CHAPTER 3

The Impact of Open Data on Developing Economies

Capturing impact and developing evidence

Many studies of open data are concerned with proving a case—either that open data can spur rapid social transformation (the positive case), or that it has a negligible or harmful effect (the negative case). In truth, the evidence is mixed and emergent; the impact of open data is, in fact, far more ambiguous. Rather than just asking Does open data spur development? we seek in this paper to also ask, How and under what conditions can it work?

To answer this question, we’ve examined a wide and variable range of attempts to provide evidence to develop a plausible “theory of change” that would explain the role of open data in development. Theories of change are important. A recent report from the United Kingdom’s Institute of Development Studies points to “the persistence of poorly articulated theories of change that fail to specify realistic causal pathways at the outset” in relation to transparency initiatives.27 Weak theories of change can lead to a variety of false assumptions and misconceptions when it comes to understanding how open data works; these, in turn, can lead to missed opportunities, spending inefficiencies, and a general failure to live up to open data’s potential.

A review of the literature shows that numerous pathways and theories of change have in fact been proposed. For example, a recent study conducted by IDRC, the World Wide Web Foundation, and the Berkman Klein Center for Internet and Society at Harvard University cites at least thirteen “theories of change,” including open data’s ability to reduce transaction costs, generate new forms of economic growth and prosperity, generate new revenue models, and disrupt traditional business models.28 Others point to the social and environmental benefits of open data. For example, Martin Hilbert draws

attention to the potential of opening geospatial-, education-, and housing-related information. Based on a review of 180 pieces of literature related to Big and Open Data, he concludes—with caveats we discuss further below—that open data does in fact contain true opportunities for development.  

**Open Data in Developing Economies Logic Model**

In what follows, we describe a change theory for open data using the Open Data in Developing Economies Logic Model (Figure 3). This model suggests that:

*Open data (supply), when analyzed and leveraged by both governmental and non-governmental actors (demand), can be used in a variety of ways (actions and outputs), within the parameters established by certain enabling conditions (and disabling factors), to improve government, empower citizens and users, create economic opportunity and/or solve societal problems (impact).*

It is important to reiterate, however, that these positive impacts are always subject to certain local, context-sensitive enabling conditions and disabling factors. While this logic model presents a general outline of how open data can work, Table 1, below, presents a more detailed explanation of how it interacts with various sector-specific opportunities and challenges to create genuine impact on the ground.

The logic model is built around the premise (informed by the case studies and primed for further experimentation and research) that higher impact open data projects are the result of matching the supply to the demand of data actors who can operationalize open data toward specific activities and outputs. These outputs and activities can, in turn, serve a broader and more diverse group of users and objectives.

Having a logic model allows for a more detailed analysis within and across sectors of open data toward providing a number of highly specific lessons about actors and conditions, opportunities and challenges.

---


Examining the Open Data in Developing Economies Logic Framework

The logic model presented in Figure 3 above describes the elements in place across the lifespan of an open data initiative, from the initial input or supply of data through its use and impact, with several enabling conditions and disabling factors influencing impact. We explore each of the elements introduced in the table in more detail below, with attention to the types of impact and the enabling conditions and disabling factors that inform the Periodic Table of Open Data Elements.
Table 1. In-Depth Logic Framework

<table>
<thead>
<tr>
<th>Input (Supply)</th>
<th>Actors (Demand)</th>
<th>Activity</th>
<th>Output</th>
<th>Use/Users</th>
<th>Indicators</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Government Data</td>
<td>NGOs &amp; Interest Groups Researchers and Academia Journalists and Media Outlets Donor Organizations Private Sector–Entrepreneurs and Corporations Government Officials</td>
<td>Data Analysis (Methods) Presentation (Visualization) Aggregation and Commingling (Mashups) Dissemination—toward, for instance: • Benchmarking • Hotspotting • Comparing services and performance • Resource allocation • Diagnosis and assessment</td>
<td>Decision Trees Maps Apps and Platforms Dashboards Process improvements Data-driven journalism Infographics Searchable databases Policies Advocacy Alerts</td>
<td>Other Government Agencies and Officials Nongovernment Actors including, for instance:</td>
<td>The outputs of open data have had several real-world effects, which can be assessed according to a number of indicators, for instance: Accountability • Court cases on corruption Improved service delivery: • Increase in citizens receiving services Increased information sharing: • Collaborations between public entities Enhanced decision-making capacity and choice: • Satisfaction among citizens Social mobilization: • Data-driven advocacy efforts Job creation: • Employment in data-driven fields Frugal innovation: • Data-driven startups Economic growth: • Industry-level growth figures Improved situational awareness: • Time saved responding to emerging situations and crises More expertise and knowledge brought to bear: • Number of disciplines engaged in decision-making processes Targeting interventions and tracking impact: • Assessment of change against baseline indicators</td>
<td>Improving Government • Accountability • Improved service delivery • Increased information sharing Empowering Citizens • Enhanced decision-making capacity and choice • Social mobilization Innovation and Creating Economic Opportunity • Job creation • Frugal innovation • Economic growth Solving Public Problems • Improved situational awareness • More expertise and knowledge brought to bear • Targeting interventions and tracking impact</td>
</tr>
</tbody>
</table>

