially pose given the massive amounts of data collection that occurs continuously and often surreptitiously and without individual approval. With the government able to access a great deal of the privately-collected data collected on any given individual, concerns over deanonymized data are highly relevant to future public perceptions of open data.

Possibly one of the largest challenges of this paper was assembling and working entirely with secondary data. As the indices I employed are from a variety of sources, each with a parent organization with its own agenda, I was cautious in the choice of statistical tests I used, as well as was cautious in overstating findings. I was also wary of speculating on why a certain variable shows a significant association with the strength of open data initiatives, whether the test was on a global or regional level. Proper assessments of what is happening on such a detailed level

[211] Ibid.

require extensive regional expertise as well as in depth qualitative studies. Instead, I focused on evaluating the larger trends in my analysis, as well as describing the implications taking a less conventional methodological approach in assessing open data globally.

5.5.1 Key Findings and Their Implications

While current research and reports on open data look forward, evaluating impact and potential benefits of open data adoption, there is a little written on the factors that influence the path of a national government in creating an open data initiative. Literature that does address factors influencing open data adoption usually does so using qualitative methods and expert interviews.[212]

Key findings from analyses of cultures of transparency and accountability tests indicate that: (1) When examining all countries with open data initiatives, the year FOI legislation was introduced, democratic freedoms, and levels of perceived corruption all show strong relationships with the strength of an open data initiative. (2) However, when tested by region, no region exhibited a statistically significant relationship with the year FOI legislation was introduced and a country's overall open data initiative score. (3) Additionally, when tested by region, Asia and Europe exhibit strong positive associations between democratic freedoms and the strength of their open data initiatives, (4) and when testing against perceived corruption, the Americas and Europe exhibit strong positive associations between lower levels of corruption and stronger open data portals. These findings support existing academic and NGO literature which include FOI legislation, democratic freedoms, and transparency and corruption as important

[212] M. Janssen, Y. Charalabidis & A. Zuiderwijk, "Benefits, Adoption Barriers and Myths of Open Data and Open Government. Information Systems Management (ISM)," 2012, vol. 29, no.4, pp. 258.
Noor Huijboom, & Tijs Van den Broek, "Open Data: An International Comparison of Strategies," *European Journal of ePractice,* April 2011, pg. 3.

factors influencing open data initiatives globally.[213] At the same time, analyses suggest caution in assuming these factors are equally influential across all countries and regions. While the reason a region like MENA does not exhibit a relationship to democratic freedoms may be due to countries' similar open data initiative scores across all MENA countries, the varying strengths of association depending on region are relevant figures to report alongside the global level of association in understanding the influence of cultures of transparency and accountability on open data initiatives.

Within government involvement and support for technology, key findings include (1) While the involvement of the President/Prime Minister in the development of an OGP action plan does not exhibit association with the strength of the open data initiative in question, involving the agency that will implement the action plan in the development of the plan exhibits a positive association with the strength of the open data initiative. This finding indicates that against all the media noise surrounding the development of open data initiatives, examining the specific agency that will implement the OGP action plan is a stronger predictor of the outcome of an initiative than the involvement of the president/prime minister or MOFA. (2) Sub Saharan Africa and the Americas exhibit strong associations between the Open Service index and the strength of open data initiatives. In comparison to the other three topic areas tested, Sub Saharan Africa shows significant association not only in its overall country score but also with several types of datasets. This finding specifically pertains to government familiarity with web technologies and its ability to create government websites with easy-to-use tools and features, and supports existing literature that discusses government familiarity with various technologies

[213] Schwegmann, Claudia, "Open Data in Developing Countries," *eSPI*, February, 2012, https://www.europeandataportal.eu/sites/default/files/2013_open_data_in_developing_countries.pdf. "ODB Methodology – v1.0 28th April 2015," *Open Data Barometer*, April 28th, 2015.

associated with open data (e.g. spreadsheet software, website development) as potential barriers to creating robust open data initiatives.[214]

