

# OPEN DATA IMPACT

## WHEN DEMAND AND SUPPLY MEET

Key Findings of the Open Data Impact Case Studies



By Stefaan Verhulst and Andrew Young\*

*March 2016*

# OPEN DATA IMPACT

## WHEN DEMAND AND SUPPLY MEET

Key Findings of the Open Data Impact Case Studies

By Stefaan Verhulst and Andrew Young\*

*March 2016*

[www.odimpact.org](http://www.odimpact.org)

*\* With important conceptual, research and editorial input from Ali Clare, Auralice Graft, Akash Kapur, Juliet McMurren, Beth Simone Noveck, Christina Rogawski and David Sangokoya.*



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

# TABLE OF CONTENTS

Executive Summary .....	4
Introduction .....	6
<b>I. What Is Open Data? .....</b>	<b>7</b>
<b>II. The Case Studies.....</b>	<b>8</b>
<b>III. What Is the Impact of Open Data on People’s Lives?.....</b>	<b>14</b>
<b>IV. What Are the Enabling Conditions that Significantly Enhance the Impact of Open Data?.....</b>	<b>16</b>
<b>V. What Are the Challenges to Open Data Making an Impact?.....</b>	<b>19</b>
<b>VI. Recommendations: Toward a Next Generation Open Data Roadmap .....</b>	<b>22</b>
Key Remaining Questions .....	32
Open Data Definitions – What’s in a Name?.....	34
Open Data Glossaries .....	36
Open Data Resources .....	38
Open Data Portal Repositories .....	49
Intermediary Repositories .....	50
List of Interviews and Peer Reviewers .....	52



# OPEN DATA IMPACT

## WHEN DEMAND AND SUPPLY MEET

### Key Findings of the Open Data Impact Case Studies

## EXECUTIVE SUMMARY

Recent years have witnessed considerable enthusiasm over open data. Several studies have documented its potential to spur economic innovation and social transformation, and to usher in fresh forms of political and government accountability. Yet for all the enthusiasm, we know little about how open data actually works, and what forms of impact it is really having.

This report seeks to remedy that informational shortcoming. Supported by Omidyar Network, the GovLab has conducted 19 detailed case studies of open data projects around the world. The case studies were selected for their sectoral and geographic representativeness. They were built in part from secondary sources (“desk research”), but also from a number of first-hand interviews with important players and key stakeholders. They are presented at length, in narrative format, on an online repository, Open Data’s Impact ([odimpact.org](http://odimpact.org)). In this paper, we consider some overarching lessons that can be learned from the case studies and assemble them within an analytical framework that can help us better understand what works, and what doesn’t, when it comes to open data.

The paper begins (Part I) with an overview of open data. Like many technical terms, open data is a contested and dynamic concept. The GovLab has conducted a study of ten widely used definitions to arrive at the following working definition, which guides our discussion here:

Open data is publicly available data that can be universally and readily accessed, used and redistributed free of charge. It is structured for usability and computability.

Part II includes a brief summary of our 19 case studies, each of which is detailed at considerably greater length in each stand-alone case study. Parts III-V represent the core of our analytical framework: They identify the key parameters and variables that determine the impact of open data.

Part III discusses what we have identified as the four most important dimensions of impact. Based on the case studies, GovLab has determined that open data projects are *improving government*, primarily by making government more accountable and efficient; *empowering citizens*, by facilitating more informed decision-making and enabling new forms of social mobilization; *creating new economic opportunities*; and *helping policymakers and others find solutions to big, previously intractable public problems* (e.g., related to public health or global warming).

These types of impact cannot be taken for granted. They are evident to varying degrees across our case studies, and sometimes not at all. Our research also identified four enabling conditions that allow the potential of open data to manifest (Part IV). Overall, we found that open data projects work best when they are *based on partnerships and collaborations* among various (often inter-sectoral) organizations; when they emerge within what we call an “*open data public infrastructure*” that enables the regular release of potentially impactful data; when they are accompanied by *clear open data policies*, including performance metrics; and when they address or attempt to solve *a well-defined problem or issue* that is an obvious priority to citizens and likely beneficiaries.

Part V identifies the key challenges that open data projects face. These include *a lack of readiness*, especially evident in the form of low technical and human capacity in societies or nations hosting open data initiatives; projects that are *unresponsive* – and thus inflexible – to user or citizen needs; projects that result in *inadequate protections for privacy or security*; and, finally, projects that suffer from *a shortage of resources*, financial and otherwise. None of the 19 initiatives we studied was immune to these obstacles; the most successful ones had found ways to surmount them and build applications or platforms that were nonetheless able to tap into the potential of open data.

The report ends (Part VI) with a set of 10 recommendations directed at policymakers, entrepreneurs, activists and others contemplating open data projects. Each of these broad recommendations is accompanied by more specific and concrete steps for implementation. Together, these recommendations and steps for implementation add up to something of a toolkit for those working with open data. Although preliminary, they are designed to guide the open data community in its ongoing efforts to launch new initiatives that achieve maximum societal, economic, political and cultural change.



# INTRODUCTION

Recent years have witnessed considerable enthusiasm over the opportunities offered by open data. Across sectors, it is widely believed today that we are entering a new era of information openness and transparency, and that this has the potential to spur economic innovation, social transformation, and fresh forms of political and government accountability. Focusing just on economic impacts, in 2013, for example, the consulting firm McKinsey estimated the possible global value of open data to be over \$3 trillion per year.<sup>1</sup> A study commissioned by Omidyar Network has likewise calculated that open data could result in an extra \$13 trillion over five years in the output of G20 nations.<sup>2</sup>

Yet despite the evident potential of open data, and despite the growing amounts of information being released by governments and corporations, little is actually known about its use and impact. What kind of social and economic transformations has open data brought about, and what transformations may it effect in the future? How – and under what circumstances – has it been most effective? How have open data practitioners mitigated risks (e.g., to privacy) while maximizing social good?

As long as such questions remain unanswered, the field risks suffering from something of a mismatch between the supply (or availability) of data and its actual demand (and subsequent use). This mismatch limits the impact of open data, and inhibits its ability to produce social, economic, political, cultural and environmental change. This report begins from the premise that, in order to fully grasp the opportunities offered by open data, a more full and nuanced understanding of its workings is necessary.

Our knowledge of if, how and when open data actually works in practice is lacking because there have been so few systematic studies of its actual impact and workings. The field is dominated by conjectural estimates of open data's hypothetical impact; those attempts that have been made to study concrete, real-world examples are often anecdotal or suffer from a paucity of information. In this report, we seek to build a more systematic study of open data and its impact by rigorously examining 19 case studies from around the world. These case studies are chosen for their geographic and sectoral representativeness. They are built not simply from secondary sources (e.g., by rehashing news reports) but from extensive interviews with key actors and protagonists who possess valuable and thus far untapped on-the-ground knowledge. They go beyond the descriptive (what happened) to the explanatory (why it happened, and what is the wider relevance or impact).

- 
- 1 Manyika, James, Michael Chui, Diana Farrell, Steve Van Kuiken, Peter Groves, and Elizabeth Almasi Doshi. "Open Data: Unlocking Innovation and Performance with Liquid Innovation." McKinsey Global Institute. November 12, 2013.  
[http://www.mckinsey.com/insights/business\\_technology/open\\_data\\_unlocking\\_innovation\\_and\\_performance\\_with\\_liquid\\_information](http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information).
  - 2 Gruen, Nicholas, John Houghton, and Richard Tooth. "Open for Business: How Open Data Can Help Achieve the G20 Growth Target." Omidyar Network. June 2014.  
[https://www.omidyar.com/sites/default/files/file\\_archive/insights/ON%20Report\\_061114\\_FNL.pdf](https://www.omidyar.com/sites/default/files/file_archive/insights/ON%20Report_061114_FNL.pdf)

In order to provide these explanations, we have assembled an analytical framework (see Picture – TBD) that applies across the 19 case studies and allows us to present some more widely applicable principles for the use and impact of open data. Impact – a better understanding of how and when open data really works – is at the center of our research. Our framework seeks to establish a taxonomy of impact for open data initiatives, outlining various dimensions (from improving government to creating economic opportunities) in which open data has been effective. In addition, the framework lays out some key conditions that enable impact, as well as some challenges faced by open data projects.

## I. WHAT IS OPEN DATA?

It is useful to begin with an understanding of what we mean by open data. Like many technical terms, open data is a contested concept. There exists no single, universally accepted definition. The GovLab recently undertook an analysis of competing meanings, with a view to reaching a working definition. **Appendix I** contains ten widely used definitions and our matrix of analysis.

Based on this matrix, we reached the following working definition, which guides our research and discussion throughout this report:

*Open data is publicly available data that can be universally and readily accessed, used and redistributed free of charge. It is structured for usability and computability.*

It is important to recognize that this is a somewhat idealized version of open data. In truth, few forms of data possess all the attributes included in this definition. The openness of data exists on a continuum, and while many forms of information we discuss here may not be strictly open in the sense described above, they may nonetheless be shareable, usable by third parties, and capable of effecting wide-scale transformation. The 19 case studies included here therefore include a variety of different kinds of data, each of which is open in a different way, and to a different degree. For example:

- Brazil's Open Budget Transparency Portal exemplifies the most "traditional" type of open data project: a downloadable set of open government data accessible to the public.
- Mexico's Mejora Tu Escuela is the result of a nongovernmental organization compiling and presenting data (including open government data) in easily digestible forms;

- The Global Positioning System (GPS) is arguably not an “open data” system at all, but rather a means for providing access to a government-operated signal.
- The UK Ordnance Survey offers a combination of free and paid spatial data, suggesting the possibilities (and limitations) of a mixed model of open and closed data.

In each of these cases, “open” has different meanings and connotations. Many – but not all – of the cases, however, demonstrate the importance of shared and disseminated information, and highlight open data’s potential to enhance the social, economic, cultural and political dimensions of our lives.

## II. THE CASE STUDIES

### METHODOLOGY

To select our case studies, we undertook a multi-step process that involved several variables and considerations. To begin with, we examined existing repositories of open data cases and examples in order to develop an initial universe of known open data projects (see **Appendix III**). This initial scan of existing examples allowed us to identify gaps in representation – those sectors or geographies that often remain underrepresented in existing descriptions of open data and its impact (or lack thereof). In order to fill in some of these gaps (and more generally widen our list of case study candidates), we also reached out to a number of experts in relevant subject areas, for example open data, open governance, civic technology and other related fields. We also attended and conducted outreach at a number of open-data related events, notably the 2015 International Open Data Conference in Ottawa, Canada and ConDatos in Santiago, Chile.