Enabling Conditions and Disabling Factors
- Problem and Demand Definition: whether and how the problem to be addressed and/or the demand for open data are clearly defined and understood
- Capacity and Culture: whether and how resources, human capital and technological capabilities are sufficiently available and leveraged meaningfully
- Partnerships: whether collaboration within and, especially, across sectors using open data exist
- Risks: whether and how the risks associated with open data are assessed and mitigated;
- Governance: whether and how decisions affecting the use of open data are made in a responsive manner
Input (Supply)

A diversity of data types make up the supply side of open data in developing economies and, as the input, plays a key role in determining the ultimate impact. Data types being made available in developing countries range from information about the planet, such as geospatial and weather information, to information about the workings of government itself, like financial and administrative data. Of importance within a developing country context, involves data that is collected and potentially supplied by international (donor) organizations and civil society (often through crowd-sourced means) that may complement the supply of domestic government data. Much of the focus in the early days of open data (in both developed and developing economies) has fallen on improving the supply side of open data, with the Open Data Charter and Open Government Partnership, for example, pushing government data holders to make certain types of data accessible according to a number of principles and in a standardized way.

Actors (Demand)

As the global open data ecosystem matures, a greater focus is being placed on understanding the demand side of open data—the actors who will make use of the information the governments released. As the supply side of open data continues to improve thanks to international standardization efforts, including the Open Data Charter and advocacy at the national and regional level, the demand side stands to benefit through greater engagement of existing demand-side actors, and the identification of additional stakeholders who could make use of that information. Some of the key yet distinctive segments or constituencies (often with different interests and needs) on the demand side of open data include nongovernmental organizations (NGOs), including not only government watchdog groups but also service providers, researchers and scholars, data-driven journalists, entrepreneurs and businesses, and government officials themselves who benefit from more liquid data that has escaped internal silos.

Activities

The activities enabled by access to open data are in many ways only limited by the imagination and skills of the actors on the demand side of the equation. Some of the most common ways that open data is used include data analysis to uncover new insights, presentation and visualization to make the information more comprehensible, aggregation and commingling of multiple datasets to
gain a more multi-faceted view of an issue, and eventually dissemination of processed open data toward benchmarking efforts, hotspotting (e.g., data-driven crime or healthcare maps), or informing future resource allocation decision making.

Output

Like the activities open data enables, the output of processed open data can take any number of forms depending on the problem or opportunity the data is meant to address and the priorities of the actors on the demand side. Although the output of open data initiatives is often some form of data or technology—such as searchable databases, information dashboards or smartphone applications—they can also take the form of evidence-based policies, advocacy, or activism efforts or data-driven journalism pieces.

Use/Users

The existence of actors on the demand side of open data and users of the open data-driven outputs those actors create complicates the lifespan of open data and makes clear the need for responsiveness and feedback loops on the supply side. In many cases, the types of actors representing the demand side of open data are also present as users of open data outputs—such as NGOs or journalists. The community of user includes a broader swath of the population, however, with individuals and entities that lack any data science capabilities still able to make use of the outputs of open data—whether in the form of smartphone weather applications, data-driven infographics in the newspaper, or government process optimizations.

Indicators

In many ways, open data is a double-edged sword. How open data is made accessible to the public ensures that anyone can use open data for any reason. It also means that identifying those usages and capturing their impacts is extremely challenging, especially for resource-strapped governments. To gain some meaningful sense of the impact of open data releases, data holders can seek to develop indicators tied to the problems open datasets stand to address. After the release of open data on the financial dealings of government officials, for instance, an uptick in court cases on corruption could provide a window into open data’s impact on accountability. Similarly, the creation of more data-driven startups, increased investment from international donors, or increases in hiring among technology companies that use (or are likely to use) open
data can act as indicators of open data’s effect on economic development. Especially in developing economies where government resources are often limited, meaningfully capturing the impacts of open data through indicators of success will likely prove essential for maintaining the political will needed for open data efforts to be sustainable.

**Types of Impact**

As reflected in Table 1, our research indicates that open data has four main types of impact, and that each type of impact requires different indicators. Although the four types of impact described below provide a framework of analysis, it is also important to understand that these types of impact can manifest in different ways, and some projects might seek to achieve more than one type of impact. In the discussion below, we point to a diversity of open data initiatives aimed at having a positive effect in one or more of these impact areas, but in some cases, impact remains largely aspirational; in others, impact was negligible, or in fact, negative. Rather than focusing exclusively on gold-standard open data projects with unquestionable and consequential on-the-ground impacts—like the oft-referenced Global Positioning System or opening of weather data in the United States—we examine initiatives across the spectrum of impact to develop a more detailed understanding of the current reality for open data in developing economies, and more importantly, to provide testable premises of how to create an impact based on lessons learned from efforts to date, even some efforts that have not (yet) created major positive impacts.

To be sure, much of the evidence provided below is emergent, and in some cases largely speculative. Collecting and organizing these signals of what is known (and believed) about open data for development, however, provides for a systematic understanding of the current field, and informs more strategic, analytical assessments of open data’s impact going forward. So although the evidence here is unquestionably variable—ranging from concrete, clearly demonstrable on-the-ground impacts to largely ideological assertions of impact—they provide a frame for understanding the field and taking the next step toward meaningful impact assessment.