Analyses of education attainment and technology access and adoption indicate that across all regions, EGDI exhibits a strong relationship with the strength of open data initiatives, both in their overall open data initiative country scores as well as specific dataset types. Across all four topics and tests, EGDI exhibited the highest number of statistically significant associations with dataset types. This measure of total number of dataset types exhibiting significant associations with a given independent variable is a method of analysis to indirectly examine the influence of the independent variable across government ministries who are responsible for creating and providing access to these specific datasets. Analysis of EGDI's relationship with dataset types, in comparison to all other tests in this paper, suggests that if improved upon, higher capacities for e-government development could have the most widespread impact improving the strength of a country's open data initiative.

Analyses of the involvement of international organizations indicate that Asia and Sub Saharan Africa exhibit statistically significant associations between citizen/local NGO participation in international NGO open data initiatives and the strength of the national open data initiative. Although citizen-driven open data programs are lauded within the open data community,[215] further qualitative research is necessary to understand why Asia and Sub Saharan Africa exhibit this association as well as what sociopolitical and technology environments promote this association.

---

[214] Audrey Ariss, "Machine-readable open data: How It's Applicable to Developing Countries," *The DATA Blog*, March 02, 2017, https://blogs.worldbank.org/opendata/machine-readable-open-data-how-it-s-applicable-developing-countries.

[215] Erin Cusack, "Citizen-Driven Data is Key to the Success of Nepal's Open Data Revolution, *AidData*, August 13, 2014, http://aiddata.org/blog/citizen-driven-data-is-key-to-the-success-of-nepals-open-data-revolution.
Niklas Kossow, "Albania: Civil Society Going at it Alone?" *Open Data in Europe and Central Asia,"* October 19, 2015, http://aiddata.org/blog/citizen-driven-data-is-key-to-the-success-of-nepals-open-data-revolution.

As the number of countries with open data initiatives continues to increase, those engaged in designing and evaluating the initiatives will continue to develop innovative methods of synthesizing and analyzing data from these countries in order to match the growing volume of local, state, and national open data initiatives. While past methods of global analysis that leverage descriptive statistics will have benefits in their comprehensibility and ease-of-testing, the robustness of the raw datasets collected by GODI and ODB in assessing open data worldwide have reached a point where new interrogative methods of analysis, examining both the factors influencing open data initiatives as well as open data initiatives' impacts, are possible. Additionally, while the findings of this paper begin to explore the expansion of the motivating factors behind open data initiatives as a new wave of countries invest in open data initiatives, the methods of analysis used in this paper must be complimented by in depth qualitative research on national and regional levels in order to better understand this expansion.

Based within an existing wealth of qualitative literature from NGOs, IGOs, and academia, the tests of association conducted in this paper between overall scores and specific dataset scores on a regional level can (1) better inform researchers on the numerical landscape of open data on a regional level (2) provide a comparative analysis of different regional paths toward open data that can serve as a stepping stone for future in-depth qualitative research (3) offer a new measurement of assessing how various factors affect open data initiatives across ministries, therefore providing open data advocates with more precise routes of engagement that will maximize impact when creating an open data project.

Appendix

Abbreviations and Acronyms

API	Application Program Interface
FOI	Freedom of Information
ICT	Information and Communication Technology
IGO	Intergovernmental Organization
IoT	Internet of Things
NGO	Non-government Organization
ODH	Open Data Handbook
OGI	Open Government Information
RTI	Right to Information

CLD	Centre for Law and Democracy
CTG	Center for Technology in Government
EGDI	E-Government Readiness Index
EITI	Extractive Industries Transparency Initiative
EU	European Union
GODI	Global Open Data Index
NYPD	New York Police Department
OD4D	Open Data for Development
ODB	Open Data Barometer
OECD	Organization for Economic Co-operation and Development
OGP	Open Government Partnership
OKI	Open Knowledge International

