



www.odimpact.org

OPEN DATA IN DEVELOPING ECONOMIES

Toward Building an Evidence Base
on What Works and How

Stefaan G. Verhulst and Andrew Young
The GovLab

JULY 2017



mobile Solutions
Technical Assistance
and Research

fhi360
THE SCIENCE OF IMPROVING LIVES

GOVLAB

www.odimpact.org

OPEN DATA IN DEVELOPING ECONOMIES*

Toward Building an Evidence Base
on What Works and How

Stefaan G. Verhulst and Andrew Young**

JULY 2017



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

* This document was written under the Mobile Solutions Technical Assistance and Research (mSTAR) project, United States Agency for International Development Cooperative Agreement No. AID-OAA-A-12-0073. The content and views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

** Special thanks to Akash Kapur who provided crucial editorial support for this case study, and to the peer reviewers who provided input on a pre-published draft.

ABOUT THE GOVLAB

The Governance Lab (GovLab) is an action research center based at New York University. Our work seeks to build evidence on how to strengthen the ability of institutions and people to work more openly, collaboratively, effectively, and legitimately to make better decisions and solve public problems. GovLab studies and curates new findings relevant to institutional innovation, and develops and tests new methodologies for institutions to solve problems in a more open and collaborative manner. To increase the innovation capacity of institutions, GovLab also trains public entrepreneurs in new approaches to tackle societal challenges, and builds networks of innovators to help match the supply of expertise present in society to demand in the form of public problems.

ABOUT THE AUTHORS

Stefaan G. Verhulst is Co-Founder and Chief Research and Development Officer of the GovLab where he is building a research foundation on how to transform governance using advances in science and technology. @sverhulst

Andrew Young is the Knowledge Director at The GovLab, where he leads research efforts focusing on the impact of technology on public institutions. Among the grant-funded projects he has directed are a global assessment of the impact of open government data; comparative benchmarking of government innovation efforts against those of other countries; and crafting the experimental design for testing the adoption of technology innovations in federal agencies. @_AndrewYoung

CONTENTS

| | |
|--|------------|
| EXECUTIVE SUMMARY | 7 |
| INTRODUCTION | 23 |
| I. What makes open data uniquely relevant to developing economies? | 27 |
| What is Open Data?..... | 27 |
| What makes open data uniquely relevant to developing economies? | 29 |
| II. How can the impact of open data on developing economies be captured and evidence be developed?..... | 37 |
| Open Data in Developing Economies Logic Model..... | 38 |
| Examining the Open Data in Developing Economies Logic Framework..... | 43 |
| Types of Impact | 44 |
| Enabling Conditions and Disabling Factors | 50 |
| The Challenge of Scaling and Replication..... | 58 |
| III. The Realized and Intended Impacts of Open Data in Developing Economies across Sectors..... | 61 |
| Health | 62 |
| Humanitarian Aid..... | 67 |
| Agriculture and Nutrition..... | 69 |
| Poverty Alleviation | 73 |
| Energy | 74 |
| Education..... | 77 |
| IV. How can open data be leveraged as a new asset for development? | 81 |
| APPENDIX A: List of Open Data in Developing Economies Case Studies | 95 |
| APPENDIX B: Input Received..... | 101 |
| APPENDIX C: Periodic Table of Open Data Impact Elements..... | 107 |
| APPENDIX D: Selected Readings on Open Data in Developing Economies | 123 |



OPEN DATA IN DEVELOPING ECONOMIES

Toward Building an Evidence Base
on What Works and How

Hanny Naibaho - Unsplash



EXECUTIVE SUMMARY

Across the world, governments are acting on the belief that systematically making data more accessible can provide an important new asset to usher in positive social and economic transformation. This trend is not limited to countries with more developed economies. Although the bulk of data in terms of quantity has thus far been released in developed countries, a growing number of developing economies — in Asia, in Africa, in Latin America — have also been adopting open data plans and policies, and publishing government datasets that previously remained locked away in closed databases. This move toward open data is part of a broader global trend toward more data-driven decision making in policymaking and development — a manifestation of what is sometimes called the “data revolution.”