*Improving governance*

One of the most consistent ways in which open data has an impact on development, across countries and regions, is in *improving governance*. This impact manifests in several ways:
• Greater transparency and citizen involvement can make governments more accountable to their citizens.
• A focus on data use and data-driven decision making engendered by the institutional process of opening data (i.e., cleaning and making liquid government datasets) can result in better and more efficient service delivery.
• In addition to making data accessible to entities outside of government, open data efforts can increase information sharing between departments and agencies within government, improving coordination and knowledge-sharing.

Emergent evidence:

• **Elections in Burkina Faso:** To ensure elections in Burkina Faso were conducted fairly, poll results were made available in real time via an official election website, which tracked candidates leading in each of the provinces. This project, run by the Burkina Open Data Initiative (BODI, http://data.gov.bf/about) with the support of the ODI, sought to promote democracy and trust between Burkina Faso’s citizens and elected officials. For a country in transition like Burkina Faso, opening electoral data was seen as an important first step toward establishing longer term political stability and citizen trust in the electoral process, though the number of citizens or organizations who actually accessed and acted upon the data is unclear.31

• **Elections in Indonesia:** In a similar initiative, Indonesia’s Kawal Pemilu (“guard the election,” in Indonesian) was launched in the immediate aftermath of the 2014 presidential elections, as the country was riven by political polarization and the two leading contenders for the presidency traded allegations of vote rigging. A globally dispersed group of technologists, activists, and volunteers came together to create a website that would allow citizens to compare official vote tallies with the original tabulations from polling stations. These tabulations were already made public as part of the Elections General Commission’s commitment to openness and transparency. Kawal Pemilu’s organizers, however, played a critical role in assembling a team of over 700 volunteers to digitize the often-handwritten forms and make the data more legible and accessible.

31 The author of a case study from the participating Open Data Institute, which disseminated the results of the effort, notes that, “Whether or not the improved information flow and accessibility to results—enabled by the application of open data—led to increased trust in the election process, or even an improved process, is not something that we can or indeed set out to prove or show empirically, although the outcome of this case study is certainly supportive of that conclusion.” Anna Scott, “Case Study: Burkina Faso’s open elections,” Open Data Institute, October 2016, https://theodi.org/case-study-burkina-fasos-open-elections#1.
The site was assembled in a mere two days, with a total budget of just $54. Not only did the site enable citizen participation in monitoring the election results, but Kawal Pemilu’s vote tallies also played an important role in court hearings confirming the election winner.\(^{32}\)

- **Data Journalism in Kenya:** In Kenya, journalists leveraged open data to report on a “freeze” in the dissemination of welfare support to the elderly and disabled. The freeze was traced back to a government failure to build an effective system for distributing such funds and, as a result, a significant amount of public money went missing. Media attention and public pressure that grew out of this open data-driven journalism effort led to an audit of the program and the implementation of reforms.\(^{33}\)

*Empowering citizens*

Open data also has a powerful role to play in *empowering citizens*. This role is evident in several ways:

- With more access to information in hand (including information on, for example, health care or education choices), citizens can have improved decision-making capacity and choice.
- As a result of increased transparency, open data can act as a *social mobilization* tool when information made available to the public can inform advocacy efforts, including those related to corruption or perceived corruption, consumer advocacy, or health care and other service delivery.

*Emergent evidence:*

- **Follow the Money Nigeria:** In Nigeria, a consortium of activists, journalists, researchers and NGOs use open data to track and visualize government expenditures. Based on knowledge drawn from open data regarding current spending practices, the group successfully pushed the Nigerian government to allocate $5.3 million to help address a lead poisoning crisis in the village of Bagaga that affected thousands of children. Follow the Money Nigeria in fact demonstrates how open data can both improve governance (as a result of enabling better, more evidence-
based policy decisions) and empower citizens to have an impact on their communities.34

- **Seeking to Improve Voter Turnout in Kenya with Open Data:** Kenya’s national Independent Electoral and Boundaries Commission (IEBC) released information about polling center locations on its website in the lead up to Kenya’s 2013 general election. The information, however, was difficult to access and reuse. Seizing on the gap between opening government data and citizens’ actual ability to use that data, two Code 4 Kenya fellows conducted an experiment in unlocking government data to make it useful to the public. The fellows scraped the released IEBC data and built a simple website where it could be more easily accessed. The result was the initial version of GotToVote!, a site that provided citizens with voter registration center information, and also helped them navigate the sometimes complex world of registration procedures. This first version was developed in just 24 hours at minimum cost, garnered over 6,000 site visits in just its first week of existence, and has since been replicated across sub-Saharan Africa.

- **Social Movements in Brazil:** The availability of open data has helped to inform the community organizing and advocacy efforts of several social movements in Brazil. Efforts to fight the use of pesticides latched onto the fact that “each Brazilian citizen is exposed to [5.2 litres] of pesticides every year.” Similarly, an effort to fight school closures has rallied around the 24,000 schools that open data shows have been closed over the last 10 years. Efforts to fight violence against women and the consolidation of land ownership are similarly using open data to aid in their advocacy. There’s little causal evidence that these efforts directly created significant policy impact, however, though the resignation of the Minister of Promotion of Racial Equality is credited to open data-driven reporting and advocacy.35 So although these advocacy efforts demonstrate how open data can empower citizens and advocates through access to important factual information, more work needs to be done (and responsiveness in the public sector must be engendered) to create more direct, tangible impacts.36