Additional Tables

OGP Member Data	0	1	Mann-Whitney U	Z	Asymp. Sig. (2-tailed)	Exact Sig. [2*(1-tailed Sig.)]
President/Prime Minister	17	9	66.000	-.566	0.571	.597
Change in Exec	16	10	75.000	-0.264	0.792	.816
Exec Lead Agency	14	7	47.000	-0.149	0.881	.913
Executive	1	25	3.000	-1.268	0.205	.308
Implementing Agency	12	14	39.000	-2.316	0.021	.020*
Legally Mandated	22	4	18.000	-1.849	0.064	.069
Official Mandate	5	21	40.000	-0.814	0.416	.447
Ministry for Affair	14	12	76.000	-0.412	0.681	.705
MOFA Lead Agency	7	1	0.000	-1.528	0.127	.250
Multiple Agencies	8	18	65.000	-0.389	0.697	.724
Multiple Arrangements	16	10	66.000	-0.738	0.460	.484
Only One Branch	4	22	40.000	-0.284	0.776	.811

39 Countries
** = p≤0.01, * = p≤0.05
Table 39:Full Table of Table 19 (pg. 94)

1	The government is a member/participant/adopter of an NGO's program
2	Individual citizens/local NGOs within a given country are members of an international NGO
3	A country has an open data project occurring in their country, but the government is not part of it
4	Both local NGOs and the government are members of an international NGO/multilateral open data initiative

Table 40: Organization Classification and Breakdown of NGO Variable

OKFN Open Knowledge Foundation Membership	2
OpenDRI Open Data for Resilience Initiative Membership	3
GPSA Global Partnership for Social Accountability	1
IBP Open Budget Index	3
OGP	1
GIFT Global Initiative for Fiscal Transparency	4
OCP Open Contracting Partnership	1
Financial Transparency Coalition	2
BOOST	1
Open Aid Partnership	1
Open Aid Partnership	1
African Development Bank Open Data	3
EITI	1
Public Administration and Development Management Department of Economic and Social Affairs	1
World Bank Open Government Data Community of Expert Consultants	2
Open Data Charter	1
Omidyar	2
Center for Open Data Enterprise	2

Table 41: Organizations Included in NGO Variable

Bibliography

"114 FOI Regimes." *Freedom Info.* April 5, 2016. http://www.freedominfo.org/2016/04/105-foi-regimes-freedominfo-org-count-shows/.

"A World That Counts: Mobilizing the Data Revolution for Sustainable Development." *UN Independent Expert Advisory group on a Data Revolution for Sustainable Development (IEAG).* 2015. http://www.undatarevolution.org/report/.

"About." *Michael Culture Association. http://www.michael-culture.eu/about.*

"Advancing Access to Information in MNA: Supporting Coalitions & Networks." *World Bank Institute and Middle East and North Africa Social Development Department.* "May, 2012. http://documents.worldbank.org/curated/en/945061468299331196/pdf/NonAsciiFileName0.pdf.

Ackerman, John M, Sandoval-Ballesteros, Irma E. "The Global Explosion of Freedom of Information Laws." *Administrative Law Review.* vol. 58, no. 1. 2006. http://www.jstor.org/stable/40712005.

Adler, Steven. "Why the Open Government Partnership Needs a Reboot." *Global Investigative Journalism*. December 12, 2015. http://gijn.org/2015/12/12/why-the-open-government-partnership-needs-a-reboot/.

Ariss, Audrey. "Machine-readable open data: How It's Applicable to Developing Countries." *The DATA Blog.* March 02, 2017. https://blogs.worldbank.org/opendata/machine-readable-open-data-how-it-s-applicable-developing-countries.

Ben Yacoub, Hatem. "Why OKF Global Open Data Index 2015 is a Failure." December 10, 2015. http://www.hbyconsultancy.com/blog/why-okf-global-open-data-index-2015-is-a-failure.html.