The growing enthusiasm surrounding open data gives rise to several questions about open data’s unique features to foster change. Can it truly improve people’s lives in the developing world — and, if so, how and under what conditions?

The goal of this paper is to map and assess the current universe of theory and practice related to open data for developing economies, and to suggest a theory of change that can be used for both further practice and analysis. We reviewed the existing literature, consulted with the open data community and sought to collect evidence through a series of 12 in-depth case studies spanning multiple sectors and regions of the world. In particular, we sought to answer the following three questions:

What makes open data uniquely relevant to developing economies?

How can the impact of open data in developing economies be captured and evidence be developed?

How can open data be leveraged as a new asset for development?

WHAT MAKES OPEN DATA UNIQUELY RELEVANT TO DEVELOPING ECONOMIES?

Although there has been much focus on the potential of data to usher in change, what makes open data different is often not as well recognized, with supporting analysis often speculative or anecdotal. Understanding the distinguishing features of open data is important to subsequently document how open data works — its mechanisms and pathways — and in doing so, to build a more solid evidence foundation for open data in development. Based on our review of the existing narratives and theories in the literature, we identified six unique features that are believed to make open data specifically relevant — or potentially powerful — in the context of developing countries. They reflect some of the features that are often also associated with the value of open source or open innovation in a developing context; and explain why open data is closely connected with the twin trends of open government and open development.

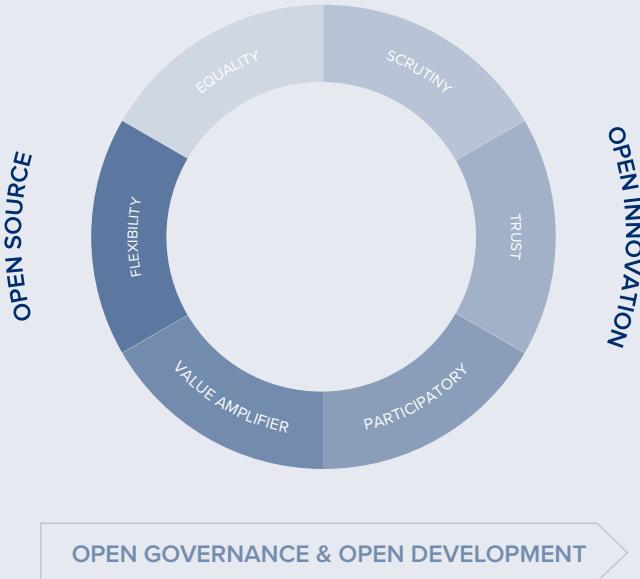
HOW CAN THE IMPACT OF OPEN DATA IN DEVELOPING ECONOMIES BE CAPTURED AND EVIDENCE BE DEVELOPED?

To analyze and capture existing evidence and to inform future comparative analyses across sectors and countries, we developed, expanded, and sought to start testing the following Change Theory and/or Logic Model:

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).



THE UNIQUE FEATURES OF OPEN DATA



Participation

By facilitating citizen participation and mobilization, open data can allow a wider range of expertise and knowledge to address and potentially solve complex problems.

Trust

Because open data increases transparency and avenues for citizen oversight, unlocking data can lead to higher levels of trust throughout societies and countries.

Equality

Open data can lead to more equitable and democratic distribution of information and knowledge — though, several observers have also pointed out that just releasing open data can play a role in further entrenching power asymmetries related to access to technology and data literacy.

Scrutiny

Because open data is subject to greater scrutiny and exposure than inaccessible institutional data, there is potential for enhanced review and improvement in the quality of government data by actors outside government.

Value Amplifier

Opening government datasets in a flexible and equal manner can amplify the value of data by filling — and identifying — important data gaps in society.

Flexibility

When released in an interoperable, machine-readable manner, open data is easier to repurpose and combine with other pieces of information, which in turn means that it is more flexible, with secondary data potentially yielding innovative insights.

LOGIC MODEL OF OPEN DATA

This logic model is built around the premise (informed by the case studies) that high-impact open data projects are the result of matching supply and demand so that open data can be effectively used to inform specific activities and outputs aimed at improving development. These outputs and activities can, in turn, serve a broader and more diverse group of users and objectives.