Based on this process, we identified a long list of approximately 50 case studies from around the world. These included examples from the private sector, civil society and government, and spanned the spectrum of openness mentioned above. The next step was to conduct a certain amount of preliminary research to arrive at our final list of 19 case studies. To do this, we took into account several factors: the availability and type of evidence in existence; the need for sectoral and geographic representativeness; and the type of impact demonstrated by the case study in question (if any). We also considered whether previous, detailed case studies existed; as much as possible, our goal was to develop case studies for previously unexplored and undocumented examples.



Having selected our 19 cases, we then began a process of more in-depth researching. This involved a combination of desk research (e.g., using existing media and other reports) and interviews (usually over the telephone). For many of our examples, there existed very little existing research; the bulk – and certainly the most useful – of our evidence came from a series of in-depth interviews we conducted with key participants and observers who had been involved in our various cases. **Appendix IV** includes a full list of interviewees.

Upon completing drafts of each case study, and in the spirit of openness that defines the field under examination, we open-sourced the peer review process for each case and this paper. Rather than sharing drafts only with a select group of experts, we made our report and each of the case studies openly accessible for review in the interest of gaining broad input on our findings and collaboratively producing a common resource on open data's impacts for the field. Through broad outreach at events like the 2015 Open Government Partnership Summit in Mexico City, Mexico and through social media, over 50 individuals from around the world signed up to peer review at least one piece.

During the monthlong open peer review process, more than two dozen of those who signed up shared their input as Recognized Peer Reviewers through in-line comments and in-depth responses to the ideas and evidence presented in the case studies and this paper (see **Appendix IV**). Additionally, the paper and case studies were made openly accessible to the public, allowing anyone to share suggestions, clarifications, notes on potential inaccuracies and any other useful input prior to publishing. Much of this input was integrated into the final versions of the documents.

## THE 19 CASES

The standalone impact case studies include detailed descriptions and analyses of the initiatives listed below. In addition, the table below summarizes their main features and key findings. Here, we include a brief summary of each example:

IMPROVING GOVERNMENT	
<b>Brazil – Open Budget Transparency Portal</b>	
Public Sector Tackling Corruption and Transparency	A tool that aims to increase fiscal transparency of the Brazilian Federal Government through open government budget data. As the quality and quantity of data on the portal have improved over the past decade, the Transparency Portal is now one of the country's primary anti-corruption tools, registering an average of 900,000 unique visitors each month. Local governments throughout Brazil and three other Latin American countries have modeled similar financial transparency initiatives after Brazil's Transparency Portal.
<b>Canada – T3010 Charity Information Return Data</b>	
Philanthropy and Aid Improving Services	In 2013, the Charities Directorate of the Canada Revenue Agency (CRA) opened all T3010 Registered Charity Information Return data since 2000 via the government's data portal under a commercial open data license. The resulting data set has been used to explore the state of the nonprofit sector, improve advocacy by creating a common understanding between regulators and charities, and create intelligence products for donors, fundraisers and grant-makers.
<b>Denmark – Consolidation and Sharing of Address Data</b>	
Geospatial Services Improving Services	In 2005, the Building and Dwelling Register of Denmark started to release its address data to the public free of charge. Prior to that date, each municipality charged a separate fee for access, rendering the data practically inaccessible. There were also significant discrepancies between the content held across different databases. A follow-up study commissioned by the Danish government estimated the direct financial benefits alone for the period 2005-2009 at EUR 62 million, at a cost of only EUR 2 million.
<b>Indonesia – Kawal Pemilu</b>	
Politics and Elections Tackling Corruption and Transparency	A platform launched in the immediate aftermath of the contentious 2014 Indonesian presidential elections. Kawal Pemilu's organizers assembled a team of over 700 volunteers to compare official vote tallies with the original tabulations from polling stations and to digitize the often handwritten forms, making the data more legible and accessible. Assembled in a mere two days, with a total budget of just \$54, the platform enabled citizen participation in monitoring the election results, increased public trust in official tallies, and helped ease an important democratic transition.
<b>Slovakia – Open Contracting Projects</b>	
Public Sector Tackling Corruption and Transparency	In January 2011, Slovakia introduced a regime of unprecedented openness, requiring that all documents related to public procurement (including receipts and contracts) be published online, and making the validity of public contracts contingent on their publication. Over 2 million contracts have now been posted online, and these reforms appear to have had a dramatic effect on both corruption and, equally important for the business climate, perceptions of corruption.
<b>Sweden – openaid.se</b>	
Philanthropy and Aid Tackling Corruption and Transparency	A data hub created by the Swedish Ministry of Foreign Affairs and the Swedish International Development Cooperation Agency (Sida) built on open government data. The website visualizes when, to whom and why aid funding was paid out and what the results were. The reforms are seen to be an important force for enhanced transparency and accountability in development cooperation at an international level and increased cooperation and involvement of more actors in Swedish development policy.

EMPOWERING CITIZENS	
<b>Kenya – Open Duka</b>	
Public Sector Informed Decision-making	A platform developed by the civil society organization, the Open Institute, that aims to address issues of opacity in governance in the private and public sectors, promoting corporate accountability and transparency by providing citizens, journalists and civic activists with insight into the relationships, connections (and, to some extent, the dynamics) of those in and around the public arena. As a case study, it exemplifies the challenge for open data initiatives to generate sufficient awareness and use necessary to achieve impact.
<b>Mexico – Mejora Tu Escuela</b>	
Education Informed Decision-making	A platform created by the Mexican Institute for Competitiveness (IMCO) that provides citizens with information about school performance. It helps parents choose the best option for their children, empowers them to demand higher-quality education and gives them tools to get involved in their children's schooling. It also provides school administrators, policymakers and NGOs with data to identify hotbeds of corruption and areas requiring improvement. Data available on the site was used in a report that uncovered widespread corruption in the Mexican education system and stirred national outrage.
<b>Tanzania – Shule and Education Open Data Dashboard</b>	
Education Social Mobilization	Two recently established portals providing the public with more data on examination pass rates and other information related to school performance in Tanzania. Education Open Data Dashboard is a project established by the Tanzania Open Data Initiative; Shule was spearheaded by Arnold Minde, a programmer, entrepreneur and open data enthusiast. Despite the challenges posed by Tanzania's low Internet penetration rates, these sites are slowly changing the way citizens access information and make decisions, and are encouraging citizens to demand greater accountability from their school system and public officials.
<b>Uruguay – A Tu Servicio</b>	
Health Informed Decision-making	A platform that allows users to select their location and then to compare local health care providers based on a wide range of parameters and indicators, such as facility type, medical specialty, care goals, wait times and patient rights. A Tu Servicio has introduced a new paradigm of patient choice into Uruguay's health care sector, enabling citizens not only to navigate through a range of options but also generating a healthy and informed debate on how more generally to improve the country's health care sector.

## CREATING OPPORTUNITY

## U.K. – OS OpenData

Geospatial Services	Data from Ordnance Survey (OS), Britain's mapping agency, supports essentially any U.K. industry or activity that uses a map: urban planning, real estate development, environmental science, utilities, retail and much more. OS is required to be self-financing and, despite the launch of its OS OpenData platform in 2010, uses a mixed-cost model, with some data open and some data paid. OS OpenData products are estimated to deliver between a net £13 million - £28.5 million increase in GDP over its first five years.
Economic Growth	

## U.S. – New York City Business Atlas

Business	Developed by the Mayor's Office of Data Analytics (MODA), the Business Atlas is a platform designed to alleviate the market research information gap between small and large businesses in New York. The tool provides small businesses with access to high-quality data on the economic conditions in a given neighborhood to help them decide where to establish a new business or expand an existing one.
Economic Growth	

## U.S. – NOAA: Opening Up Global Weather Data in Collaboration with Businesses

Weather	Opening up weather data through NOAA has significantly lowered the economic and human costs of weather-related damage through forecasts; enabled the development of a multi-billion dollar weather derivative financial industry dependent on seasonal data records; and catalyzed a growing million-dollar industry of tools and applications derived from NOAA's real-time data.
Economic Growth	

## U.S. – Opening GPS Data for Civilian Use

Geospatial Services	Over the past 20 years, Global Positioning System (GPS) technology has led to a proliferation of commercial applications across industries and sectors, including agriculture, construction, transportation, aerospace and – especially with the proliferation of portable devices – everyday life. Were the system to be somehow discontinued, losses are estimated to be \$96 billion. In addition to creating new efficiencies and reducing operating costs, the adoption of GPS technology has improved safety, emergency response times and environmental quality, and has delivered many other less-readily quantifiable benefits.
Economic Growth	

## SOLVING PUBLIC PROBLEMS

## New Zealand – Christchurch Earthquake GIS Clusters

Emergency Services  
Data-Driven  
Engagement

In February 2011, Christchurch was struck by a severe earthquake that killed 185 people and caused significant disruption and damage to large portions of a city already weakened by an earlier earthquake. In the response to the quake, volunteers and officials at the recovery agencies used open data, open source tools, trusted data sharing and crowdsourcing to develop a range of products and services required to respond successfully to emerging conditions, including a crowdsourced emergency information Web app that generated 70,000 visits within the first 48 hours after the earthquake, among others.

## Sierra Leone – Battling Ebola

Health  
Data-Driven  
Engagement

In 2014, the largest Ebola outbreak in history occurred in West Africa. At the start, information on Ebola cases and response efforts was dispersed across a diversity of data collectors, and there was little ability to get relevant data into the hands of those who could leverage it. Three projects – Sierra Leone's National Ebola Response Centre (NERC), the United Nations' Humanitarian Data Exchange (HDX) and the Ebola GeoNode – significantly improved the quality and accessibility of information used by humanitarians and policymakers working to address the crisis.

## Singapore – Dengue Cluster Map

Health  
Data-Driven  
Engagement

In 2005, the Singapore National Environment Agency (NEA) began sharing information on the location of dengue clusters as well as disease information and preventive measures online, through a website now commonly known as the "Dengue Website." Since then, the NEA's data-driven cluster map has evolved, and it became an integral part of the campaign against a dengue epidemic in 2013.

## U.S. – Eightmaps

Politics and Elections  
Data-Driven  
Engagement

A tool, launched anonymously in 2009, that provided detailed information on supporters of California's Proposition 8, which sought to bar same-sex couples from marrying. The site collected information made public through state campaign finance disclosure laws and overlaid that information onto a Google map of the state. Users could find the names, approximate locations, amount donated and, where available, employers of individuals who donated money to support Prop 8. Eightmaps demonstrates how the increased computability and reusability of open data could be acted upon in unexpected ways that not only create major privacy concerns for citizens, but could also lead to harassment and threats based on political disagreements.

