ATHENA INFONOMICS

Barriers to Data Use in Sustainable Development

Barriers to Data Use in Sustainable Development

Shruti Viswanathan[1], Deepa Karthykeyan[2], and Wyatt Williams[3]

August 2021

Context

Today, there is global consensus that responsible data access, sharing, and use can lead to social and economic 'good.' Several studies estimate the economic impact of the data market to be in the hundreds of millions USD, a number that grows with every year. A 2017 study found that the use of open data from the London transport authority led to annual savings of nearly GBP 130 million. This was attributed to the time saved by passengers and road users, reuse of Transport for London's data for commercial purposes, job creation, and improved operations.[4] Citizen access to and use of data can help improve transparency and accountability of governments. In India, the Right to Information Act (2005) gave citizens access to information held by public authorities "to promote transparency, contain corruption, and hold the Government accountable to citizens."[5] In addition to government data, data generated by the private sector has also had a huge impact on social good. A 2021 report suggests that 60 percent of private sector organisations see progress on the UN Sustainable Development Goals (SDGs) as a key reason to participate in shared data ecosystems.[6] Successful collaborations with the private sector can help foster social and economic good through the use of data.

While the demand for and supply of data have grown exponentially in the past decade, the data ecosystem has become more complex with an expanding number of actors and limited progress on the data governance frameworks needed for this rapidly evolving field.[7]

---

[1] Shruti Viswanathan leads the Governance and Technology practice at Athena Infonomics

[2] Deepa Karthykeyan is Co-Founder and Director at Athena Infonomics

[3] Wyatt Williams is a Summer Associate at Athena Infonomics

[4] Deloitte (2017), 'Assessing the value of TfL's open data and digital partnerships', http://content.tfl.gov.uk/deloitte-report-tfl-open-data.pdf.

[5] Centre for Internet and Society, 'Transparency and Privacy', https://cis-india.org/internet-governance/blog/transparency-and-privacy.pdf/view.

[6] Capgemini Research Institute (2021), 'Data Sharing Masters: How smart organisations use data ecosystems to gain an unbeatable competitive edge', https://www.capgemini.com/wp-content/uploads/2021/07/Final-Web-Version-of-Report-Data-Ecosystems.pdf

[7] The World Bank (2017), 'Conversations: Strengthening the role of data in policy making-Perspectives from the World Bank Group, OECD, and AidData', https://ieg.worldbankgroup.org/news/role-of-data-in-policy-making.

Additionally, the ability to produce data has grown at a faster pace than the ability to understand and use this data. This mismatch has led to abundant 'data graveyards' where large quantities of data remain unused.[8]

About the Data Values Project

The Data Values Project is a policy consultation and advocacy campaign to develop the data norms needed to achieve the SDGs. Bringing more than 270 partners together across geographies and sectors, the Data Values Project aims to harness collective global action on data use, paving the way for policies that unlock the power of data for better lives. This framing paper fits within the Data Values Project's working group focused on data use that drives sustainable and equitable development. To harness the full potential of data for the SDGs and to deliver better policies and services to citizens around the world, it is important to identify and unpack the barriers and drivers for sustained data use and to establish sustainable and inclusive engagement of key stakeholders around this topic.

The Global Partnership for Sustainable Development Data's Technical Advisory Group leads the Data Values Project and has identified three thematic tracks of action focusing on inclusion and equity, sustainable and equitable development, and data governance. We aim to gather available evidence and knowledge around these themes, with the objective of mapping key pathways for sustained data use and recognising their fundamental components. Sharing knowledge on what roots effective data use in policy and service delivery practices will help accelerate achievement of the SDGs and ensure that the data revolution serves citizens and communities.

This paper provides a brief overview of the key barriers to optimal data use.[9] We discuss the common challenges faced, the existing incentives for data use, and the downside to generating data that is never used. The paper ends with key questions we are seeking to address through the Data Values Project.

