PART 1
Open Data for Development
CHAPTER 2

The Relevance of Open Data for Developing Economies

What is Open Data?

In this book, open data is defined as follows:

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.

Not all forms of data shared actually possess all the attributes included in this definition, nor do they necessarily conform to all the principles found in the Open Data Charter. In many ways, this is a gold-standard definition of open data, an important target to work toward. In fact, the openness of data exists on a continuum, and many forms of data that are not strictly “open” in the sense defined above are nonetheless shareable and usable by third parties. It is this broader sense of “open” that is used in this paper.

Open data exists in a wide variety of fields and domains. Three sectors in particular are responsible for producing the bulk of open data: governments, scientists, and corporations. In this paper, we focus mainly on the release and use of government data. We acknowledge, however, the importance and often untapped potential of more open access to science data and corporate data. Those other data sources, as well as crowdsourced data collection are also often mashed up with open government data, supplementing official public datasets to create new insights, opportunities, and impacts as a result. In what follows, we deconstruct the main reasons why open government data matters to developing economies.


6 Open by Default; Timely and Comprehensive; Accessible and Usable; Comparable and Interoperable; For Improved Governance and Citizen Engagement; For Inclusive Development and Innovation. http://opendatacharter.net.
The literature on open data reflects considerable enthusiasm about the potential for open government data in development. For example, a recent report published by the Open Data for Development Network suggests that open data is central to the development community’s goals of “enabling widespread economic value, fostering greater civic engagement and enhancing government transparency and accountability to citizens.” The report goes on to argue that “open data is increasingly recognized as a new form of infrastructure that is transforming how governments, businesses, and citizens are organized in an increasingly networked society.”

We do find some evidence to support such enthusiasm: across sectors, we see signs that open data can indeed spur positive economic, political, and social change. On the other hand, we also find grounds for caution; the impacts of many of the projects we examined remain largely aspirational or speculative, and some cases even led to harms (or potential harms). Although the real-world impacts of open data in developing economies remain emergent, it is important to distil these early lessons and develop a frame of analysis to support the current window of opportunity to increase access to data sources—a window that is likely to close absent any further evidence of open data’s impact or a better, more targeted description of the value proposition and change theory driving the field.

What makes open data uniquely relevant to developing economies?

We live in an era of big data. Every day, an unprecedented amount of information is being generated by an ever-increasing diversity of devices and appliances. Today, a growing consensus exists that this data, if applied correctly, and with attention to the attendant risks, can help spur positive social change. Sometimes called the “data revolution,” this new paradigm often fails to distinguish between the benefits of data per se and the complementing benefits of unlocking government data.

---

9 At the same time, there is a growing recognition of open data’s potential to meet the Sustainable Development Goals (SDGs). The World Bank, for instance, has explored the various ways in which open data could help to make progress toward the SDGs. Similarly, a recent White Paper by the Open Data Institute (ODI) concludes: “Open data can make an impact across the globe. Its role in combating development challenges of the next 15 years, both as a tool for measuring progress and in finding solutions, is becoming more clear.” See The World Bank, “Open Data for Sustainable Development,” Policy Note, August 2015, http://bit.ly/2aGjaJ4; and Open Data Institute, Supporting Sustainable Development with Open Data, 2015, 3, http://theodi.org/supporting-sustainable-development-with-open-data.
Based on our examination of the narratives and evidence provided in the existing literature, six distinguishing features seem to be credited to open data. Although these characteristics are unique to open data, in many cases, they would not be possible without a broader data, technology, and innovation ecosystem. Across our case studies, we’ve seen that the existence of a strong information and communication technology for development (ICT4D) sector in a country, for example, tends to result in higher impact, more quickly developing open data efforts.