---

Creating economic activity

Under the right conditions, open data can help create economic activity. If harnessed properly, this is a particularly important form of impact in developing economies. Our case studies indicate that open data can have a positive impact on economic activity in the following ways:

- As the global economy becomes increasingly reliant on data and information, the accessibility of open data can enable business creation, foreign investment, and meaningful job creation.
- Open data is increasingly seen as a new business asset but, unlike many such assets, it is available free of charge, opening the door to more frugal innovation efforts in the private sector.
- More than just an asset to individual businesses or entrepreneurs, many predictive analyses have pointed to open data’s potential for creating more systemic and far-reaching economic growth, particularly when commingled with proprietary data held by private sector entities.37

Emergent evidence:

- Market Research in Kenya and Nigeria: Sagaci Research is a market intelligence firm based in Kenya that works across countries in Africa. The firm’s strategic knowledge offerings—spanning sectors like consumer goods, agriculture, and telecom—are built from researchers and field surveyors active across Africa and, importantly, open census and national statistical data from the Kenyan and Nigerian governments. According to the Sagaci website, 90 percent of its clients have pursued follow-on work with the firm, demonstrating the value of its open data-driven offerings.38
- Data Mapping Consultation in India: Excel Geomatics is a private consultancy firm that leverages open data to provide geospatial insights to private and public sector clients. The company’s offerings—including ward maps of more than 700 towns and cities and satellite image-enabled population distribution maps—would not be possible without access to data from the Indian census, as well as publicly accessible village and

district boundary maps. Importantly, Excel Geomatics uses the Earth Observing System Data and Information System (EOSDIS) and ASTER database from NASA for its products and services—demonstrating how the opening of data in developed countries often creates impacts far afield.  

- **Open Data to Benefit Tourism in Jamaica**: Like much of the Caribbean, the Jamaican economy is strongly dependent on the health of its tourism industry. Influenced by the rise of all-inclusive resorts and a general disincentive for tourists to stray far from a few highly trafficked areas, tourists rarely experience much of Jamaica’s unique culture, and the economic benefits of tourism are often concentrated in a few areas. To increase tourism, spread its positive impacts and provide useful skills to citizens, a community mapping project\(^{40}\) combined open government data with crowdsourced, volunteer-collected mapping data to enable the more participatory development of the tourism sector. Built around open tourism data and the engagement of government agencies, civil society organizations, developers, and an interested group of community mappers, the initiative has created new artefacts aimed at better spreading the economic impacts of the tourism industry in Jamaica, though impact remains primarily aspirational.\(^{41}\)

**Solving public problems**

Finally, open data’s impact is evident in the contribution it makes to solving public problems. Open data can help address complex problems in the following ways:

- Especially in crisis situations where geospatial information can prove essential,\(^{42}\) open data can play a role in improving situational awareness.
- In some developing countries, government is the primary data holder and data user, limiting the number of people capable of creating value with data. The accessibility of open data can help to bring a wider range of expertise and knowledge to bear on public problems.

---

40 See http://icm.msbm-uwi.org/.
41 Empowering Local Communities with Open Data and Interactive Community Mapping,” Caribbean Open Institute, http://caribbeanopeninstitute.org/node/133.
In many cases the result of improved situational awareness and more expertise brought to bear, open data can play an important role in targeting interventions and meaningfully tracking impact.

Emergent evidence:

- **Stopping Deforestation in Brazil and Indonesia**: To monitor deforestation in Brazil and Indonesia, Global Forest Watch consolidates satellite imagery datasets to monitor global deforestation in real time. Monitoring on this scale has produced several observable positive effects. For instance, data from the project has been used in legal proceedings related to illegal logging. Although causation cannot be proved directly, deforestation has declined in both countries—deforestation in Indonesia is at its lowest levels in a decade and has declined by 18 percent in Brazil. It is important to note that prior to this project, deforestation levels were consistently rising in both nations. The Indonesian government also uses GFW to monitor forest and peat fires and target response. In Brazil, firefighters have reduced their response time to forest fires from 36 hours to 4 hours. This project is a forceful demonstration that intelligent use of open data can be used for successful advocacy—and could even provide additional benefits that may not have been anticipated.

- **Fighting Ebola in Sierra Leone**: In the parts of West Africa affected by the Ebola epidemic, roads, village names, and villages were missing on many online maps. OpenStreetMap (OSM), a free, crowdsourced mapping tool provided critical mapping information to Sierra Leone’s National Ebola Response Centre (NERC), the United Nation’s Humanitarian Data Exchange, and to the Ebola GeoNode to assist them in coordinating public health strategies in response to the epidemic. The OSM data was then often mashed up with open data from affected governments and international organizations. Although the direct impact of open data in the Ebola response was difficult to empirically measure, those working on the ground during the response made clear that providing missing data in open formats played an important role in fighting a complex epidemic and coordinating relief efforts of those working in a chaotic, fast-developing context.
• **Targeting Disaster Risk Funding in the Philippines:** The Philippines was one of the eight founding members of the Open Government Partnership launched in 2011, endorsing an Open Government Declaration to commit to open data. As part of its commitment, in 2014 the Philippine government launched data.gov.ph, which publishes data from government agencies for the public to access. Though some federal agencies have been hesitant to disclose their data in an open and accessible way, at the local government level the open data initiative is making more headway. The disclosure of spending data in Bohol province, for example, allowed civil society groups to notice that insufficient funds were allocated to disaster risk reduction projects. As a result, organizations are now drafting new disaster-reduction proposals to lobby the government to provide more support to this area.46