We categorized open data's (intended or realized) impact on development along the following pathways:

- ▶ **Creating economic opportunity**, by enabling business creation, job creation, new forms of innovation, and more generally spurring economic growth
- ▶ **Helping to solve complex public problems** by improving situational awareness, bringing a wider range of expertise and knowledge to bear on public problems, and by allowing policymakers, civil society, and citizens to better target interventions and track impact
- ▶ **Improving governance**, for instance by introducing new efficiencies into service delivery, and increasing information sharing within government departments
- ▶ **Empowering citizens** by improving their capacity to make decisions and widen their choices, and by acting as a catalyst for social mobilization

We found that there is wide variability of documented evidence. Much of the literature remains focused on the potential of open data to bring about positive benefits — absent real evidence and data. Little distinction is made between intent, implication, and impact, blurring our understanding of what the true value may be of open data. Several of the cases we analyzed did provide evidence that open data had an impact on people's lives while some failed to achieve notable, scalable impact, or worse, created new harms or risks.

To address the variability in evidence, and start focusing on testable premises of how, and under what conditions, open data works best in developing economies, we subsequently identified 27 enabling or disabling factors within the following five categories:

- ▶ **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 that affect the use of open data are made in a responsive manner

By combining these factors we developed a “Periodic Table of Open Data’s Impact Factors” that can be used (and further fine-tuned) as a canvas or checklist toward designing open data initiatives. Together with the logic model, the periodic table can also be used as a framework for analyzing and capturing key evidence of what contributes to “successful” open data efforts in developing economies.

PERIODIC TABLE OF OPEN DATA IMPACT

| Problem and Demand Definition | Capacity and Culture | Governance | Partnerships | | Risks |
|-------------------------------|--------------------------------------|---------------------|------------------------|----------------|--|
| | | | Pr | Ds | Dr |
| U | C | Di | Od | Dh | Ds |
| User Research | Causes and Context | Data Infrastructure | Open by Default | Data Holders | Data Security |
| Rf | Pu | Se | Fi | I | Dm |
| Refinement | Public Infrastructure | Skills & Expertise | Freedom of Information | Intermediaries | Poor decision-making due to faulty information |
| Bg | Lp | M | Dq | De | Pa |
| Benefit and Goals | Tech Literacy & Internet Penetration | Performance Metrics | Data Quality | Domain Experts | Entrenching power asymmetries |
| Da | Rs | Rm | R | Co | Ow |
| Data Audit and Inventory | Cultural/ Institutional Roadblocks | Risk Mitigation | Responsiveness | Collaborators | Open washing |



Chepe Nicoli - Unsplash

HOW CAN OPEN DATA BE LEVERAGED AS A NEW ASSET FOR DEVELOPMENT?

Despite the variability in evidence, we identified several key take-aways enabling us to make the following specific recommendations for open data practitioners and decision makers, including donors like USAID, on how to better leverage open data as a new asset for development.

Focus on and define the problem, understand the user, and be aware of local conditions. The most successful open data projects are those that are designed and implemented with keen attention to the nuances of local conditions; have a clear sense of the problem to be solved; and understand the needs of the users and intended beneficiaries. Projects with an overly broad, ill-defined or “fuzzy” problem focus, or those that have not examined the likely users, are less likely to generate the meaningful real-world impacts, regardless of funds available. Too often, open data projects have less impact because they are overly data-focused rather than problem- and user-focused.

Focus on readiness, responsiveness, and change management. Implementing open data projects often requires a level of readiness among all stakeholders, as well as a cultural transformation, in the way governments and institutions collect, share, and consume information. For development funders, however, this important determinant of success can imply difficult decisions regarding high-potential open data initiatives in developing economies that lack clear institutional readiness or demonstrated responsiveness to feedback.



Nurture an open data ecosystem through collaboration and partnerships. Data does not exist in isolation. The success of open data projects relies on collaboration among various stakeholders, as well as collaboration with data scientists and topic or sector experts. During the problem definition and initial design phase, practitioners and funders should explore the types of collaborations that could increase uptake and impact. Such partnerships could, for example, take place with other data providers (perhaps from different sectors), like-minded international or local organizations, and established intermediaries such as journalists or industry groups.