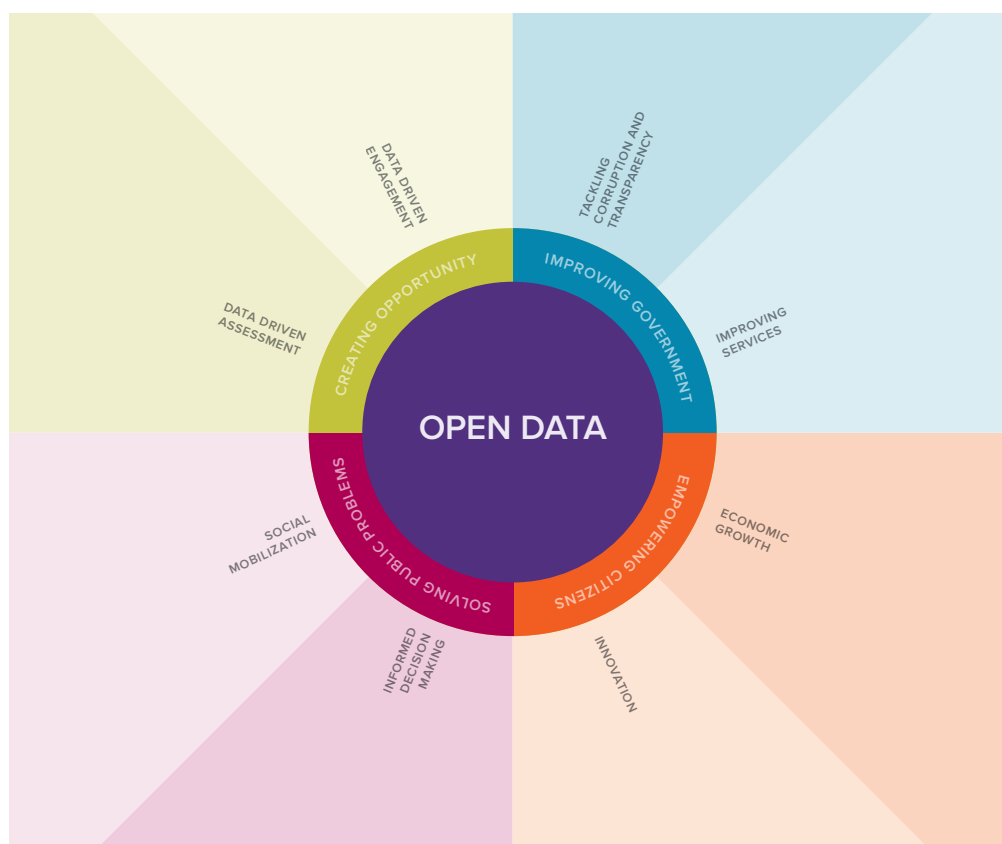
## U.S. – Kennedy vs. the City of Zanesville

Law  
Data-Driven  
Assessment

For over 50 years, while access to clean water from the City of Zanesville water line spread throughout the rest of Muskingum County, residents of a predominantly African-American area of Zanesville, Ohio were only able to use contaminated rainwater or drive to the nearest water tower. One of the key pieces of evidence used during the court case was a map derived from open data that showed significant correlation between the houses occupied by the white residents of Zanesville and the houses hooked up to the city water line. The case went in favor of the African-American plaintiffs, awarding them a \$10.9 million settlement.

### III. WHAT IS THE IMPACT OF OPEN DATA ON PEOPLE'S LIVES?

What lessons can we learn from these examples of open data applications, platforms and websites? In this and the following sections, we outline some overarching insights derived from our 19 case studies. First, we focus on impact. What is the impact of open data on people's lives? What are the real, measurable and tangible results of our case studies? And, just as important, who (which individuals, institutions, demographic groups) are most affected?



*Taxonomy of Open Data Impact*

Determining impact requires taking certain nuances into account. In many cases, open data projects show results in more than one dimension of impact. In addition, the impact of our case studies on people's lives is often indirect (and thus somewhat more subtle), mediated by changes in the way decisions are made or other broad social, political and economic factors. Nonetheless, despite these nuances, our analysis suggests that there exist four main ways in which open data is having an impact on people's lives:



i) First, open data is *improving government* around the world. It is doing so in various ways, but in particular by: a) making governments more accountable, especially by helping tackle corruption and adding transparency to a host of government responsibilities and functions (notably budgeting); and b) making government more efficient, especially by enhancing public services and resource allocation.

Improvements in governance are evident in six of our 19 case studies. Notable examples include the Brazil Open Budget Transparency Portal, which brings accountability and citizen oversight to the country's budget processes; Slovakia's Central Registry, which is a global model for the open contracting movement; and Canada's opening of tax return data submitted by charities, the first move in a broader global effort to increase the transparency and accountability of philanthropies.

ii) Open data is *empowering citizens* to take control of their lives and demand change by enabling more informed decision-making and new forms of social mobilization, both in turn facilitated by new ways of communicating and accessing information.

This dimension of impact plays a role in four case studies. Some notable examples in this category of impact include Uruguay's A Tu Servicio, which empowers citizens to make more informed decisions about health care and education dashboards in Mexico (Mejora Tu Escuela) and Tanzania (Shule and Education Open Data Dashboard), each of which enables parents to make more evidence-based decisions about their children's schools.

iii) Open data is creating *new economic opportunities* for citizens and organizations. Around the world, in cities and countries, greater transparency and more information are stimulating economic growth, opening up new sectors, and fostering innovation. In the process, open data is creating new jobs and new ways for citizens to prosper in the world.

This category of impact often follows from applications and platforms built using government data. It is evident in four of our case studies, each of which relies for its underlying data on information released by governments. Two notable examples include New York's Business Atlas, which allows small businesses to use data to identify the best neighborhoods in which to open or grow their companies; and the various platforms and companies built around data released by the National Oceanic and Atmospheric Administration (NOAA) in the United States.

iv) Finally, open data's impact is evident in the way it is helping *solve several big public problems*, many of which have until recently seemed intractable. Although most of these problems have not been entirely solved or eliminated, we are finally seeing pathways to improvements. Open data is allowing citizens and policymakers to analyze societal problems in new ways and engage in new forms of data-driven assessment and engagement.

Open data has created notable impacts during public health crises and other emergencies. In Sierra Leone, open data helped to inform the actions of people working on the ground to fight Ebola. The government and citizens of Singapore are using a Dengue Fever Cluster Map to try to limit the spread

of dengue fever during outbreaks like that experienced in 2013. The efforts to rebuild following devastating earthquakes in Christchurch, New Zealand were also aided by open data. It is important to recognize, however, that attempts to solve problems can also have unintended consequences. We see this, for example, in the case of Eightmaps, where efforts to address discrimination and other issues unintentionally created new privacy (and even personal security) problems.

## IV. WHAT ARE THE ENABLING CONDITIONS THAT SIGNIFICANTLY ENHANCE THE IMPACT OF OPEN DATA?

While our initial analysis told us what *types* of change open data was creating, a further round of analysis was required to understand *how* change comes about. In examining open data projects around the world, we are struck by the wide variability in outcomes. Some work better than others, and some simply fail. Eightmaps is an example of how open data can lead to unintended consequences, but there are many, many more examples that the GovLab did not select for this group of case studies due to the lack of meaningful, measurable impact to date. Some projects do well in a particular dimension of success while failing in others. If we are to achieve the believed potential of open data and scale the impact of the individual case studies included here, we need a better – and more granular – understanding of the enabling conditions that lead to success.

### ENABLING CONDITIONS

**PARTNERSHIPS**

**PUBLIC INFRASTRUCTURE**

**POLICIES AND  
PERFORMANCE METRICS**

**PROBLEM DEFINITION**

Based on our research, we identified four key enabling conditions, each of which allows us to articulate a specific “premise” for success:

**i) Partnerships:** The *power of collaboration* was evident in many of the most successful open data projects we studied. Effective projects were built not from the efforts of a single organization or government agency, but rather from partnerships across sectors and sometimes borders. Two forms of collaboration were particularly important: partnerships with civil society groups, which often played an important role in mobilizing and educating citizens; and partnerships with the media, which informed citizens and also played an invaluable role in analyzing and finding meaning in raw open data. In addition, we saw an important role played by so-called “data collaboratives,” which pooled data from different organizations and sectors.

Virtually all the case studies we examined were the products of some form of partnership. Uruguay’s A Tu Servicio was an important example of how civil society can work with government to craft more effective open data initiatives. NOAA’s many offshoots and data initiatives are an equally important example of collaboration between the private and public sectors. New York City’s Business Atlas was similarly an illustration of a public-private partnership; its data set, built both from government and private-sector information (supplied by the company Placemeter), is an example of an effective data collaborative.

**Premise #1: Intermediaries and data collaboratives allow for enhanced matching of supply and demand of data.**

**ii) Public Infrastructure:** Several of the most effective projects we studied emerged on the back of what we might think of as an *open data public infrastructure* – i.e., the technical backend and organizational processes necessary to enable the regular release of potentially impactful data to the public. In some cases, this infrastructure takes the form of an “open by default” system of government data generation and release. The team behind Kenya’s Open Duka, for example, is responding to its lack of impact to date by attempting to build such an infrastructure with county-level governments to improve the counties’ internal data capacity, improving the data available on Open Duka as a result.

An open data public infrastructure does not, however, only involve technical competencies. As part of the push around Brazil’s Open Budget Transparency Portal, for example, organizers not only developed an interoperable infrastructure for publishing a wide variety of data formats, but also launched a culture-building campaign complete with workshops seeking to train public officials, citizens and reporters to create value from the open data.

**Premise #2: Developing open data as a public infrastructure enables a broader impact across issues and sectors.**

**iii) Policies and Performance Metrics:** Another key determinant in the success of open data projects was the existence of *clear open data policies, including well-defined performance metrics*. The need for clear policies (and more generally an enabling regulatory framework) is a reminder that technology does not exist in a vacuum. Policymakers and political leaders have an essential role to play in creating a flexible, forward-looking legal environment that, among other things, encourages the release of open data and technical innovation; and that spurs the creation of fora and mechanisms for project assessment and accountability.

In addition, high-level political buy-in is also critical. It is not sufficient simply to pass enabling laws that look good on paper. Policymakers and politicians must also ensure that the letter of the law is followed, that vested interests are adequately combated, and that there are consequences for working against openness and transparency.

Among the many case studies that benefited from the right policy environment, a few stand out. In Mexico, we saw how an open data initiative (in this case, the Mejora Tu Escuela project) can benefit from high-level government commitments to opening data that trickles down to – and empowers – local and regional governments. Slovakia's Central Registry is another good example; it shows how laws can be redesigned, in this case to encourage transparency by default in contracting, and in the process greatly increase openness. The openness of GPS, though ingrained in daily life for many, was the subject of questions following the terrorist attacks of September 11, 2001; those questions were put to rest with the enactment of a new policy commitment in 2004 to maintain unfettered global access to the geospatial system.