• **PakReport Crisis Mapping:** In the wake of the worst flooding in Pakistan in decades, several crisis mapping organizations (led by PakReport) teamed with relief organizations to map affected areas in real time. This project was a piece of a broader trend toward crisis mapping, particularly after natural disasters. Collaboration of this kind allows aid organizations to survey areas that may be difficult to landscape because of the disaster and correctly understand where needs are greatest and what kind of assistance populations across the affected area require. The efforts of the PakReport team demonstrate the complicated nature of international disaster relief and the need for comprehensive and proactive data responsibility assessments. Pointing to the risks and unintended consequences of open data, PakReport was forced to restrict access to its crowdsourced maps, which were intended to be open and freely accessible, after the Taliban threatened to attack foreign aid workers in the country, whose presence they deemed “unacceptable.”47

**Enabling Conditions and Disabling Factors**

Based on the existing literature and case studies, we have developed a Periodic Table of Open Data Elements (Table 2) detailing the enabling conditions and disabling factors that often determine the impact of open data initiatives.

---


Although the importance of local variation and context is, of course, paramount, current research and practice shows that the elements included in five central issue categories—Problem and Demand Definition, Capacity and Culture, Partnerships, Risks, Governance—are likely to either enable or disrupt the success of open data projects when replicated across countries.

As discussed above, there is a large variability as it relates to evidence of open data's impact, so we provide these enabling conditions and disabling factors not as a concrete, certain drivers of success or failure, but as an aggregated set of premises to be tested as the field of practice and research of open data in developing economies continues to expand and mature. We examine these enabling conditions and disabling factors in more detail below.
Problem and Demand Definition

Particularly in developing economies, where resources to put toward data release or data use can be in short supply, a clear, detailed understanding of the problem to be addressed by open data can help to ensure that efforts are targeted and optimized. Some of the most effective open data projects examined here are laser-focused on a specific user group (e.g., smallholder farmers in Colombia or Ghana), or identified gap (e.g., the lack of power quality in the Indian energy sector). Clearly defining the problem can also aid in the development of metrics of success and a strategy for monitoring progress against a well-defined baseline. Many of the initiatives studied as part of this project lacked such a monitoring strategy, making assessments of impact, evidence-driven iteration, and the demonstration of return on investment more challenging.

Users and Audiences

Open data initiatives tend to be more successful and avoid the notion of, “if you build it, will they come,” when they are clearly optimized for an intended audience or user base from the start. The upfront identification, mapping and understanding of relevant constituencies, and a similar examination of their needs can enable more targeted open data-driven interventions.

Causes and Context

In many open data initiatives, and in governance innovation efforts more generally, practitioners can find themselves addressing symptoms rather than the root causes of problems. Open data projects, such as the effort to predict dengue outbreaks in Paraguay, tend to be more successful when they seek to address underlying problems (mosquito breeding and transmission) rather than the symptoms of those problems (high levels of dengue fever).

Refinement

To move from a well-understood problem area, to a granular, actionable, and quantifiable path forward, successful practitioners often look to refine their understanding of the problem to be addressed by seeking to understand, for instance, why the problem exists in its current form, what contributing factors could be at play, what potential knock-on effects of addressing the problem might be, and why the problem has not yet been solved by some other interested party.
**Bg Benefits and Goals**
Open data projects often fail to build an audience or continue to evolve and expand successfully over time if they do not successfully define the intended benefits of the open data use and set clear target goals. These deficiencies often can create difficulty in the development of metrics and indicators—important drivers of iteration and impact.

Many of the projects studied, including notably Kenya’s GotToVote! project did not have a clear baseline against which to measure the success of the project. Without an understanding of the current baseline, measuring progress toward identified goals and demonstrating whether and how open data efforts actually benefited the public remains a challenge.

**Da Data Audit and Inventory**
Once the problem and value proposition are in place, practitioners are able to explore the availability of datasets, both in the form of open government data, and from other potentially useful and relevant data sources, like NGOs, the private sector, or crowdsourcing efforts. A clear problem definition can help to uncover which data sources could add value and inform strategies for collecting or accessing that data. Colombia’s Aclímate Colombia, for instance, identified the types of data it needed for its agriculture algorithms and engaged the semi-public industry groups that had it. The Prayas Energy Group in India, on the other hand, found that no one collected or stored the type of energy usage information it needed for its power quality monitoring efforts, so it launched its own (open) data collection effort across 18 Indian states.

**Capacity and Culture**
The lack of available resources, insufficient human capital and immature technological capabilities can create major challenges to achieving meaningful impact with open data projects. These challenges can exist both within a country’s open data ecosystem—that is, the capacity of government, civil society, tech community, and the general public—as well as within the actors on the demand side using open data toward certain objectives and the donor organizations funding them.

*Open Data Ecosystem Elements*

**Di Data Infrastructure**
On the supply side of open data the lack of a strong data infrastructure—that is, hardware and software platforms to make data consistently accessible
and machine-readable in a timely manner—often creates major challenges to positive impact.