Have a risk mitigation strategy. Open data projects need to be mindful of some of the important risks associated with even the most successful projects. Notably, these risks include threats to individual privacy (for example, through insufficiently anonymized data or commingling multiple datasets to create new privacy issues), and data security. Funders should ensure that projects that deal in information that is potentially personally identifiable (including anonymized data) have audited any data risks and developed a clear strategy for mitigating those risks before proceeding with the project.

Secure resources and focus on sustainability. Open data projects can often be initiated with minimal resources, but require funding and additional sources of support to sustain themselves and scale up. It is important to recognize that access to funding at the outset is not necessarily a sign that open data projects are destined for success. A longer term, flexible, business model or strategy is a key driver of sustainability, and should be developed in the early stages of the design process.

Build a strong evidence base and support more research. As demonstrated in this paper, of perhaps even greater importance is the development of an evidence base that can provide feedback and indications of what works and what doesn't work, to maximize the impact of the (often scarce) resources available. The analytical framework provided in this report could act as a starting point for developing more systematic and comparative research that would allow for more evidence-based open data practice. In addition, to turn insights gathered into more evidence-based practice we need to translate these findings into new methodologies or checklists of key steps and variables (leveraging the Periodic Table provided in this paper).

In conclusion, given the nascent nature of existing open data initiatives, the signals of open data's impact in developing economies are still largely muted, as evidenced in the examples discussed in our paper. Our goal in this paper was not to use these examples as the ultimate proof of open data's importance for development; rather, we have picked up these signals and placed them into an analytical framework to enable further practice and analysis going forward. It is only with this type of structured analysis that we can gain a systematic and comparative evidence base of if and how open data is having meaningful impact on conditions on the ground in developing economies.

THE OPEN DATA FOR DEVELOPING ECONOMIES CASE STUDIES

In this table, we list the case studies developed to inform this paper, identify the intended impact of each, and note which of the key take-aways described above are represented in the projects studied.

| Case Study | Summary | Intended Impact | Risk Mitigation Strategy | Resources and Sustainability |
|--|--|----------------------|--------------------------|------------------------------|
| Burundi: Open Results and Performance-Based Financing | Results-based financing (RBF) in Burundi is an instrument that links development financing with pre-determined results, aimed to strengthen accountability and transparency in government expenditure. | Improving Government | ✓ | ✓ |
| Cambodia: Opening Information on Development Efforts | Open Development Cambodia helps to improve the public's awareness of current and historical information on development efforts through a data-driven online platform. | Improving Government | ✓ | ✓ |
| Colombia: Establishing Climate Resilience in Agriculture | The Aclimate Colombia project processes, analyzes, and publishes open government datasets to help farmers access climate and market data to improve their decision making and livelihoods. | Creating Opportunity | ✓ | ✓ |
| Ghana: Empowering Smallholder Farmers | Esoko is a for-profit communication tool that provides farmers repackaged data from different sources (including government and crowd-sourced data) via mobile phones to empower smallholder farmers. | Creating Opportunity | ✓ | ✓ |
| India: Open Energy Data | The Electricity Supply Monitoring Initiative (ESMI) collects real-time power quality information in an effort to monitor power quality in urban and rural India. | Improving Government | ✓ | ✓ |
| Jamaica: Open Data to Benefit Tourism | Jamaica uses open data to increase and improve its tourism sector, particularly by combining crowdsourced mapping data with open government data to create more participatory tourist maps. | Creating Opportunity | ✓ | ✓ |