Barker, Sara. "Govt Calls Forth Kiwi Developers & Data Scientists for Open Data Fellowships." *FutureFiveNZ.* August 5, 2016. https://futurefive.co.nz/story/govt-calls-forth-kiwi-developers-data-scientists-open-data-fellowships/.

Bonina, Carla M. “New Business Models and the Value of Open Data: Definitions, Challenges, and Opportunities.” *London School of Economics and Political Science.* September 27. 2013. http://www.nemode.ac.uk/wp-content/uploads/2013/11/Bonina-Opendata-Report-FINAL.pdf.

Bovens, Mark. “Information Rights: Citizenship in the Information Society.” *The Journal of Political Philosophy.* http://dx.doi.org/10.1111/1467-9760.00155.

Bozeman, Barry. *Public Values and Public Interest: Counterbalancing Economic Individualism.* Georgetown University Press. 2007.

“Briefing Paper on Open Data and Privacy.” *The Center for Open Data Enterprise.* 2016. http://reports.opendataenterprise.org/BriefingPaperonOpenDataandPrivacy.pdf.

Canes-Wrone, Brandice. *Who Leads Whom? Presidents, Policy, and the Public*. The University of Chicago Press. 2006.

Castells, Manuel. *The Rise of the Network Society: The Information Age: Economy, Society, and Culture Vol. 1.* Blackwell. 1996.

Clark, Cal, Veal, Don-Terry. *Advancing Excellence and Public Trust in Government.* Lexington Books. 2011.

Cohen, Jeffrey E. *Presidential Leadership in Public Opinion.* Cambridge University Press. 2015.

“Corruptions Perceptions Index 2010: Long Methodological Brief.” *Transparency International.* http://transparency.ee/cm/files/cpi2010_long_methodology_en.pdf.

“Corruptions Perceptions Index 2015.” Transparency International. http://www.transparency.org/news/feature/corruption_perceptions_index_2016.

Costa, Stefano. “Italian Government Ditches Transparency and Open Data.” *Open Knowledge International.* March 1, 2013. http://blog.okfn.org/2013/03/01/italian-government-ditches-transparency-and-open-data/.

Cume, Ian, Klaaren, Jonathan. “An Update on Access to Information in South Africa: New Directions in Transparency.” Freedom of Information Review. 2003.

Cusack, Erin. “Citizen-Driven Data is Key to the Success of Nepal’s Open Data Revolution.” *AidData.* August 13, 2014. http://aiddata.org/blog/citizen-driven-data-is-key-to-the-success-of-nepals-open-data-revolution.

Dany, Charlotte. *Global Governance and NGO Participation: Shaping the Information Society in the United Nations.* Routledge (2013).

"Data for Africa, by Africa: Rebooting Open Data and the High Level Conference on the Data Revolution." *World Wide Web Foundation.* April 2, 2015. http://webfoundation.org/2015/04/data-for-africa-by-africa-rebooting-open-data-and-the-high-level-conference-on-the-data-revolution/.

Data.world, "Understanding data.world's features." https://data.world/features-overview.

"East Africa's Open Data Revolution." *Global Partnership For Sustainable Development Data.* September 8, 2016. http://www.data4sdgs.org/master-blog/2016/9/7/east-africas-open-data-revolution.

E-Government Readiness Index," *Wikiprogress,* "http://wikiprogress.org/articles/governance-human-rights/e-government-readiness-index/#Online_Service_Index

"Executive Order – Making Open and Machine Readable the New Default for Government Information." Office of the Press Secretary. *The White House*. May 09, 2013. https://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-.

Fagan, Craig. *Transparency International Blog.* "Why Open Data Can Stop Corruption." March 24, 2016. http://blog.transparency.org/2016/03/24/why-open-data-can-stop-corruption/.

*"*Fighting Corruption in Eastern Europe and Central Asia: The Istanbul Anti-Corruption Action Plan." OECD (2008).

"Finland." *Open Data Partnership.* http://www.opengovpartnership.org/country/finland.