Burundi’s OpenRBF platform is an example of working around issues related to data infrastructure. Burundi provided access to data on its results-based financing efforts around healthcare through the OpenRBF platform, a digital infrastructure for collecting and publishing such data. The existence of an “out-of-the-box” tool for making results-based funding (RBF) data public in reusable formats catalyzed the widespread opening of RBF data across many developing countries in Africa.

Public Infrastructure
Similar to the ICT4D environment, much of the literature and practice of open data in developing economies points to the importance of a strong public infrastructure—human capital (including data science and statistical knowledge), public services (including education and libraries), and civil society—to ensure that data is collected, cleaned, and released in a usable manner and that updates and feedback are seamlessly incorporated into open datasets. Supply side efforts to leverage these public infrastructures can increase the demand for open data and establish touchpoints with users.

An active ecosystem of data users and international open mapping platforms and individuals helped to ensure that Nepal’s open data-driven crisis response efforts could be quickly developed and put into practice. The challenges experienced by Ghana’s Esoko platform as a result of unreliable electricity access in the country, on the other hand, shows the many ways that public infrastructure can affect the success of open data projects.

Tech Literacy and Internet Penetration
Even as access to the Internet continues to expand across the developing world, especially through smartphones and other portable devices, many open data projects are being launched into communities that suffer from low Internet penetration and a persistent digital divide. Several of the initiatives studied struggled to achieve their transformative potential, particularly when practitioners failed to engage intermediaries or civil society groups capable of reaching unconnected audiences.

Stakeholders involved in South Africa’s Medicine Price Registry Application (MPRApp) and Tanzania’s open education dashboards pointed to low Internet penetration rates, and the related challenge of low tech literacy, as major barriers they confronted to achieving greater positive impacts.

---

**Cultural/ Institutional Roadblocks**

As is often the case in developed countries, too, cultural and institutional roadblocks can limit the impact of open data. These roadblocks can manifest in the form of an institutional culture that remains skeptical of openness, or the absence of well-trained individuals and professionals capable of recognizing and acting on the potential of open data (readiness)—beyond the prevalence of engaging volunteers in the development of open data initiatives. In all cases, a more concerted culture- and capacity-building effort is often necessary to create an impact.

In Burundi, for instance, efforts to create transparency and accountability around its results-based financing efforts were slowed and complicated by a lack of readiness for technology-enabled openness within key institutions. Jamaica’s open data tourism efforts relied on the readiness of outside volunteers to supplement open data through crowdsourcing—with the impact of the project dependent on their capacity to collect data and information in a strategic, usable manner.

**Open Data User/Donor Elements**

**Skills and Expertise**

Especially for more technical uses of open data—such as sophisticated data analytics—actors on the demand side of open data need to possess certain skills and expertise. Employees at CIAT, the organization behind Aclimate Colombia, for instance, possess high-level data science capabilities that enabled them to leverage open data to create sophisticated algorithmic tools to inform agricultural decision making. Other projects, like crowdsourcing efforts from Jamaica and Nepal, relied on the skills of a few important institutional actors on the demand side and the less-technical efforts of volunteer data collectors.

**Feedback Loops**

Open data initiatives tend to be less successful when they do not create mechanisms for users and beneficiaries to provide input to demand-side practitioners. Tanzania’s open education dashboards are a notable example. The platforms were launched into an environment with low Internet penetration and digital literacy, with seemingly little opportunity for the intended users and beneficiaries of the tools, like parents or education advocates, to suggest ways to make the platforms more usable (and useful) for the community.
**Rs Resource Availability and Sustainability**

The availability of funding and resources are a key variable of success on both the supply and demand sides of open data. Focusing on the demand side, although many open data projects can be stood up quickly on a tight budget (such as Kenya’s GotToVote! an initial prototype of which was created for only $500), sometimes with a very small team (Paraguay’s dengue prediction efforts were championed by researcher Juan Pane and a small team under his direction), establishing sustainability and scaling use often requires more sustained funding and/or well-defined business models. This was the dynamic at work for example in South Africa, where the MPRApp relied almost entirely on the time and effort of a single person. Likewise, in Uganda, CIPESA, the developers of the iParticipate open health data and citizen engagement effort, struggled to proactively elevate health service delivery concerns to relevant government officials because of funding issues affecting both data collection and outreach efforts.

The agriculture information tool Esoko, on the other hand, has managed to take hold in Ghana in large part due to its for-profit, largely business to business (B2B) model, as well as significant investments from foundations and international organizations.

**Partnerships**

In many high-impact open data projects, partnerships within and especially across sectors play a key role in enabling success. Whether creating touchpoints with citizens through partnerships with civil society, informing the public through media partnerships, or filling important data gaps through partnerships with private sector entities, open data suppliers and users often improve outcomes through collaboration.

**Dh Data Holders**

Although open data is meant to provide value to data users without any direct engagement with data holders necessary, partnering with entities on the supply side (including government) can help to fill data gaps and enable higher impact data use.

Aclimate Colombia is a strong example of the potential of such partnerships. The initiative, aimed at providing farmers with a better ability to plant crops in a way that is resilient to the effects of climate change, would not be possible without collaboration between the driver of the initiative (a civil society organization), key data holders (government ministries and agencies), and
a second group of key data holders (private and semi-private crop growers’ associations). GotToVote! in Kenya, on the other hand, did not establish such cross-sector partnerships, and its long-term sustainability is now in question.