| Case Study | Summary | Intended Impact | Clear Problem Definition | Readiness, Responsiveness and Change Management | Open Data Ecosystems and Partnerships | Risk Mitigation Strategy | Resources and Sustainability |
|--|---|-------------------------|--------------------------|---|---------------------------------------|--------------------------|------------------------------|
| Kenya: Improving Voter Turnout with Open Data | In the lead up to Kenya's 2013 general election, Code 4 Kenya published election data on the website GotToVote!, which provided citizens with voter registration center information, and easy-to-access information about registration procedures. | Empowering Citizens | ✓ | ✓ | ✓ | ✓ | ✓ |
| Nepal: Open Data to Improve Disaster Relief | A number of crowdsourced and open mapping data platforms helped humanitarian relief efforts in the wake of Nepal's 2015 earthquake. | Solving Public Problems | ✓ | ✓ | ✓ | ✓ | ✓ |
| Paraguay: Predicting Dengue Outbreaks with Open Data | The National Health Surveillance Department of Paraguay opened data related to dengue morbidity, which was used to create an early warning system. | Solving Public Problems | ✓ | | ✓ | | |
| South Africa: Code4SA Cheaper Medicines for Consumers | Code for South Africa uses data from the National Department of Health for its Medicine Price Registry Application (MPRApp), an online tool that helps patients identify and access cheap medicine. | Empowering Citizens | ✓ | ✓ | ✓ | | |
| Tanzania: Open Education Dashboards | Two portals in Tanzania publish education data in efforts to improve Tanzania's schools. The first, the Education Open Data Dashboard (educationdashboard.org), supports open data publication, accessibility and use. The second, Shule (shule.info), attempts to use open data to catalyze social change in Tanzania. | Empowering Citizens | | | | ✓ | |
| Uganda: Opening Health Data to Improve Outcomes | Uganda's iParticipate project uses open data available from government portals and other sources to analyze health service delivery and public investments in health projects. | Improving Government | ✓ | ✓ | ✓ | ✓ | |

LIST OF ACRONYMS

| | |
|----------------|---|
| BBW | Banana Bacterial Wilt |
| BODI | Burkina Open Data Initiative |
| CIAT | International Center for Tropical Agriculture |
| DFID | Department for International Development, United Kingdom |
| EITI | Extractive Industries Transparency Initiative |
| ELOG | Elections Observation Group |
| EOSDIS | Earth Observing System Data and Information System |
| ESM | Electricity Supply Monitor |
| ESMI | Electricity Supply Monitoring Initiative |
| EITI | Extractives Industries Transparency Initiative |
| GODAN | Global Open Data Initiative for Agriculture and Nutrition |
| IDRC | International Development Research Centre |
| IEBC | Independent Electoral and Boundaries Commission, Kenya |
| IMCO | Mexican Institute for Competitiveness |
| MERC | Maharashtra Electricity Regulatory Commission |
| MPR | Medicine Price Registry, South Africa |
| MPRApP | Medicine Price Registry Application |
| NERC | National Ebola Response Centre |
| NGO | Non-Governmental Organization |
| OD4D | Open Data for Development |
| ODC | Open Development Cambodia |
| ODI | Open Data Institute |
| OGP | Open Government Partnership |
| OSM | Open Street Map |
| PEG | Prayas Energy Group |
| PHC | Primary Health Care Centre |
| PII | Personal Identifiable Information |
| RBF | Results Based Financing |
| SDG | Sustainable Development Goals |
| TAP | Transparency, Accountability and informed Participation |
| UNICEF | The United Nations Children's Fund |
| USAID | United States Agency for International Development |
| WOUGNET | Women of Uganda Network |



Matt Jones - Unsplash

ACKNOWLEDGEMENTS

The authors would like to thank our partners at USAID: Mark Cardwell, Samir Doshi, Priya Jaisinghani, Merrick Schaefer, Vivian Ranson, Josh Machleder, Brandon Pustejovsky, Subhashini Chandrasekharan; and at FHI 360: Hannah Skelly and Abdul Bari Farahi who provided essential guidance and input throughout the project.

We would also like to acknowledge the great team responsible for the case study research that informs this paper: Michael Canares and Francois Van Schalkwyk from the Web Foundation; and Anirudh Dinesh, Auralice Graft, Juliet McMurren and Robert Montano at the GovLab. Editorial support for this paper was provided by Akash Kapur and David Dembo. The members of our Advisory Committee (see Appendix C) were a great resource to the project for which we are grateful. Finally, special thanks to the stakeholders we interviewed to gain on-the-ground and expert perspectives on the use of open data in developing economies, as well as the peer reviewers who provided input on a pre-published draft (see Appendix C).

This paper was authored by Stefaan G. Verhulst and Andrew Young of the GovLab.



www.odimpact.org