"France Becomes 64th Country to Join the Open Government Partnership." *Open Government Partnership.* April 24, 2014. http://www.opengovpartnership.org/blog/blog-editor/2014/04/24/france-becomes-64th-country-join-open-government-partnership-press.

Freedom House. "FIW 2016 Data." https://freedomhouse.org/report/freedom-world-2016/download-fiw-2016-data.

"FY06 LICUS List (Low-Income Countries Under Stress). *World Bank.* http://pubdocs.worldbank.org/en/359521410886172040/FY6toFY9Fragile-States-List-formerly-LICUS.pdf.

Gao, George. "What Americans Think About NSA Surveillance, National Security and Privacy." *Pew Research Center*. May 29, 2015. http://www.pewresearch.org/fact-tank/2015/05/29/what-americans-think-about-nsa-surveillance-national-security-and-privacy/.

Galka, Max. "How an Open Data Blogger Proved the NYPD Issued Parking Tickets in Error." *The Guardian.* July 26, 2016. https://www.theguardian.com/cities/2016/jul/26/open-data-blogger-parking-tickets-new-york-nypd.

"Global Report." *Open Data Barometer.*2015. http://opendatabarometer.org/doc/3rdEdition/ODB-3rdEdition-GlobalReport.pdf.

"Global Right to Information Rating." Centre for Law and Democracy. http://www.rti-rating.org/country-data/.

Gordon, Donald. *Transparent Government: What It Means and How You Can Make It Happen.* Prometheus Books. 2014.

Goutham, Sudheer. "KTR's Mission Hyderabad as India's Silicon Valley." *eGov.* October 04, 2016. http://egov.eletsonline.com/2016/10/ktrs-mission-hyderabad-as-indias-silicon-valley/.

Granickas, Karolis. "Open City: Local Government & Open Data." *European Data Portal/* August, 2015. https://www.europeandataportal.eu/sites/default/files/2015_open_city_local_government_and_open_data.pdf.

Gurin, Joel, Manley, Laura. "Open Data Transition Report: An Action Plan for the Next Administration." *Center for Open Data Enterprise.* October 2016. http://opendataenterprise.org/transition-report.

Gurstein, Michael. "Open Data: Empowering the Empowered or Effective Data Use For Everyone?" *First Monday.* February 7, 2011. http://firstmonday.org/ojs/index.php/fm/article/view/3316/2764#p3.

"Haiti." *Open Data Watch*. http://odin.opendatawatch.com/Report/countryProfile/HTI?appConfigId=1.

"Highlighting Use Cases From the Open Data Impact Map: Farmerline Empowering Farmers Through Mobile Technology and Information." *Center for Open Data Enterprise.* September 23, 2016. http://opendatacon.org/highlighting-use-cases-from-the-open-data-impact-map-farmerline-empowering-farmers-through-mobile-technology-and-information/.

"Highlighting Use Cases From the Open Data Impact Map: Haezoom Accelerating the Adoption of Solar Energy in South Korea." *Center for Open Data Enterprise.* September 28, 2016. http://opendatacon.org/highlighting-use-cases-from-the-open-data-impact-map-farmerline-empowering-farmers-through-mobile-technology-and-information/.

Hoehner, Christian. "The OPEN Government Data Act: A Sweeping Open Data Mandate for All Federal Information." *Data Coalition*. April 18, 2016. http://www.datacoalition.org/the-open-government-data-act-a-sweeping-open-data-mandate-for-all-federal-information/.

Huijboom, Noor, Van den Broek, Tijs. "Open Data: An International Comparison of Strategies." *European Journal of ePractice*. April 2011. http://www.tandfonline.com/doi/full/10.1080/10580530.2012.716740.

"Intergovernmental Organizations (IGOs)." *Harvard Law School.* http://hls.harvard.edu/dept/opia/what-is-public-interest-law/public-international-law/intergovernmental-organizations-igos/.