Intermediaries

In many developing economies, as mentioned above, Internet penetration and, especially, data literacy are low among the citizenry. The presence of intermediaries—including journalists and others with relevant skills—can help to determine whether or not the available open data-driven outputs reach a community of users, and the intended impact is achieved. The continued advancement of open data intermediaries can be seen as a key area of capacity building in developing economies.

To encourage the use of Code for South Africa’s MPRApp, doctors and pharmacists played an important intermediation role with citizens. These trusted advisors—with nothing to gain from helping patients spend less money on their prescriptions—helped to alert citizens to the database and the potential for identifying much cheaper generic drugs to treat their ailments.

In addition, the open data-driven offerings of Open Development Cambodia are often presented on the initiative’s website in a comprehensible manner (similar to data-driven Wikipedia articles on topics of public concern, like forest cover or development aid spending), but reach a much wider audience when taken up by journalists in the country and abroad in reporting on conditions in the country.

Both of Tanzania’s open education dashboards, on the other hand, failed to attract a regular user base, likely as a result of a failure to engage consistently with intermediaries that could make the sites’ offerings useful to an intended audience with low digital literacy and access.

Domain Experts

In many cases, demand-side open data actors’ expertise lies in technology or data science rather than the problem areas they seek to address through the use of open data. Tapping into the knowledge of stakeholders with relevant sector-specific expertise can improve efforts to optimize and target open data efforts based on a true understanding of needs, opportunities, and barriers. Nepali NGOs and businesses using open government data and crowdsourced data during the response to a major earthquake in the country, for instance, engaged with on-the-ground experts in crisis response who came to Nepal from around the world to help target its offerings.

Co Collaborators

Open data practitioners can extend their capacity by collaborating with like-minded organizations, institutions, or individuals, including foreign actors. Ghana’s Esoko agricultural information service, for example, is part of the Global Open Data for Agriculture and Nutrition (GODAN) network, enabling the company to tap into the knowledge of similar organizations from around the world seeking to leverage open agriculture data for business development and/or public benefit.

Risks

The release and use of open data in developing economies are not without risks. An upfront mapping and consideration of risks associated with intended uses of open data can allow practitioners to design programs from the outset in a way that is well-positioned to overcome or mitigate those risks. The risks listed here, however, should not be considered arguments against using open data in development. Rather, they are reasons for taking a more fine-grained approach that pays close attention to the empirical evidence, sifting out what works and what does not, and identifying conditions for scaling and replication.

Privacy Concerns

Privacy concerns probably rank among the most commonly cited worries over opening up data. Especially in conflict-stricken regions, individuals’ anonymity can be of life-or-death importance. Potential privacy harms can arise even from the release of ostensibly anonymized personally identifiable information (PII). 50 Although the vast majority of open data efforts seek to anonymize or otherwise limit the release of PII, it is important to recognize that a lack of sophistication in anonymization or aggregation efforts can result in the inadvertent release of sensitive information. 51 In addition, in some instances information that itself poses no privacy concerns can be combined with other openly available datasets; the aggregated and linked information can lead to unexpected disclosure of personal data, such as bringing together open data on political activities with separately accessible information on a person’s location or place of work, for example. 52

**Ds Data Security**

Because much government data contains sensitive information regarding individuals, industries, and national security, opening that data often leads to quite reasonable questions about data security. Cybersecurity remains a challenge across the world, and perhaps especially so in developing countries, which may lack the technical expertise to adequately protect information from sophisticated hackers and other intrusions.\(^{53}\) At the same time, though security concerns are very real and important, they must be balanced against the opportunity cost or risk of not sharing data; often, government decision makers can lean on tenuous security concerns to justify keeping data closed and restricting access, potentially limiting the solution space.

**Dm Poor decision-making due to faulty information**

Whether related to humanitarian efforts, crisis relief, or the livelihoods of vulnerable populations, data-driven efforts in developing economies can be literally life-or-death affairs. Given the many challenges and obstacles involved in open data projects, it is important to recognize the risks inherent in basing such life-and-death decisions on information that could be incomplete, out-of-date or otherwise faulty. The broader point is this: insights generated from data are only as good—and their impacts only as positive—as the quality of the underlying data.\(^{54}\)

**Pa Entrenching (power) asymmetries**

Although data can be empowering, it can also consolidate or reinforce existing privileges and authority inherent in societies. This problem is closely linked (though not restricted) to digital divide challenges; when only the elite of a society have access to data and/or data science capabilities, releasing data is likely to disproportionally benefit that elite.\(^{55}\) We found numerous examples,\(^{56}\) and they are important reminders that open data projects need to work hard to ensure that their social and economic benefits are widely, and evenly, distributed.

---


56 See odimpact.org.
Open washing

The term “open washing” has taken hold in practitioner circles over recent years describing the risk that governments may seek to leverage the enthusiasm for open data to avoid more difficult and potentially transformative openness and transparency efforts. The Extractives Industries Transparency Initiative, for instance, is a laudable effort to push for more energy-related openness around the world, which has had demonstrable impacts on accountability. There is a growing belief, however, that a subset of still largely closed governments is joining the initiative only “in order to increase their international reputation and bolster their access to foreign aid.”

Governance

A diversity of governing decisions affect the use and impact of open data efforts. Issues of governance exist at both the ecosystem level—especially related to standards and policies of data release—and on the demand side, with questions of risk mitigation and impact assessment leading the way.