Barriers to data use

To better understand 'data for public good,' we need to understand who uses this data and for what purposes. Data users can be broadly categorised as: governments/the public sector, the private sector, individuals/the public, local civil society organisations, and international institutions engaged in policy and programming (i.e. multilateral institutions, global civil society, thinktanks, etc.). The use of data for public good is seen in a few major focus areas: policy/programme planning and implementation, monitoring and evaluation, accountability, and advocacy. The types of data needed, along with the incentives and capacities for data use vary for each data user and use type. For example, the barriers

---

[8] AidData (2017), 'Avoiding Data Graveyards', http://docs.aiddata.org/ad4/pdfs/avoiding_data_graveyards_full_report.pdf.
[9] This paper builds on an earlier GPSDD study entitled 'Barriers and enablers on the use of data for decision making.'

experienced by local civil society organizations (user segment) using data to hold governments to account and to advocate for changes may be different from the challenges associated with building a culture of data use within government. We hope to systematically unpack these challenges by user segments and use cases through a series of consultations under the Data Values Project.

In the following pages, this paper seeks to provide a high-level overview of the five types of barriers that may hinder data use for development across user segments and uses: *intrinsic data barriers, stakeholder capacity, institutional incentives for data use, data cultures,* and *resourcing barriers*. This broad typology helps classify the types of barriers encountered in data use and identify strategies to overcome them.

A. Intrinsic data barriers: One of the most important barriers to the use of data is that the data produced is not 'fit-for-purpose.' This includes the following challenges:

- Inconsistent data quality: While data collection is consistently increasing, the timeliness, credibility, accuracy, and completeness of data determine its usability. Data quality within these parameters remains inconsistent in public and private sector data, making it difficult to use this data for decision making. Several governments have introduced management information systems to streamline data collection and quality. Adequate physical and digital infrastructure is necessary for collecting, analysing, and sharing data. As new modes of data collection and generation become popular, internet access and digital infrastructure will become important in generating and using good quality data.

- Insufficient communication between data producers and data users: Another key challenge to collection and use of data is the relationship, or lack thereof, between data producers and data users, especially in the public sector. Often, there is a disconnect between those who are collecting and producing the data and those who are going to use the data. This results in cases where the data collected is not fit for decision making by users, and data collectors are not involved in identifying what data is needed. Overall, this leads to limited or inefficient data use.[10] An intervention in Cote d'Ivoire aimed to strengthen data use within the national health information system. This included engagements with data users and data producers, with membership from government program staff, partner organisations, and Monitoring and Evaluation (M&E) experts. The results from this intervention found that continuous engagement of data users and producers through the data cycle had a significant impact on the use of the data.[11]

> **IMPROVING THE UTILITY OF DATA COLLECTED IN THE DRC[12]**

---

[10] Masaki, et al. (2017), 'Decoding Data Use', http://docs.aiddata.org/ad4/pdfs/Decoding_data_use--How_do_leaders_source_data_and_use_it_to_accelerate_development.pdf

[11] Nutley, T., Gnassou, L., Traore, M., Bosso, A. E., & Mullen, S. (2014). Moving data off the shelf and into action: an intervention to improve data-informed decision making in cote d'Ivoire. *Global health action*, 7(1), 25035.

12 Technical Working Group Uses Data to Identify Data Collection Tool Stockouts and Find a Solution in DRC, https://www.measuremalaria.org/news/technical-working-group-uses-data-to-identify-data-collection-tool-stockouts-and-find-a-solution-in-drc/

> Engaging data collectors and users is a crucial strategy in overcoming barriers to data use. In the Democratic Republic of Congo (DRC), the President's Malaria Initiative (PMI) worked to improve the utility of data collected in the healthcare system by engaging in quarterly meetings with provincial staff, implementing partners, and Technical Working Groups (TWGs). During a meeting in December 2020, a shortage of data collection tools was identified in multiple health zones in the DRC. Many existing tools were outdated, incomplete, or non-standardized, and as a result, the quality of malaria data suffered. The Department of Provincial Health collaborated with partners to identify a solution for the financing of new collection tools. During this meeting, the TWGs created a mechanism for stakeholders to identify data-quality challenges and empowered these stakeholders to find a solution that would strengthen the quality of routine malaria data. These meetings serve as a space for data collectors and users to identify existing problems inherent to data use, find solutions to such problems, and to develop a framework for other stakeholders to find solutions at a smaller scale.