**Premise #3: Clear policies regarding open data, including those promoting regular assessments of open data projects, provide the necessary conditions for success.**

**iv) Problem Definition:** We have repeatedly seen how the most successful open data projects are those that address *a well-defined problem or issue*. It is very challenging for open data projects to try to change user behavior or convince citizens of a previously unfelt need. Effective projects identify an existing – ideally widely recognized – need, and provide new solutions or efficiencies to address that need.

Singapore's Dengue Fever Cluster Map is a good example in this regard. Its core area of activity (public health) has clear, tangible benefits; it seeks to limit the spread of an illness that policymakers widely recognize as a problem, and that citizens dread. Uruguay's A Tu Servicio is another good example – it provides clear, tangible benefits to citizens, allowing them to take action that improves their health care. It is perhaps no coincidence that both these examples are in the health sector: The most successful projects often touch on the most basic human needs (health, pocketbook needs, etc.). In a case involving one of the most essential human needs, the use of open data in Kennedy vs. the City of Zanesville accomplished its singular goal: demonstrating beyond a reasonable doubt that water access decisions were being made on the basis of citizens' race.

**Premise #4: Open data initiatives that have a clear target or problem definition have more impact.**

## V. WHAT ARE THE CHALLENGES TO OPEN DATA MAKING AN IMPACT?

The success of a project is also determined by the obstacles and challenges it confronts. The challenges are themselves the function of numerous social, economic and political variables. In addition, some regions may face more obstacles than others.



As with the enabling conditions, we found widespread geographic and sectoral variability in our analysis of challenges. Broadly, we identified four challenges that recurred the most frequently across our 19 case studies:

**i) Readiness:** Perhaps unsurprisingly, countries or regions with *overall low technical and human capacity or readiness* often posed inhospitable environments for open data projects. The lack of technical capacity could be indicated by several variables: low Internet penetration rates, a wide digital divide, or overall poor technical literacy. In addition, technical readiness can also be indicated by the existence of a group of individuals or entities that are technically sophisticated, and that believe in the transformative potential of technology, particularly of open data. Repeatedly, we have seen that such “data champions” or “technological evangelists” play a critical role in ensuring the success of projects.

Low technical capacity did not necessarily result in outright project “failures.” Rather, it often stunted the potential of projects, making them less impactful and successful than they could otherwise have been. In Tanzania, for instance, the Shule and Education Open Data Dashboard portals were limited by low Internet penetration rates, and by a general low awareness about

open data. Slovakia's Central Registry was in many ways very successful; yet it, too, was restricted by a lack of technical capacity among government officials and others (particularly at the lower level). In these projects and others, we see that success is relative, and that even the most successful projects could be enhanced by greater attention to the overall technical environment or ecosystem.

**Premise # 5: The lack of readiness or capacity at both the supply and demand side of open data hampers its impact.**

**ii) Responsiveness:** Success is also limited when projects are *unresponsive to feedback and user needs*. As we saw in the previous section, the most successful projects address a clear and well-defined need. A corollary to this finding is that project sponsors and administrators need to be attuned to user needs; they need to be flexible enough to recognize and adapt to what their users want.

For Sweden's OpenAid project, for example, user experience was not a core priority at launch, and much of the information found on the site is too complex for most citizens to digest. Despite this high barrier to entry, the site only offers limited engagement opportunities – namely, a button for reporting bugs on the site. Moreover, project titles found on the site often contain cryptic terms interpretable only to those with close familiarity of the project at hand.

NOAA, on the other hand, has some of the most mature and wide-reaching open data efforts in any of the cases studied here. But given that breadth, for the agency's essential information to remain useful to the evolving needs of its users, an increased focus needs to be placed on customer analytics and user behaviors. U.K.'s Ordnance Survey has very sophisticated user analytics and prioritizes customer satisfaction; however, the separation of OS OpenData from its other data sets and products is potentially limiting.

**Premise #6: Open data could be significantly more impactful if the release of open data would be complemented with a responsiveness to act upon insights generated.**

**iii) Risks:** A major challenge arises from the trade-offs between the potential of open data and the *risks posed by privacy and security violations*. These risks are inherent to any open data project – by its very nature, greater transparency exists in tension with privacy and security. When an initiative fails to take steps to mitigate this tension, it risks not only harming its own prospects, but more broadly the reputation of open data in general.

Concerns about privacy and security dogged many of the projects we studied. In Brazil, over 100 legal actions were brought against the Open Budget Transparency Portal when it inadvertently published the salaries of public servants. In New York, despite steps being taken to mitigate such harms, there has been concern that citizen privacy might be violated as cameras collect data for the project in public spaces. Without question, the clearest example



of open data leading to privacy concerns (and even outright violations) can be found in the Eightmaps case study, which used public campaign finance disclosure laws to publish various identifying information about and home addresses for donors to California's Proposition 8, leading to instances of intimidation and harassment.

For all the very real – and legitimate – concerns, our case studies also show that the scope for privacy and security abuses can be mitigated. For example, NOAA stood out for its creation of a dedicated Cyber Security Division to address data security challenges when collecting and releasing data (the sole instance of such a dedicated division in our 19 case studies). Singapore, too, took proactive steps to anonymize data to protect the privacy of citizens. Addressing risks to privacy and security, though important, can also work against the goals of openness and transparency – for example, in the city of Zanesville, Ohio, security concerns have been raised (controversially) to restrict access to data that has proven essential in addressing decades-old civil rights violations. Such examples are an important reminder of the tensions that exist between openness and security/privacy, and of the need for careful, judicious policymaking to achieve a balance.

**Premise #7: Open data does pose a certain set of risks, notably to privacy and security; a greater, more nuanced understanding of these risks will be necessary to address and mitigate them.**

**iv) Resource Allocation:** Finally, we found that *inadequate resource allocation* was one of the most common reasons for limited success or outright failure. Many of the projects we studied were “hackable” – easily put together on a very limited budget, often created by idealistic volunteers. Indonesia's Kawal Pemilu, for example, was assembled with a mere \$54. Over time, though, projects require resources to succeed; while they may emerge on the backs of committed (and cheap) idealists, they are fleshed out and developed with real financial backing.

The limited success of Kenya's Open Duka is a good example. Although the project was well-conceived and based on a sound premise, it has been limited by the unanticipated effort involved in data collection; more resources would almost certainly have helped address this challenge. In addition, Mexico's Mejora Tu Escuela is just one project that relies on foundation funding in order to operate – leading to some level of uncertainty about the long-term sustainability of such projects should any of those funding streams be discontinued. U.K.'s Ordnance Survey, meanwhile, is required to be self-financing, forcing the agency to rely heavily on private sector customers paying to access the more sophisticated data products not included in OS OpenData.

Even an initiative as central and widely used as GPS experiences funding challenges. In a government climate focused on budget cuts at every corner, new features and capabilities, even for a “global public utility,” can be difficult to finance through public money.

**Premise #8: While open data projects can often be launched cheaply, those projects that receive generous, sustained and committed funding have a better chance of success over the medium and long term.**

## VI. RECOMMENDATIONS: TOWARD A NEXT GENERATION OPEN DATA ROADMAP

Our case studies clearly indicate the tremendous potential and possibilities offered by open data. Around the world, open data has improved governments, empowered citizens, contributed solutions to complex public problems, and created new economic opportunities for companies, individuals and nations.

But despite this clear potential, the hurdles are also apparent. We outline several of the particular issues faced by open data projects above. In addition to these specific challenges, there is the more general problem of scaling: How do we move beyond a “points of light” narrative that celebrates individual case studies to a broader narrative about the social, economic and political transformation that could result from a far broader deployment of open data? In this section, we outline 10 steps or recommendations for policymakers, advocates, users, funders and other stakeholders in the open data community that we believe could usher in such wholesale transformation. For each step, we describe a few concrete methods of implementation – ways to translate the broader recommendation into meaningful impact.

Together, these 10 recommendations and their means of implementation amount to a **Next Generation Open Data Roadmap**. They allow us to better understand how the potential of open data can be fulfilled – across geographies, sectors and demographics.

### SUPPLY VS. DEMAND TRAJECTORIES

In studying the ways in which open data has been made available, we’ve found consistent trajectories depending on whether the data is pushed from the government, or made available by users in civil society or the general public extracting that data from reluctant institutions. Interestingly, we’ve found that as both open data push and pull trajectories advance, the optimal end point is the same: greater collaboration between data holders and data users.

## TRAJECTORY OF OPEN DATA

### PUSH

**Data release** – simply making some amount of data available

**Open by default** – creating the infrastructure and processes needed for constant, automatic data release

**Demand-driven collaboration** – working with users to make the most useful data available in the most useful ways

### PULL

**Data audit and gap identification** – outside assessment of where data could have an impact if made accessible

**Creation and demand** – through scraping, Freedom of Information requests, data leaks or other methods, data users finding ways to make government data accessible without the direct involvement (and often without the blessing) of the data holding institution

**Collaboration** – working with government to craft impactful data release strategies

## EIGHT PREMISES THAT DETERMINE THE IMPACT OF OPEN DATA

**Premise #1:** Intermediaries and data collaboratives allow for enhanced matching of supply and demand of data.

**Premise #2:** Developing open data as a public infrastructure enables a broader impact across issues and sectors.

**Premise #3:** Clear policies regarding open data, including those promoting regular assessments of open data projects, provide the necessary conditions for success.

**Premise #4:** Open data initiatives that have a clear target or problem definition have more impact.

**Premise #5:** The lack of readiness or capacity at both the supply and demand side of open data hampers its impact.

**Premise #6:** Open data could be significantly more impactful if the release of open data would be complemented with a responsiveness to act upon insights generated.

**Premise #7:** Open data does pose a certain set of risks, notably to privacy and security; a greater, more nuanced understanding of these risks will be necessary to address and mitigate them.

**Premise #8:** While open data projects can often be launched cheaply, those projects that receive generous, sustained and committed funding have a better chance of success over the medium and long term.

### **Recommendation #1: Focus on and define key problem areas where open data can add value.**

A core premise offered by our case studies is that the impact of open data is often dependent on how well the problem it seeks to address is defined and understood. It is therefore essential for open data advocates and practitioners to clearly define their goals, the problem they are seeking to address, and the steps they plan to take. Some possibilities for how this focus can be achieved:

- Set up a crowdsourced “Problem Inventory” to which users can contribute specific questions and answers, both of which can help define open data projects. The UK Ordnance Survey’s GeoVation Hub is an interesting model focusing on the latter. It poses very specific questions (e.g., How can we improve transport? How can we feed Britain?) for users to answer using OS OpenData.
- Facilitate user-led design exercises to help define important public and social problems and how open data can help solve them.
- To guide such exercises, it may be useful to establish Problem and Data Definition toolkits – potentially modeled on and informed by Freedom of Information requests – that help formulate clearly defined public issues and connect them with potentially useful open data streams.