With the understanding in mind that open data must exist in a strong ecosystem, the six distinguishing features that are most quoted with regard to open data in a development context include:

- **Scrutiny**: Because open data is subject to greater scrutiny and exposure than inaccessible institutional data, there is potential for enhanced review and improvement of government data quality (e.g., by data-literate civil society groups or other crowdsourced methods). This can result in more useful data—again, a benefit that is relevant in less developed countries and societies, where data is scarce, and of limited quality and usefulness.¹⁰

---

¹⁰ Tim Davies, for instance, notes that, “researchers and other users outside of government may highlight inaccuracies and inconsistencies between datasets” as a result of access to open data, and thereby improving data quality and usefulness. Tim Davies, “Open Data in Developing Countries: Emerging
• **Equality**: Open data can lead to an inherently more equitable and democratic distribution of information and knowledge. This is a key intended benefit in all countries, but particularly salient in many developing economies that struggle with large socioeconomic and digital divides. It is important to keep in mind, however, that the lack of Internet penetration and access to tools for using and accessing open data still present challenges in many contexts—and, indeed, such technological inequities can be further entrenched through open data in some cases.

• **Flexibility**: Open data is open with regard not only to the information it contains, but also to its format. This means that, when released in a usable manner, open data can be easier to repurpose and combine with other pieces of information than data institutions fail to make accessible, which in turn means that it is more flexible, with secondary uses that are likely to yield innovative insights. This is true of data from all sectors, but perhaps especially of government data, which often exists in vast, untapped silos; opening that data (turning it “liquid”) can play a key role in generating new insights and policies. Such liquidity can only become a reality if data, and the tools used to manipulate it, are interoperable and adhere to agreed upon standards. Creating such technical capacity can, however, lead to opportunity costs and require significant upfront resource allocation on the supply side, potentially slowing progress at the outset.

• **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. This quality of “open innovation” can allow resource-starved developing economies to access and benefit from the best global minds and expertise. It can offer a more participatory way of solving complex public dilemmas, with pathways toward more easily tapping into previously inaccessible knowledge (e.g., those related to social and economic development).

---


Chapter 2 – The Relevance of Open Data for Developing Economies

- **Trust:** Because it increases transparency and avenues for citizen oversight, unlocking data can lead to higher levels of accountability and trust throughout societies and countries. This “sunlight” or “trust” quality of open data can have powerful ripple effects, including incentives for better government practice, and the enhancement of the quality of public life and citizenship. Such increases to trust and accountability rely on meaningful data being made open, however, rather than governments participating in “open washing” where largely useless datasets are made accessible toward boosting institutional reputation.

- **Value amplifier:** Finally, it is now widely recognized that data is a new kind of asset or knowledge is a form of wealth. The opening of government datasets in a flexible and equitable manner can amplify the value of data thanks to data filling important data gaps felt in society. Though this attribute is important across the world, it may have a particularly important role to play in developing economies. In its 2016 *World Development Report*, The World Bank pointed out that technology can play an “accelerator” role in developing countries. But while the inherent scarcity of resources (data and otherwise) in the developing world increases the apparent value and potential impact of open data, cultural and political barriers to timely and well-targeted open data provision efforts could slow progress.

These narratives surrounding the open data movement reflect those associated with the cross-sector paradigm shift from closed processes to open ones, and how it applies to governance and development. Software, for example, is increasingly developed in an open source manner. With the rise of the collaborative coding platform GitHub, a notable driver, the open source movement, similar to open data, is seen to be providing for more equal and flexible ways to create and access code—resulting in distributed coders, not just tech company employees, creating and improving exciting new products. Similarly, businesses and governments alike are embracing open innovation techniques, posing opportunities to the crowd to provide input on important challenges and absorbing the best ideas—providing for enhanced participation.

---


and scrutiny, other features of open data. The emerging fields of open governance and open development are also built on similar principles and techniques (see text box below).