"International Open Data Roadmap." *OD4D,* 2016. http://od4d.com/roadmap/assets/files/report-iodc-2016-web.pdf.

Janssen, M, Charalabidis Y, Zuiderwijk A. "Benefits, Adoption Barriers and Myths of Open Data and Open Government." *Information Systems Management (ISM).* 2012. vol. 29, no.4. http://www.tandfonline.com/doi/full/10.1080/10580530.2012.716740.

Jennings, Ralph. "How Taiwan Fostered the World's Most Open Government." *Forbes.* December 15, 2015. http://www.forbes.com/sites/ralphjennings/2015/12/15/how-taiwan-rose-to-world-no-1-in-government-transparency/#54d5020a7179.

Kash, Wyatt. "Number of Countries Embracing Open Government Surges." *FedScoop*. August 1, 2016. http://fedscoop.com/number-of-countries-embracing-open-government-surges.

Karklins, Rasma. *The System Made Me Do It: Corruption in Post-Communist Societies.* M.E. Sharpe (2005).

Knell, Noelle, et al., "Open Data Policies in State and Local Government." *GovTech.* March 17, 2014. http://www.govtech.com/data/are-governments-committed-to-open-data-interactive-map.html.

Kossow, Niklas. "Albania: Civil Society Going at it Alone?" *Open Data in Europe and Central Asia.* October 19, 2015. http://aiddata.org/blog/citizen-driven-data-is-key-to-the-success-of-nepals-open-data-revolution.

Ladd, Bruce. *Crisis in Credibility*. The New American Library, Inc. 1968.

Ma, Jinxin Yolanda. *"Why Open Data is Good For China." Global Investigative Journalism.* March 28, 2017. http://gijn.org/2017/03/28/why-open-data-is-good-for-china/.

Mamadou Dia. *World Bank, A Governance Approach to Civil Service Reform in Sub-Saharan Africa.* International Bank for Reconstruction. (1993).

Manyika, James, et al. "Open Data: Unlocking Innovation and Performance with Liquid Information." *Mckinsey Global Institute.* October 2013. http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information.

Mehboob, Ahmed. "The Right to Information." February 20, 2016. *Dawn*. http://www.dawn.com/news/1240632.

"Methodology." *Open Data Index.* 2016. http://index.okfn.org/methodology/.

Mohanty, Hrushikesha, Bhuyan, Prachet, Cehnthati. Deepak. *Big Data: A Primer*. Springer. 2015.

"New Country Classifications by Income Level." World Bank Data Team. July 1, 2016. https://blogs.worldbank.org/opendata/new-country-classifications-2016.

"New Initiative Aimed at Increasing Trade and Investments in Armenia." *International Finance Corporation.* http://www.ifc.org/wps/wcm/connect/news_ext_content/ifc_external_corporate_site/news+and+events/news/new+initiative+aimed+at+increasing+trade+and+investments+in+armenia.

"NGO." *Oxford Dictionary.* https://en.oxforddictionaries.com/definition/NGO.

"Norway." *Open Data Partnership.* http://www.opengovpartnership.org/country/norway.

"ODB-3rdEdition-Datasets-Scored.csv." *Open Data Barometer 3rd Edition.* http://opendatabarometer.org/3rdEdition/data/.

O'Donnell, Guillermo. "Delegative Democracy?" *The Helen Kellogg Institute for International Studies.* April 1993. https://kellogg.nd.edu/publications/workingpapers/WPS/172.pdf.

OpenDataSoft. "A Comprehensive List of 2600+ Open Data Portals Around the World." https://www.opendatasoft.com/a-comprehensive-list-of-all-open-data-portals-around-the-world/.

"Open Budgets Portal: Tools and Resources." *The World Bank.* http://wbi.worldbank.org/boost/tools-resources/intl-initiatives.

"Open Budget Survey." *International Budget Partnership.* http://www.internationalbudget.org/opening-budgets/open-budget-initiative/open-budget-survey/.