### **Recommendation #2: Encourage collaborations across sectors (especially between government, private sector and civil society) to better match the supply and demand of open data.**

Large public problems are by definition cross-sectoral and inter-disciplinary. They define boundaries and require a variety of expertise, knowledge and data in order to be successfully addressed. It therefore stands to reason that the most successful open data projects will similarly be collaborative and work across sectors and disciplines. Working in a collaborative manner can help draw on a diverse pool of talent, and can also lead to innovative, out-of-the-box solutions. Perhaps most importantly, by allowing data users and data suppliers to work together and interact, collaborative approaches can improve the match between data demand and supply, thus enhancing the overall efficiency of the demand-use-impact value chain for open data.

Some pathways to achieving the required collaborative and cross-sectoral approaches:

- Create data collaboratives to improve the efficiency and effectiveness of the demand-use-impact cycle. The value of data collaboratives is clearly illustrated by New Zealand’s Canterbury Earthquake Recovery Authority’s data sharing with construction companies, which is projected to deliver NZ\$40 million in savings. In addition, NOAA’s Big Data Partnership, which formalized a sector partnership with five leading private-sector data and cloud technology companies, is also a good example.
- Engage and nurture data intermediaries, especially from civil society, to help spread awareness and disseminate data (and their findings) more widely. Data intermediaries play a particularly important role in countries with low technical capacity (e.g., as evident in our Tanzanian case study); they offer a vital link between technology and society, helping citizens maximize and make real, effective use of data in their everyday lives.

### Recommendation #3: Approach and treat data as a form of vital 21st century public infrastructure.

Too often, policy- and decision-makers focus solely on opening up data, as if open data on its own provides a silver bullet for a society's problems. In fact, as repeatedly evidenced in our case studies, data – in its raw form – needs to be supplemented by a host of other commitments: sustained and sustainable funding, skills training among those charged with data collection and use, and effective governance structures for every step of the data collection and use cycle. Approaching data in this broader, more holistic way means treating it as a vital form of public infrastructure, one at the heart of a society or nation, essential for its success, and embedded within wider social, economic and political structures.

There are several steps policymakers can take to advance a “data-as-infrastructure” approach. These include:

- Developing a **systems design and mapping methodology**. Mapping the public and private sector data infrastructure, as well as local, national and global data infrastructures that may impact the value creation of open data is a first and necessary step to approach data as infrastructure. A systems map could enable the more targeted, coordinated, collaborative development of open data technical standards and best practices across sectors.
- Embracing and implementing the **Open Data Charter**,<sup>3</sup> which seeks to “foster greater coherence and collaboration” around open data standards, practices and, in particular, the following principles:

- Open by default
- Timely and comprehensive
- Accessible and usable
- Comparable and interoperable
- For improved governance and citizen engagement
- For inclusive development and innovation

- Leveraging existing public infrastructure, such as libraries, schools and other cultural and education institutions, so that data is more firmly embedded into other forms of public investment and public life. Open Referral, for example, is creating a data backend for the social safety net, allowing pilot partners, including libraries, to tap into a wide, interconnected range of potentially impactful data on civic and social services.
- Developing skills and capacity around data collection, cleaning and standardization to ensure better quality data is being released. This is especially important within agencies and

3 <http://opendatacharter.net/>



organizations releasing data (to ensure the quality of data), but also, to the extent possible, within the community of users.

- Viewing and treating open data as a public good, something to which citizens and taxpayers are entitled. Moving toward a view of open data as a public good requires as much of a cultural change as a policy change: As our case studies have repeatedly shown, the success of open data initiatives depends crucially on government stakeholders accepting that citizens – whether researchers, journalists or just average individuals – have a right to demand access to government data.

#### **Recommendation #4: Create clear open data policies that are measurable and allow for agile evolution.**

Our research illustrates the vital enabling role played by a national legal and regulatory framework that supports open data. Well-articulated internal rules and priorities are equally important when the releasing entity is a company or other organization. In both cases, clarity is essential: Open data thrives when there is an unambiguous commitment to its cause. Importantly, open data policies should include provisions to measure the success (or otherwise) of an initiative; systems for measurement and assessment are vital to ensuring accountability.

There are several steps policymakers can take to ensure the necessary clarity of open data policies. These include:

- Co-creating open data policies with citizen and other groups, which can be an important way not only of drafting inclusive (and thus more legitimate) policies, but also of ensuring that policies are responsive to actual conditions and needs. Our research repeatedly shows that policies drafted without adequate public input and participation are less effective than those that draw on a wider range of experiences and expertise. Of course, attention must be paid to knowledge and power asymmetries involved in such co-creation processes.
- Engaging the public in defining and monitoring metrics of success: Citizen participation in measuring the results of open data initiatives is as important as in drafting policies, and for the same reasons. It is a vital part of ensuring accountability and in enhancing the legitimacy and effectiveness of open data projects.
- Creating a “Metrics Bank” of important indicators, with input from stakeholders, researchers and experts in the field. Such a Metrics Bank could be built around the variety of categories of open data’s impacts, such as economic concerns (like return on investment or private sector economic revenues generated), public problem solutions (lives saved, increases in the efficiency of service delivery), and others. In line with the previous suggestion, the Metrics Bank should be reviewed on a regular basis by a citizens’ group or panel created specifically for that purpose.

## **Recommendation #5: Take steps to increase the capacity of public and private actors to make meaningful use of open data.**

Repeatedly, we have seen how open data initiatives are limited by a lack of capacity and preparedness among those who could potentially benefit most. Often, this manifests quite simply as a lack of awareness: Those who do not know about the potential of open data are likely to use and benefit less from it. It is important to recognize that low capacity is a problem both on the demand side and supply side of the open data value chain – policymakers and those tasked with releasing data are often as unprepared as intended beneficiaries.

Several steps can be taken to increase capacity and preparedness:

- Set up coaching and training centers to teach policymakers and key stakeholders among citizens about the potential benefits and applications of open data. Brazil's Open Budget Transparency Portal, for instance, benefited tremendously from TV campaigns and regular workshops designed to train citizens, reporters and public officials on how to use the Open Budget Transparency Portal. In addition, a combined overview or searchable directory of coaching opportunities already in place and provided by, for instance, the GovLab Academy and the Open Data Institute, could enable easier navigation and matching of interests and needs worldwide.
- Establish mentor and expert networks for those seeking to use open data. Such networks can serve as valuable resources, providing guidance on the optimal uses of open data and helping citizens and policymakers overcome hurdles or navigate obstacles.
- Invest in and promote user-friendly data tools, such as data visualizations and other analytic tools. While raw data can often be overwhelming for novice users, platforms and apps that include analytics and visualizations are often far more accessible. Notable examples from our case studies include the UK Ordnance Survey's OS OpenMap, NYC's Business Atlas and Mexico's Mejora Tu Escuela.
- Use online and offline meet-ups and similar tools to create a culture that encourages knowledge sharing and collaboration. Many off-the-shelf tools already exist; if integrated within open data initiatives or data labs – like the Justice Data Lab in the United Kingdom – they can provide a helpful online supplement to the types of training efforts and expert-mentor networks mentioned above.

## **Recommendation #6: Identify and manage risks associated with the release and use of open data.**

As our case studies have shown, open data can be a force for good, but it is not without risks. Two of the most important risks involve potential violations of privacy and security that can result from widespread releases of data. Such risks were apparent in a number of our case studies, notably Eightmaps, Brazil's Open Budget Transparency Portal, and New York's Business Atlas. Mitigating such risks is essential not only for its inherent value, but also because privacy and security violations undermine trust in open data and, over the long run, limit its potential.

Several steps can be taken to mitigate risks:

- Develop data governance “decision trees” to help decision-makers track the potential risks and opportunities around certain types of data releases. These decision trees can also help weigh the pros and cons and relative risks of data releases.
- Create innovative, collaborative open data risk management frameworks so that governments and other institutions releasing data can draw on a clear, structured, step-by-step process to strategically respond to breaches of privacy, security or other risks. NOAA, for example, is working with outside experts to crowdsource new frameworks for data management.
- Involve all stakeholders (including citizen groups) in developing data quality and risk standards. A participatory, collaborative approach to mitigating risks can build trust and help achieve the right balance between social goods like innovation, on the one hand, and risks like privacy and security, on the other hand. Crowdsourcing can be a valuable tool here, allowing policymakers to solicit a wide range of responses from diverse stakeholder groups.

## **Recommendation #7: Be responsive to the needs, demands and questions generated from the use of open data.**

We have seen that public participation is essential in the drafting of open data policies and in decisions about what data to release. It is equally important in understanding the impact of open data and in taking advantage of the opportunities it offers. For example, open data can generate insights that require government action; open data can likewise reveal inefficiencies that need concrete steps in order to be addressed. And as we have seen in the Brazilian case study on preventing government corruption, meaningful responsiveness requires the ability to take such steps and actions; what's required are communities focused on problem solving, not simply on releasing data.

Meaningful responsiveness can be achieved through the following methods:

- Develop open and online feedback mechanisms, including Q&As, ratings and feedback tools to gauge public opinion and solicit insights from citizens. For example, Denmark's Open Address Initiative has a single portal for users to correct data errors across all agencies. Simplified mechanisms such as this help establish a virtuous open data cycle, allowing open

data to generate insights and ensuring meaningful action on those insights.

- Designate an open data ombudsman function to consistently track the usefulness of open data and whether necessary follow-up actions are being taken. This ombudsman should itself be open and transparent, and ideally include a wide range of stakeholder inputs.

**Recommendation #8: Allocate and identify adequate resources to sustain and expand the necessary open data infrastructure in a participatory manner.**

As noted, open data initiatives are often cheap to get off the ground, but require resources and investment over time. Goals such as increased participation and transparency are laudable, but without resource commitments, they may remain unachievable. Kenya's Open Duka project is a good example of a laudable open data initiative that has been limited by a lack of resources. Similarly, Canada's Open Charity Initiative T3010 has not yet been updated since its original 2013 release, in part due to a lack of funding, with the result that anyone seeking recent data on Canadian charities must now scrape information independently.

Adequate resource allocations can be achieved by:

- Participatory budgeting initiatives, which allow citizens to choose their priorities and how public funds are allocated. Such initiatives can ensure that the most useful open data initiatives receive the most funding.
- Undertaking more rigorous cost/benefit analyses of open data initiatives, which would allow policymakers and other stakeholders to assess the relative opportunities offered by projects against their costs and possible risks. Among our case studies, NOAA and the UK Ordnance Survey both commissioned cost/benefit studies before launching their projects – this played a vital role in bolstering support and long-term commitments from policymakers and government stakeholders.
- Exploring innovative avenues for funding, especially crowdsourcing, which may offer the public (and other interested parties) an avenue not only for funding initiatives but also for establishing and ensuring the sustainability of their priorities.