**OPEN GOVERNMENT AND OPEN DEVELOPMENT**

*Open Governance/Government*\(^\text{18}\)

Definitions of open governance or open government vary not only across sectors but within them. Definitions focus to varying degrees on the key elements of transparency, citizen participation, and collaboration, among others, depending on the context. Some illustrative examples of “open government” definitions\(^\text{19}\) include:

- **Wallace Parks**, “Open Government Principle: Applying the right to know under the Constitution” (1957)—“The general availability of government information is the fundamental basis upon which popular sovereignty and the consent of the governed rest, subject to several important restrictions on this general rule (i.e., to allow for the carrying out of the constitutional powers of the Congress and the President; to protect the personal and property rights of individuals, corporations and associations; to acknowledge administrative complications as to whether to release, to withhold, or to partially release particular types of information under particular conditions; to protect confidentiality of communications internal to government; to acknowledge the difficulty of segregating information when parts of a document should be released and parts withheld).”\(^\text{20}\)

- **White House**, “Transparency and Open Government: Memorandum for the Heads of Executive Departments and Agencies” (2009)—“Open government is defined as a system of transparency (information disclosure; solicit public feedback), public participation (increased opportunities to participate in policymaking), and collaboration (the use of innovative tools, methods, and systems to facilitate cooperation among Government departments, and with nonprofit organizations, businesses, and individuals in the private sector).”\(^\text{21}\)

---


Chapter 2 – The Relevance of Open Data for Developing Economies

- Open Government Partnership, Open Government Partnership Declaration (2011)—“Open government involves: Increasing the availability of information about governmental activities; Supporting civic participation; Implementing the highest standards of professional integrity through: Increasing access to new technologies for openness and accountability, information sharing, public participation, and collaboration.”

Open Development

In the wake of open government taking hold as an organizing concept for improving and innovating governance, open development has evolved as a more networked and innovative pathway to improving international aid and development efforts. In a book on the topic, the International Development Research Centre (IDRC) seeks to gain clarity on the contours of the field of open development. Their assessment of the theory and practice involves a number of key elements present in open development work, including:

- The power of human cooperation
- Sharing ideas and knowledge
- The ability to reuse, revise and repurpose content
- Increasing transparency of processes
- Expanding participation
- Collaborative production

Based on the examination of these strands of openness in development efforts from the World Bank, ONE, African Development Bank, and others, the IDRC authors conclude that the central idea behind open development is: “harnessing the increased penetration of information and communications technologies to create new organizational forms that improve the lives of people.”

24 Ibid.
25 Ibid.
The Data Life Cycle

An important insight from the emerging research and practice is that data is not “a thing” but involves a “process”—what we call a “data life cycle."

![Figure 2. Data Life Cycle](image)

How each stage of the data life cycle is implemented—from collection to processing and sharing; to analysis and using; and back to the start of the cycle again—will determine the value of data and who ultimately benefits. Disparities among those who collect and have access to data or have the capacity to make sense of the data can reinforce existing imbalances in power or influence. This is especially true in a developing economies context where the number of data holders and data scientists is more concentrated, and this smaller group is disproportionately empowered to make meaningful use of data. Within that context, opening datasets is often characterized as a force for democratization—engaging private and civil society actors, and, often indirectly as a result of intermediation, citizens themselves in analyzing and using data.

As such, open data provides for unique efficiencies by leveraging civic-minded technologists (government and nongovernment), as well as entrepreneurs, to analyze, disseminate, and/or use data in a new, sometimes profitable way, as discussed more below.

On the other hand, each stage of the data life cycle contains risks. Risks are often the result of technological weaknesses (e.g., security flaws); individual and institutional norms and standards of quality (e.g., weak scientific rigor in analysis); legal confusion or gaps; or misaligned business and other incentives (e.g., companies seeking to push the boundaries of what is socially appropriate).

Although there are common elements across these risks, it is useful to examine them by separately considering each stage of the data value cycle.

---

When risks are not addressed at the initial stages of the value cycle (e.g., when dirty data is not cleaned at the collection stage) they may accumulate and lead to additional risks downstream (e.g., making flawed inferences from the data analysis due to inaccurate data). Therefore, it is important to consider potential risks not just at the points of opening data, but also at the data collection stage and evaluate those risks vis-à-vis the (potential) value of releasing the data. As such, to prevent harm, there may be a legitimate case—especially when there is a clear understanding of the purpose of the use and user—to share certain government datasets with those targeted audiences in a more protected manner to generate necessary insights while limiting the risks. We examine the risks introduced by open data efforts in developing economies in more detail in Part II.