"Open Data." *Council of the European Union. http://www.consilium.europa.eu/en/general-secretariat/corporate-policies/transparency/open-data/.*

"Open Data for Business in Jamaica: Initial Findings and Recommendations." *The Caribbean Open Institute.* http://caribbeanopeninstitute.org/od4b_roundtable_jamaica.

"Open Data for Sustainable Development." *World Bank Group.* August 2015, 8. http://pubdocs.worldbank.org/en/999161440616941994/Open-Data-for-Sustainable-Development.pdf.

"Open Data Impact Map." Center for Open Data Enterprise. 2016. http://www.opendataimpactmap.org/map.html.

"Open Data Portals." *European Commission.* https://ec.europa.eu/digital-single-market/en/open-data-portals.

"Open Data Readiness Assessment." *St Lucia Open Data Portal.* http://data.govt.lc/story/open-data-readiness-assessment.

"Open Data Roundtables." *Data.gov*. March 4, 2016. https://www.data.gov/meta/open-data-roundtables/.

"Open Data: the Extractive Industries Case-Study." *Sherpa.* December 8, 2016. https://www.asso-sherpa.org/open-data-the-extractive-industries-case-study.

"Open Data, Transparency and Accountability: Topic Guide." *GSDRC.* September. 2016. http://www.gsdrc.org/wp-content/uploads/2016/09/OpenDataTA_GSDRC.pdf.

"Open Development – Vietnam." *Open Development Initiative*. https://vietnam.opendevelopmentmekong.net/.

Open Government Partnership. "Public Access_IRM SU Commitments Database March 2017." https://docs.google.com/spreadsheets/d/1MEF5zN41_4tW1pK5Ptp0vqK0HvA0ofwMN9XaXK-QVNU/edit?usp=sharing.

Jacob Perez. "Open Data in Mexico: The Search for Civic, Private, and Public Sector Innovators." *Medium*. May 18, 2016. https://medium.com/@canada2020/open-data-in-mexico-the-search-for-civic-private-and-public-sector-innovators-a974547d1c22.

Piotrowski, Suzanne J, et al. "Key Issues For Implementation of Chinese Open Government Information Regulations." *Public Administration Review.* vol. 69. December 2009. http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6210.2009.02100.x/full.

"Place Overview." *Open Data Index.* http://index.okfn.org/place/.

Pollack, Rufus. "Open Knowledge Foundation Launched." *Open Knowledge International Blog.* May 24, 2004. https://blog.okfn.org/2004/05/24/open-knowledge-foundation-launched//.

"President Xi Stresses Innovation to Bolster Economy." *Xinhua Net.* May 10, 2016. http://news.xinhuanet.com/english/2016-05/10/c_135348451.htm.

Qiu, Linda. "Does Vladimir Putin Kill Journalists?" *PunditFact.* January 4, 2016. http://www.politifact.com/punditfact/article/2016/jan/04/does-vladimir-putin-kill-journalists/

Ramanauskaite, Egle Marija. "Technarium Hackerspace Celebrates Open Data Day in Vilnius, Lithuania." *Open Knowledge Foundation Blog.* April 27, 2017, https://blog.okfn.org/2017/04/27/technarium-hackerspace-celebrates-open-data-day-in-vilnius-lithuania/.

Ravallion, Martin. "Economy-wide Policies." Lecture. Georgetown University Economics Department. Washington D.C. November 17, 2016.

Relly, Jeannine E, Sabharwal, Meghna. "Perceptions of Transparency of Government Policymaking: A Cross-national Study." *Government Information Quarterly.* vol. 26. January 2009. http://dx.doi.org/10.1016/j.giq.2008.04.002.

"Right to Know." *The Economist.* May 3, 2014. http://www.economist.com/news/china/21601564-leaders-discover-some-transparency-can-help-make-society-more-stable-right-know.

Roberts, Alasdair. "Access to Information: A Key to Democracy." *The Carter Center.* November 2002. https://www.cartercenter.org/documents/1272.pdf.