**Recommendation #9: Develop a common research agenda to move toward evidence-based open data policies and practices.**

The most effective avenue to understanding how open data works, and how to achieve maximum positive impact, is through collaboration. Our knowledge of open data today is in many ways fragmentary, spread across organizations and individuals who are themselves scattered across the globe. There is a need for more communication and pooling of analysis (and resources). To achieve the potential of open data, we need a common research agenda, based on a wider evidential foundation. Importantly, this research

framework should integrate a better understanding of impact into its core agenda: Too often, open data research focuses simply on the best ways of releasing data, with impact – positive or negative – being simply an afterthought.

To achieve this common research agenda, we should:

- Set up mechanisms for communication and interaction among various stakeholders (individuals and organizations) currently working in the field of open data. Such mechanisms could include annual meetings or conferences, listservs, monthly hangouts, and other offline and online tools. The goal of these interactions would be to trade insights and ideas, to share evidence, and to collaboratively develop best practices. Events like the Open Data Research Summit within the context of the International Open Data Conference, may provide, for instance, the impetus toward improved exchange and collaboration among researchers in this field.
- Build on the taxonomy of impact developed through these 19 case studies and have other researchers test the premises we identified above. In addition, the Open Data research community could consider further fine-tuning of the open data common assessment framework<sup>4</sup> GovLab developed together with Web Foundation and others in order to create a standardized tool for evaluating every stage of the open data value chain.
- Create a directory (perhaps in wiki format) of various assessment frameworks (in addition to our own), spread across countries and sectors. Such a directory would also include a list of key contacts and organizations, and would help facilitate discussion by establishing a baseline of sorts toward achieving a common research agenda.

### **Recommendation #10: Keep innovating.**

Open data fuels innovation, but how can we innovate open data? We need to recognize different forms and models of open data – including big and small data, text-based data – and encourage stakeholders to think broadly about what data is and what *open* really means. Even while we work to better understand open data and its impact (for example, through exercises such as this one), we should foster a culture of proactive experimentation and innovation.

There are many ways to foster such a culture:

- Institutionally, we can look at creating new entities or intermediaries, for example a global open data innovation lab whose explicit purpose would be to think outside the box and research new models of open data that can be tested across sectors, regions and use cases.
- The need for collaborative research mentioned above can also be institutionally developed into a cross-border and interdisciplinary open data innovation network. Such a network would draw on global expertise and ideas.
- Perhaps most importantly, we need to be open to new ideas and insights, and always remain

4 <http://opendataresearch.org/sites/default/files/posts/Common%20Assessment%20Workshop%20Report.pdf>

in *question mode*. This report has outlined several recommendations and suggestions for how to maximize the value of open data. But we recognize that this is just a beginning. Our research has raised as many questions as it has suggested answers.

We end, therefore, with what we believe to be some of the most important questions we should be asking ourselves about open data – questions that can help direct future research, but perhaps most importantly fuel a culture of innovation and flexibility when it comes to how we think about open data.

## KEY REMAINING QUESTIONS

The preceding findings and recommendations for policymakers and stakeholders in the open data community are based on the examination of 19 case studies of open data initiatives from around the world. Though this effort enabled a major step forward in our understanding of open data and its real and potential impacts, key questions remain, including:

- What are the optimal value propositions (e.g., fighting corruption, spurring economic activity, citizens' right to government information) to highlight in order to spur open data activity in different contexts based on local priorities and needs?
- What are the conditions to scale the impact of open data?
- How can open data initiatives be made sustainable?
- What comparative insights are transferable in a universal manner?
- What is the optimal internal data infrastructure for enabling impactful open data initiatives?



## TOWARD A NEXT GENERATION OPEN DATA ROADMAP

**Recommendation #1:** Focus on and define key problem areas where open data can add value.

**Recommendation #2:** Encourage collaborations across sectors (especially between the state, private sector and civil society) to better match the supply and demand of open data.

**Recommendation #3:** Approach and treat data as a form of vital 21st century public infrastructure.

**Recommendation #4:** Create clear open data policies that are measurable and allow for agile evolution.

**Recommendation #5:** Take steps to increase the capacity of public and private actors to make meaningful use of open data.

**Recommendation #6:** Identify and manage risks associated with the release and use of open data.

**Recommendation #7:** Be responsive to the needs, demands and questions generated from the use of open data.

**Recommendation #8:** Allocate and identify adequate resources to sustain and expand the necessary open data infrastructure in a participatory manner.

**Recommendation #9:** Develop a common research agenda to move toward evidence-based open data policies and practices.

**Recommendation #10:** Keep innovating.

## APPENDIX I

# OPEN DATA DEFINITIONS – WHAT'S IN A NAME?

1. **Open Definition** (referenced by Open Data Handbook, ODI, Open Data Census, and OECD Open Data Analytical Framework)

“Open data is data that can be freely used, reused and redistributed by anyone – subject only, at most, to the requirement to attribute and sharealike.”

2. **The White House, 2013 OMB Memorandum**

“Open data refers to publicly available data structured in a way that enables the data to be fully discoverable and usable by end users.”

3. **Data.Gov.UK**

“Open data is data that is published in an open format, is machine readable and is published under a license that allows for free reuse.”

4. **Dbpedia: A nucleus for a web of open data**

“Open data is the idea that certain data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control.”

5. **Open Data Institute**

“Open data is information that is available for anyone to use, for any purpose, at no cost. Open data has to have a license that says it is open data. Without a license, the data can't be reused. These principles for open data are described in detail in the Open Definition.”

## 6. **LinkedGov**

**“Open data is non-personally identifiable data produced in the course of an organization’s ordinary business, which has been released under an unrestricted license. Open public data is underpinned by the philosophy that data generated or collected by organizations in the public sector should belong to the taxpayers, wherever financially feasible and where releasing it won’t violate any laws or rights to privacy (either for citizens or government staff).”**

## 7. **McKinsey Global Institute**

**“Machine-readable information, particularly government data, that’s made available to others. These open datasets share the following 4 characteristics: 1. Accessibility: A wide range of users is permitted to access the data. 2. Machine readability: The data can be processed automatically. 3. Cost: Data can be accessed free or at negligible cost. 4. Rights: Limitations on the use, transformation, and distribution of data are minimal.”**

## 8. **Open Data Now**

**“Open Data is accessible public data we can use to launch new ventures, analyze trends, make decisions, and solve problems.”**

## 9. **Open Data Barometer**

Excerpt from report indicates that researchers assessed datasets based on the

**“full Open Definition requirements of being machine readable, accessible in bulk, and openly licensed.”**

## 10. **The World Bank**

**“Data is open if it satisfies both conditions below:**

**Technically open:** available in a machine-readable standard format, which means it can be retrieved and meaningfully processed by a computer application

**Legally open:** explicitly licensed in a way that permits commercial and non-commercial use and re-use without restrictions.”

## APPENDIX II

# OPEN DATA GLOSSARIES

Below, we list and describe five key efforts to define the many terms related to open data – from API to XML – with a focus on how the available glossaries differ from one another. Additionally Becky Hogge’s “Open Data: Six Stories About Impact in the UK,” also commissioned by Omidyar Network, offers a glossary of terms related to open data and the six initiatives she studied.<sup>5</sup>

### Data.gov

**URL:** <https://www.data.gov/glossary>

**Description:** “Glossary of Terms”

**Terms:** 39

**Distinguishing Characteristics:** Definitions sourced from a diversity of linked secondary sources. Categories of definitions for general terms and for terms related to metadata.

### Data.gov.uk

**URL:** <https://data.gov.uk/glossary>

**Description:** “This glossary is intended to be an authoritative explanation of the meaning of technical terms, for all users of data.gov.uk. Users are encouraged to improve it by suggesting a better way of explaining the definitions, and by adding new definitions.”

**Terms:** 75

**Distinguishing Characteristics:** “Search for Term” functionality allowing the user to find instances of the term in the data.gov.uk catalog. Ability to suggest a new term.

<sup>5</sup> Hogge, Becky. “Open Data: Six Stories About Impact in the UK.” Omidyar Network. November 2015. [https://www.omidyar.com/sites/default/files/le\\_archive/insights/Open%20Data\\_Six%20Stories%20About%20Impact%20in%20the%20UK/Open-Data\\_CaseStudies\\_Report\\_complete\\_DIGITAL\\_102715.pdf](https://www.omidyar.com/sites/default/files/le_archive/insights/Open%20Data_Six%20Stories%20About%20Impact%20in%20the%20UK/Open-Data_CaseStudies_Report_complete_DIGITAL_102715.pdf)

## Open Data Handbook

**URL:** <http://opendatahandbook.org/glossary/en/>

**Description:** N/A

**Terms:** 121

**Distinguishing Characteristics:** The most extensive glossary focusing exclusively on open data. Available in 20 languages.

## Project Open Data

**URL:** <https://project-open-data.cio.gov/glossary/>

**Description:** “This section contains explanations of common terms referenced in Project Open Data and the Open Data Policy.”

**Terms:** 46

**Distinguishing Characteristics:** Generated through crowdsourced, wiki functionality, allowing users to edit the glossary and view its revision history.

## Gartner IT Glossary

**URL:** <http://www.gartner.com/it-glossary/>

**Description:** “IT Glossary is your trusted guide to exploring technology terms and definitions, from the world’s leading IT research and advisory company.”

**Terms:** 1,000+

**Distinguishing Characteristics:** Glossary for all manner of information technologies, not just open data. Search and alphabetical browsing functionalities. Related research displayed for terms. Term of the Day.

# OPEN DATA RESOURCES

### OPEN DATA CASE STUDY REPOSITORIES (alphabetically listed)



#### 1. ePSIplatform's Cases

**Who:** ePSIplatform

**What:** A collection of examples on Public Sector Information (PSI) Re-use, with the main focus on Europe.

**Number of Case Studies:** 221

**Countries Covered:** 31 – Austria, Belgium, Bulgaria, Canada (2), Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, New Zealand (1), Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom, United States (1)

**Contact:** Phil Archer, W3C; Martin Alvarez-Espinar, CTIC; Daniel Dietrich, OKFN; Makx Dekkers (according to website)

**Case Studies URL:** <http://bit.ly/1UiaGs6>

[illegible]



## 2. Follow the Money

**Who:** A global network of more than 100 organizations supported by ONE and the Open Knowledge Foundation. Other partners include the Transparency & Accountability Initiative, the World Bank, Development Initiatives, and the Revenue Watch Institute.

**What:** The network enables organizations working on “various aspects of the Follow the Money agenda – from resource availability, to spending, to service delivery and results in particular sectors to share information, campaign for the inclusion of open data principles in transparency policies, and identify opportunities for collaboration to drive policy change, including through Follow the Money campaigns in specific countries.”<sup>1</sup> The repository has a summary of each initiative including the stage of the project and its categorization.