Open Data Ecosystem Elements

Open by Default (and other principles)

Given the level of government resource allocation and time investment required to implement strong open data initiatives, high-level political buy-in and codified open data policies (reflecting the International Open Data Charter principles) are needed to provide the incentives and flexibility to government officials to meaningfully advance open data goals.

The ESMI effort in India, for example, is an industry- and NGO-driven effort to create and open useful data on power quality in the country. This effort, which has had relatively little discernible impact to date, is only necessary because of the lack of energy data being opened by government—an issue

---


that could be resolved with a commitment to openness by default and other internationally accepted principles.

**Fi Freedom of Information and other Policies**
Clear policies pushing forward access to information and data can act as important drivers for open data initiatives. Without explicit policy backing, the sustainability of open data efforts can be called into question, and access to necessary data can dry up at any time. The existence of Freedom of Information policies can also provide means for accessing relevant information, though often at a much slower pace than open data.

A key enabler for the MPRApp open data initiative, for example, was South Africa’s legislative framework that promotes and enacts transparency in medicine pricing. Such a framework compels the Department of Health to collect and publish data on medicine prices in South Africa, ensuring that the supply side of the MPRApp will continue to be made accessible, allowing Code for South Africa to focus on improving the tool and getting it into the hands of its intended users.

**Dq Data Quality**
A widely prevalent challenge to positive impact arises from poor data quality. Data quality is an issue in developed countries, but often presents even greater barriers to success in developing countries. Quality issues can manifest in a number of ways, like inaccurate information, a lack of completeness in official datasets, out-of-date data, or otherwise corrupted datasets.

Aclímate Colombia, for example, experienced challenges gaining access to the most complete and up-to-date information sets for its agriculture tools, slowing their development. Open Development Cambodia’s efforts are consistently challenged by not only strong restrictions in terms redistribution, reproduction, and reuse on some datasets, but also by the inconsistency and unpredictability of when updates to important official datasets occur.

In South Africa, the MPRApp was hurt by a lack of interoperability; that is, open data was not made available in standards that allowed for aggregation and manipulation. Likewise, Kenya’s GotToVote! experienced major challenges when one of its central data sources crashed unexpectedly, rendering the platform temporarily unusable.

**R Responsiveness**
Just as open data is unlikely to create a major impact without demand-side actors to act upon released data, a lack of responsiveness, often characterized by a lack of commitment to take up data-driven insights within governing institutions, can limit the impact of open data. Often, governments succumb to
the temptation to open wash data, nominally opening it up but failing to create feedback loops to ensure that users are actually using the data or that data is being released to meet a genuine demand.

In Jamaica, for example, an interactive community mapping project is supplementing open datasets with a crowdsourced effort to improve tourism in the country; the project’s clear potential has not yielded major impacts yet in part because tourism authorities have not yet acted on the generated insights. The researchers who used open data to predict dengue fever transmission in Paraguay also experienced ongoing challenges wrestling the most useful data for their algorithms from government data holders; there has been little indication that their insights will be meaningfully taken up by institutional authorities.

Open Data User/Donor Agency Elements

**M** Performance Metrics

Open data projects are better positioned for success when practitioners develop and monitor metrics of impact to inform management and iteration.

The vast majority of the open data initiatives studied in this series lacked clearly defined performance metrics. Not only does this create major challenges for iterating upon early efforts, it calls the sustainability of these interventions into question, with a demonstration of success and impact a likely requirement for continued funding and investment.

**Rm** Risk Mitigation

In some cases, open data projects can be advanced despite some level of risk. In such cases, practitioners must ensure that projects that deal in information that is potentially personally identifiable (including anonymized data) have outlined and implemented a clear, upfront strategy for addressing risks created by open data use.

Many of the projects studied in this series dealt in potentially sensitive information—e.g., health, energy consumption, political, and education data. Although each project took steps to ensure that no personally identifiable information was released to the public, all would benefit from a clearly defined—and, preferably, openly available—risk mitigation strategy to ensure that no harms inadvertently fall on data subjects.

The Challenge of Scaling and Replication

Much of this paper, including the above, seeks to identify cross-cutting lessons for open data projects—either in the form of opportunities or challenges.
As noted, however, it is important to keep in mind the diversity that exists within the broad category of “developing economies.” Differences in culture, economic, and political environments, as well as many other variables, can have a profound effect on the success or otherwise of open data projects.

In many ways, this is another challenge facing stakeholders—and perhaps the most intractable: the difficulty of finding an appropriate balance between universal lessons and certain, locally embedded conditions, when seeking to scale and replicate open data projects. The preceding discussion and the sector-specific examples detailed in Section III do suggest that certain enabling (and disabling) conditions have wide applicability—e.g., the need to include intermediaries and civil society groups, or the paramount importance of capacity and resources.

Perhaps the most critical element for scaling and replication found in the Periodic Table is Metrics: the need for open data projects should be evidence based, with clearly defined metrics and standards to evaluate performance. Only with those metrics in place can the applicability or appropriateness of lessons or principles be determined—and only then can the success or failure of projects be established. When making any funding or design decisions, it is essential to take a fine-grained approach that pays close attention to the empirical evidence, sifting out what works and what does not. That is a key goal of the 12 case studies that accompany this landscaping paper; we have tried to build a ground-up, highly empirical picture of open data projects in the developing world.

---