Romei, Valentina. *FT Data.* "Experts Raise Alarm Over Rise in "Open Data Inequality." April 21, 2016. http://blogs.ft.com/ftdata/2016/04/21/experts-raise-alarm-over-rise-in-open-data-inequality/.

"RTI Rating Data Analysis Series: Overview of Results and Trends," Centre for Law and Democracy, September 28, 2013, http://www.law-democracy.org/live/wp-content/uploads/2013/09/Report-1.13.09.Overview-of-RTI-Rating.pdf.

"SAP Advances Effort to Help Public Sector Organizations Measure Political, Social and Operational Value of IT Projects." *SAP*. September 13, 2006. http://global.sap.com/news-reader/index.epx?pressid=6714.

Sendugwa, Gilbert Ronald. "Legislating and Implementing Public Access to Information in Africa: What are the Incentives for Government and Civil Society Actors?" *Africa Freedom of Information Centre (AFIC).* 2010. https://spaa.newark.rutgers.edu/sites/default/files/files/Transparency_Research_Conference/Papers/Sendugwa_Gilbert.pdf.

Schwartz, Molly. "Democracy and Open Data: Are the Two Linked?" May 22, 2014. http://congressionaldata.org/democracy-and-open-data-are-the-two-linked/.

"Sweden." *University College London.* https://www.ucl.ac.uk/constitution-unit/research/foi/countries/sweden.

Tauberer, Joshua. "History of the Movement." *Open Government Data: The Book (2nd Edition).* *Lulu Press, Inc.* October 6, 2014. https://opengovdata.io/2014/history-the-movement/.

"Technical Assistance and Funding," *The World Bank,* http://opendatatoolkit.worldbank.org/en/technical-assistance.html.

Thomas, Rob. McSharry, Patrick. *Big Data Revolution: What Farmers, Doctors, and Insurance Agents Teach Us About Discovering Big Data Patterns.* John Wiley & Sons Ltd. 2015.

"Turkey Made Inactive in the Open Government Partnership." *Open Government Partnership.* September 21, 2016. http://www.opengovpartnership.org/blog/open-government-partnership/2016/09/21/turkey-made-inactive-open-government-partnership.

"UN E-Government Survey 2014." United Nations. https://publicadministration.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2014.

"United Nations e-Government Survey 2010." *United Nations Public Administration Network.* http://unpan1.un.org/intradoc/groups/public/documents/un/unpan038851.pdf

Voltmer, Katrin. "Chapter 6: The Media, Government Accountability, and Citizen Engagement." *Public Sentinel: News Media & Governance Reform*. The World Bank. November 2009.

Weinbull, Lennart. "Freedom of the Press Act of 1766." *Encyclopaedia Britannica.* December 21, 2015. https://www.britannica.com/topic/Freedom-of-the-Press-Act-of-1766.

"What is Open Data?" *Open Data Handbook*. 2016. http://opendatahandbook.org/guide/en/what-is-open-data/.

White, Nicola. *Free and Frank – Making the Official Information Act 1982 Work Better.* Institute of Policy Studies (Victoria University). November 2007.

"Why Use GNI per Capita to Classify Economies into Income Groupings?" World Bank. https://datahelpdesk.worldbank.org/knowledgebase/articles/378831-why-use-gni-per-capita-to-classify-economies-into.

"World Bank Open Government Data Community of Expert Consultants." *World Bank*. August 12, 2016. https://docs.google.com/spreadsheets/d/1JN1BeXagQNLnyRs3YRcIdpDbmdIH9yeYtm1hXQZm400/edit#gid=0.

"World Data Centers." *NOAA*. https://www.ncdc.noaa.gov/customer-support/world-data-centers.

Worthy, Ben. "More Open but Not More Trusted? The Effect of the Freedom of Information Act 2000 on the United Kingdom Central Government." *Governance*. vol. 23, issue 4. September 2010. http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0491.2010.01498.x/full.