**Rationale:** The aim of the initiative is to “support and strengthen the community of activists and advocacy organizations pushing for the transparency that is needed for citizens around the world to ‘follow the money’ and hold decision-makers to account for the use of public money.”<sup>2</sup>

**Number of Case Studies:** 266

**Countries Covered:** Africa, Indonesia, India, USA, Afghanistan, Pakistan, Mexico, Russia, Saudi Arabia, Iran, Venezuela, Kuwait, UAE, Algeria, Libya, Peru, Philippines, Argentina, Brazil, Germany, Bangladesh, Estonia, China, Spain, Costa Rica, Nepal, Chile, Georgia, Romania and South Korea

**Contact:** Alan Hudson, Policy Director (Transparency & Accountability) at ONE and Jonathan Gray, Director of Policy and Ideas at the Open Knowledge Foundation (@jwyg)

**URL:** <http://followthemoney.net/>

**Database:** <http://bit.ly/1RwXSXO>





### 3. Open Data Partnership for Development (ODP4D) Case Studies: Open Data Institute (ODI), The World Bank and Open Knowledge Foundation

**Who:** Partnership between ODI, The World Bank and Open Knowledge Foundation

**What:** As noted in the project announcement, the overall objectives of the partnership are to: “support developing countries to plan, execute and run open data initiatives; to increase re-use of open data in developing countries; to develop an ecosystem of skills, support and peer-learning; and grow the base of evidence on the impact of open data for development.”<sup>3</sup>

Project activities include “scoping the state of open data; assessing the readiness of countries to open up – and use – their data; training government officials, other policy makers, and civil society; undertaking research and producing guidelines on the best use of open data; and producing case studies of impact.”<sup>4</sup>

**Rationale:** “There needs to be a credible evidence base focused on developing countries and socio-economic development in order to both win the support of officials and to better execute future open data initiatives.”<sup>5</sup>

**Number of Case Studies:** 13

**Countries Covered:** Burkina Faso, Mexico, Morocco, Kyrgyzstan, Moldova, Macedonia and South Africa

**Contact:** Tom Heath, Head of Research, ODI; (@tommyh)

**Project URL:** <https://theodi.org/odp4d>

**Case Studies URL:** <https://theodi.org/case-studies>



#### 4. Open Government Partnership Case Studies

**Who:** Open Government Partnerships

**What:** The case studies are submitted by participating governments and give a concrete example of an initiative they have launched featuring open data. The case study provides a brief overview of the project, the challenges faced, results and the opportunities that could be maximized.

**Rationale:** OGP aims to secure concrete commitments from governments to drive open government reform and innovation at the country level, in an effort to stretch countries beyond their current baseline in the areas of transparency, accountability and citizen engagement.

**Number of Case Studies:** 7

**Countries Covered:** Mexico, Israel, Indonesia, Tanzania, Croatia, United States, Moldova, Chile, Brazil, Azerbaijan, Philippines, Brazil, Canada, Albania, Estonia, Greece, Italy, Montenegro, Norway

**Contact:** Munyema Hasan, Program Officer

**Case Studies URL:** <http://bit.ly/1pxanMo>

## 5. Open Knowledge Foundation's Global Initiative for Fiscal Transparency Report (GIFT)

**Who:** Open Knowledge Foundation with funding provided by Omidyar Network

**What:** The report examines two main areas: “technology for transparency and technology for accountability around public finance.” In particular it looks at “who builds these tools, who uses them, and who benefits from them.”<sup>6</sup>

**Rationale:** With an increasing emphasis on opening up government data, “it is prudent for governments to start to look into new ways to open up their data, to save time, money and effort in opening up a backlog of data and to get necessary workflows in place.”<sup>7</sup>

**Number of Case Studies:** 12 (as well as 2 examples of internationally recognized standards)

**Countries Covered:** United Kingdom, Canada, Republic of Korea, Brazil, Sierra Leone, Indonesia, India, Uganda

**Contact:** Rufus Pollock, President & Founder of OKF; @rufuspollock

**URL:** <http://community.openspending.org/research/gift/standards/>



## 6. Sunlight Foundation Case Studies on Social Impact of Open Data

**Who:** The Sunlight Foundation, with funding from the Open Data for Development Research Fund

**What:** The Sunlight Foundation has produced “a searchable repository of more than 100 examples on the outputs, outcomes and impacts of open data and digital technology projects.”<sup>8</sup>

**Rationale:** The case studies “aim to illustrate the social impact of open data and digital transparency initiatives in different countries, cities and communities. We gathered examples on how users are empowered by open government initiatives as well as how open data and digital transparency projects are changing the behavior, relationships, activities and actions of users.”<sup>9</sup> The cases create an “evidence base that can empower further generalizations in the open government field.”<sup>10</sup>

**Number of Case Studies:** 142

**Countries Covered:** Worldwide – 64 cases from the Global North, 37 from the Global South and nine with impact on both regions

**Contact:** Julia Keseru, International Policy Manager; (@jkeserue)

**Project URL:** <https://sunlightfoundation.com/blog/2015/05/05/a-new-approach-to-measuring-the-impact-of-open-data/>

**Case Studies URL:** <http://bit.ly/1EoIXJR>



## 7. Technology for Transparency Network

**Who:** A project of Rising Voices and Global Voices, sponsored by the Open Society Institute and the Omidyar Network

**What:** As noted on the project site, “the Technology for Transparency Network is a research and mapping project that aims to improve understanding of the current state of online technology projects that increase transparency and accountability in Central & Eastern Europe, East Asia, Latin America, the Middle East and North Africa, Sub-Saharan Africa, Southeast Asia, South Asia, and the former Soviet Union.”<sup>11</sup>

**Rationale:** The case studies aim to “document in-depth as many technology for transparency projects as possible to gain a better understanding of their current impact, obstacles, and future potential.”<sup>12</sup>

**Number of Case Studies:** 63

**Countries Covered:** Argentina, Burundi, Brazil, Cambodia, Chile, China, Costa Rica, Czech Republic, Ghana, Guatemala, Hong Kong, Hungary, Israel, India, Jordan, Kenya, Lebanon, Sri Lanka, Mexico, Malaysia, Philippines, Sri Lanka, Puerto Rico, Sudan, Singapore, Slovakia, Thailand, Tanzania, Uganda, Venezuela, Zimbabwe

**Contact:** Renata Avila, Project Leader, @avilarenata; Rebekah Heacock, Project Leader, @rebekahredux and/or Hazel Feigenblatt, Editorial Advisor, @HazelFeigenblatt

**URL:** <http://transparency.globalvoicesonline.org>

**Case Studies URL:** <http://transparency.globalvoicesonline.org/projects/all>



## 8. World Bank's Pipeline Project

**Who:** The World Bank Group

**What:** A collection of examples of “open data companies in emerging markets, specifically in Latin America, Southeast Asia, Africa, as well as India and Russia.”

**Number of Case Studies:** 315

**Countries Covered:** 21 – Argentina, Brazil, Chile, Colombia, Costa Rica, Ghana, India, Indonesia, Kenya, Malaysia, Mexico, Morocco, Nigeria, Paraguay, Philippines, Russia, South Africa, Tanzania, Thailand, Uruguay, Vietnam

**Contact:** Haishan Fu, Tariq Afzal Khokhar, Amparo Ballivian

**Case Studies URL:** <http://bit.ly/1zKVN2z>



## 9. The World Wide Web Foundation (WWWF): Exploring the Emerging Impacts of Open Data in Developing Countries (ODDC)

**Who:** The WWWF, with funding from the International Development Research Centre (IDRC)

**What:** The ODDC Project is a multi-country, multi-year study led by the World Wide Web Foundation. As noted in the project announcement, the ODDC Project “explores how open data can foster improved governance, support citizens’ rights, and promote more inclusive development through looking at the emerging impacts of existing open data projects in developing countries.”<sup>13</sup>

**Rationale:** The ODDC project is “designed to inform the development of planned and on-going open data initiatives in the South. The project will work through a series of open data case studies in Latin America, Africa, and Asia. These case studies examine initiatives, the governance challenges they propose to address, and emerging outcomes and impacts from the application of open data in these contexts. The project is also developing cross-cutting data collection instruments and analysis approaches to help explain if and how open data is bringing change to developing countries. Finally, it is engaging with global and local policymaking and practice in order to improve developmental outcomes of these initiatives.”<sup>14</sup>

**Number of Case Studies:** 17

**Countries Covered:** 14 (Initial case-study countries include Argentina, Brazil, Chile, India, Indonesia, Kenya, Nigeria, Nepal, Paraguay, Philippines, Sierra Leone, South Africa, Uganda and Uruguay)

**Contact:** Tim Davies; @timdavies

**URL:** <http://www.opendataresearch.org/emergingimpacts>



## 10. Open Data 500 Global Network

**Who:** The Governance Lab at New York University

**What:** Network of organizations collecting case studies on open data use and impact by companies.

**Rationale:** “Coordinated by the Governance Lab (GovLab), the OD500 Global Network enables participating organizations to analyze open data in their country in a manner that is both globally comparative and domestically specific. The OD500 Global Network starts from the assumption that only by mapping the use of open data within and across countries, can new approaches for understanding the economic and social impact of open government data be generated.”

**Number of Examples:** 1,100

**Countries Covered:** U.S., Mexico, Australia (Italy and Korea forthcoming)

**Contact:** Stefaan Verhulst, Chief of Research and Development; [stefaan@thegovlab.org](mailto:stefaan@thegovlab.org)

**URL:** <http://www.opendata500.com/>



# OPEN DATA PORTAL REPOSITORIES

Open data portals have been increasingly used by local, city and national governments to release data in open and reusable formats. These portals have encouraged the increased use of open government data and led to the development of applications, new products and services.

Launched at the 2011 Open Knowledge Foundation Conference in Berlin, DataPortals.org curates a list of over 400 open data portals from around the world, including Data.gov and New Zealand's Open Data Case Studies. These portals specifically also provide a collection of case studies on how these portals have been used and contributed to social impact.

## United States – Data.gov Impact Case Studies

<http://www.data.gov/impact/>

In addition to providing access to numerous government data sets, the site also highlights examples of open data impact, including impact in “cost savings, efficiency, fuel for business, improved civic services, informed policy, performance planning, research and scientific discoveries, transparency and accountability, and increased public participation in the democratic dialogue.”

## New Zealand – Open Data Case Studies

<https://www.ict.govt.nz/guidance-and-resources/case-studies/open-data/>

New Zealand's Office of the Chief Information Office provides a list of case studies that highlight “the innovative implementation of government ICT initiatives, the successful contribution to the delivery of ICT functional leadership, and the impact of innovative re-use of open government data.” The portal showcases over 30 open data case studies from 2012 to 2015 across various open data use categories, including apps and tools; new skills; central government releases; statistics; and local government releases.

# INTERMEDIARY REPOSITORIES

**Several Open Data intermediaries have also developed a set of client-studies, including:**

## **CKAN Case Studies**

<http://ckan.org/case-studies/>

<http://bit.ly/22CDLzv>

The CKAN data management platform helps “numerous governments, organisations and communities around the world” catalogue, search and display data. The site showcases 116 instances of use across Europe, Asia, Africa, South America, North America and Australasia as well five in-depth case studies from the U.K., the European Union, Helsinki and the International Aid Transparency Initiative (IATI).

## **DKAN Case Studies**

<https://civicaactions.com/case-studies>

DKAN is the Drupal-based open source data platform that includes a full suite of cataloging, publishing and visualization features that allows governments, nonprofits and universities to easily publish data to the public. The eight case studies featured on the site present an overview of the client, the project, the goal of the work and the impact of the project.

## **Esri Case Studies**

<http://www.esriuk.com/Resources/case-studies>

<http://bit.ly/25j8cgo>

Esri provides enterprise geographic information systems (GIS) to help its clients “make timely, informed and mission-critical decisions by leveraging the power of geography.” Esri provides over 80 case studies on how businesses, government agencies and nonprofits across sectors have used map-based data for decision-making and social impact.

## **Granicus Case Studies**

<http://www.granicus.com/customers/case-studies/>

<http://bit.ly/1praF68>

Granicus manages “the world’s largest and most reliable legislative content network—containing more than 5 million government media files and public records and maintaining a 99.98% uptime track-record.” Granicus provides 30 case studies of cities across the U.S. who are “leveraging their technology to successfully run their day-to-day government information tasks including webcasting public meetings, managing and delivering legislative information, collaborating with citizens, and more.”

### IBM Smarter Cities Case Studies

[http://www.ibm.com/smarterplanet/us/en/smarter\\_cities/article/smarter\\_cities\\_case\\_studies.html](http://www.ibm.com/smarterplanet/us/en/smarter_cities/article/smarter_cities_case_studies.html)

<http://ibm.co/1hxRs9k>

The IBM Smarter Cities projects “help government leaders manage complex city environments, incidents and emergencies with a city solution that delivers operational insights.” Tools from the IBM Intelligent Operations Center (US) offer integrated data visualization, near real-time collaboration and deep analytics to help city agencies enhance the ongoing efficiency of city operations, plan for growth and coordinate and manage response efforts. The aim of the initiative is to help power more collaborative working relationships across businesses, industries and state and town lines – and draw in an ever more active and engaged citizenry. There are five in-depth interactive case studies from cities around the world available with details on what IBM did and how they did it.

### Microsoft Case Studies

<https://customers.microsoft.com/Pages/AdvancedSearch.aspx>

<http://bit.ly/1ZoIKSB>

Microsoft open government solutions “allows governments to reach and interact with citizens when and where they need through secure and effective data sharing and open access across mixed technology environments.” The solutions help to “provide effective information between various government agencies and with constituents while not sacrificing security and privacy.” The diverse case studies set out the business needs of the organization (including governments, nonprofit organization and companies), the solution developed by Microsoft and the benefits of the initiative.

### Socrata Case Studies

<http://www.socrata.com/case-studies/>

The Socrata case studies provide descriptions of Socrata-led initiatives with governments at the municipal (e.g., Boston), state (e.g., Hawaii) and federal levels (e.g., Spain). Each of the 29 case studies provides a description of the initiative, why it was developed and the main highlights of the project.

## APPENDIX IV

# LIST OF INTERVIEWS AND PEER REVIEWERS

## STAKEHOLDER INTERVIEWS

During the development of its case studies on open data's impact, the GovLab interviewed the following stakeholders and experts:

**Neil Ackroyd**, Chief Operating Officer, Ordnance Survey

**Kim Alexander**, President and Founder, California Voter Foundation

**Robert Andrews**, Head of Corporate Communications, Ordnance Survey

**Jay Bhalla**, Director, Open Institute

**Uuf Brajawidagba**, Lecturer of Indonesian Politics, University of Wollongong

**François Brouard**, Founding Director, Sprott Centre for Social Enterprises, Carleton University

**Iain Champion**, former Application Team Leader, Environment Canterbury

**John Carpenter**, Director of Strategy, Ordnance Survey

**Daniel Carranza**, Co-Founder, DATA Uruguay

**Julian Carver**, former Chief Information Officer, Canterbury Earthquake Recovery Authority

**Pablo Clark**, Analyst, Mexican Institute for Competitiveness

**Izabela Corrèa**, Former Coordinator for the Promotion of Ethics, Transparency, and Integrity, Brazil's Directorate for Corruption Prevention

**Jeff de La Beaujardiere**, Data Management Architect, NOAA

**Leodelma de Marilac Felix**, Former Head of the General Auditing Coordination, Ministry of Finance, Brazil

**Vivien Deparday**, Disaster Risk Management Specialist, Open Data for Resilience Initiative Operational Deployment Lead, Global Facility for Disaster Risk and Reduction, World Bank

**Patrick DuFour**, Former Senior Web GIS Developer, U.S. Department of State, Humanitarian Information Unit

**Ee-Peng Lim**, Co-Director of Living Analytics Research Center, Singapore Management University

**Carl Elmstam**, Transparency Manager, Sida

**Peter Elson**, Community Development Researcher, Mount Royal University and University of Victoria

**Felipe Estefan**, Associate, Investments, Omidyar Network

**Aidan Eyakuze**, Executive Director, Twaweza

**Stephen Ferris**, GIS and Data Manager, Canterbury Earthquake Recovery Authority

**Christian Fischer**, Senior Advisor, Septima

**Mike Flowers**, former Chief Analytics Officer, Mayor's Office of Data Analytics, New York City

**Rafael García Aceves**, Open Data Project Coordinator, Transparencia Mexicana

**Zachary Goldstein**, Chief Information Officer, NOAA

**Hanna Hellquist**, Former State Secretary for International Development, Sweden

**Thorhildur Jetzek**, Department of IT Management, Copenhagen Business School

**Al Kags**, Founding Trustee, Open Institute

**Jason Kim**, Senior Advisor, National Coordination Office for Space-Based Positioning, Navigation, and Timing

**Verena Luise Knippel**, Senior Governance Specialist, World Bank

**Daniel Kreiss**, Assistant Professor, University of North Carolina School of Journalism and Media

**David Lasby**, Director of Research, Imagine Canada

**Michael Lenczner**, Chief Executive Officer, Ajah.ca

**Morten Lind**, Senior Advisor, Danish Geodata Agency

**Andrew Loveless**, Commercial Director, Ordnance Survey

**Lindsey Marchessault**, Senior Manager for Data & Engagement, Open Contracting Partnership

**Arnold Minde**, Founder, Shule.info

**Lindsay Mollineaux**, Director of Analytics, Mayor's Office of Data Analytics, New York City

**Oscar Montiel**, Director of Community Engagement, Code for Mexico

**Otávio Moreira de Castro Neves**, Coordinator for Open Government and Transparency, Brazil

**Ainun Najib**, Co-Founder, Kawal Pemilu

**Bitange Ndemo**, Chairman, Open Institute

**Bo Overgaard**, Director, Septima

**Tim Owen**, Chief of Climate Information Services Division, NOAA

**Allan Parnell**, PhD, Vice President, Cedar Grove Institute for Sustainable Communities

**Ed Parsons**, Geospatial Technologist, Google

**Maria Patterson PhD**, Scientific Lead, Open Science Data Cloud, University of Chicago, Open Cloud Consortium

**Hilary Pearson**, President, Philanthropic Foundations Canada

**Karl Peterson**, Project Manager, Aid Transparency, Sida

**Florent Peyre**, Chief Operating Officer and Co-Founder, Placemeter

**Peter Rabley**, Director, Investments, Omidyar Network

**Tara Ramchandani**, Attorney, Relman, Dane, & Colfax PLLC

**Ira Rubinstein**, Research Fellow and Adjunct Professor of Law, New York University

**Fabrizio Scrollini**, Chairman, DATA Uruguay

**Diah Setiawaty**, Program Manager for Election Application Programming Interface, Perludem

**Rupert Simons**, Former Advisor to the African Governance Initiative; CEO, Publish What You Fund

**Singapore National Environment Agency**

**Gabriel Sipos**, Director, Transparency International Slovensko

**Diego Soria**, Ministry of Health, Uruguay

**Allison Soussi-Tanani**, Digital Strategy Lead and Web Committee Co-Chair, NOAA

**Javier Teran**, Statistician, United Nations Office for the Coordination of Humanitarian Affairs, Humanitarian Data Exchange

**Samhir Vasdev**, ICT Sector Unit, World Bank

**Eva Vozárová**, Web & IT Lead, Fair-Play Alliance

**Neil Weston**, Chief Scientist, National Geodetic Survey, NOS, NOAA

**Alyssa Wright**, President of U.S. Board, OpenStreetMap

**Alexandra Zapata Hojel**, Coordinator of Education Projects, Mexican Institute for Competitiveness

## PEER REVIEWERS

**Karin Ahlin. Akrodata**

**Antonio Almansa Morales**, Diputación provincial de Málaga

**Andi Argast**, Open Data Institute; Toronto Node

**Jos Berens**, University of Leiden

**Keitha Booth**, Land Information New Zealand

**J. Albert Bowden II**, Sunlight Foundation

**Mark Cardwell**, USAID

**Corinne Cath**, Oxford Internet Institute

**Emmy Chirchir**, University of Münster

**Rafael García Aceves**, Transparencia Mexicana

**Erik Holmlund**, Alberta Data Partnerships Ltd

**Brendan Kenny**, Blackwell Burke P.A.

**Alessia Lefebure**, Columbia University–  
Alliance Program

**Ulrich Mans**, Leiden University

**Valerie Moye**, Socrata

**Dr. Alina Östling**, Centre for Media Pluralism  
and Media Freedom

**Giuseppe Reale**, University of Catania

**Fathima Rifaa**, The Boeing Company, USA

**Julian Singh**, Open Data Connect

**Rayna Stamboliyska**, RS Strategy/Open MENA

**Julian Tait**, Open Data Manchester

**Jamie Van Ymeren**, Mowat Centre

**Mario Velasco**, Coplade, Oaxaca

**Niki Virani**, City of Houston

**Johanna Walker**, Web Science Institute,  
University of Southampton

**Andrew Weller**, University of Washington

**Ian White**, TBD Co

**Raymond Yee**, unglue.it

**Maria Zuffova**, University of Strathclyde,  
Glasgow