- **Digital exclusion:** Increasingly data is 'born-digital', i.e., created and collected digitally. Given the inequitable nature of digital access, such data runs the risk of over-representing the voices of digital natives and under-representing those who do not have access to the digital world. Use of such data to guide policy making can be detrimental to goals of social and economic justice.

B. Limited stakeholder capacity to leverage and use data:

- **Inadequate incentives for data producers:** One of the key barriers to the use of data in the public sector is the lack of skills and incentives amongst data producers (e.g. data managers, front-line health workers, etc.) to continue to generate useful data and analyse, understand, and present the data collected into easily understandable messages that are relevant to decision makers.[13] Limited training of data producers and lack of incentives for data producers to understand data use and analysis is an important barrier for stakeholder capacity to use data.

- **Inadequate incentives for data users:** The World Development Report (2021) recommends "strengthening the data and statistical literacy of decision-makers" as an important step in promoting data use in the public sector. In some cases, data producers are also the data users (e.g., frontline workers). However, without adequate understanding of how data can be used in a sustained manner, data collection becomes the main focus and is often driven by top-down compliance triggers or on-off project monitoring or planning needs. Ensuring incentives for data use will be a key lever in improving capacity to leverage and use data.

---

[13] Measure Evaluation (2018), 'Barriers to Use of Health Data in Low- and Middle-Income Countries A Review of the Literature'. https://www.measureevaluation.org/resources/publications/wp-18-211/at_download/document

- **Inequalities between data producers and data users:** Another cause for concern is the geographical power imbalance between those whose data is being used and the users of the data. As mentioned previously, data use has many potential economic and social benefits. While the production and use of big data is expected to generate large economic benefits, most of these benefits are expected to be reaped by those in technologically advanced countries.[14] Inequality expresses itself not only in access but also in skills and benefits.[15] Inadequate infrastructure to collect, process, and analyse datasets can lead to a widening gap between countries. As private sector actors, governments, and international organisations promote the data agenda, ensuring equitable access to the benefits of the data revolution is important.

C. Lack of institutional strategies to incentivise data use:

- **Unclear understanding of data use:** The manner in which data influences policymaking also remains unclear, making it difficult to understand which specific data points would be most useful.[16] While studies indicate that leaders of public organisations view data as helpful to policy making, there is less clarity on the specific use of individual data sources. Including end users throughout the data process would be important in shaping the kind of data that is collected and facilitating use of this data to inform relevant policies or programs, as opposed to the current model in which data users are presented data with limited consultation and context. [17]

- **Siloed data systems:** In some organisations and government ministries, data produced is closely guarded in 'information islands' which don't allow data sharing and use. This leads to duplication of data collection and sub-optimal use of resources.

> **PROMOTING DATA SHARING: ESTONIA'S X-ROAD INITIATIVE [18]**
>
> Data infrastructure that promotes secure sharing of data across different entities and departments can help overcome 'information islands' through optimal sharing of resources. The Estonian government created X-Road: an informational ecosystem that provides a secure place for resource sharing by governmental services, the private sector, and Estonian citizens. Through X-Road, users may communicate with

[14] Mann, Laura. "Left to other peoples' devices? A political economy perspective on the big data revolution in development." *Development and Change* 49.1 (2018): 3-36.

[15] Cinnamon, Jonathan. "Data inequalities and why they matter for development." *Information Technology for Development* 26.2 (2020): 214-233.

[16] Masaki, et al. (2017), 'Decoding Data Use', http://docs.aiddata.org/ad4/pdfs/Decoding_data_use--How_do_leaders_source_data_and_use_it_to_accelerate_development.pdf

[17] Masaki, et al. (2017), 'Decoding Data Use', http://docs.aiddata.org/ad4/pdfs/Decoding_data_use--How_do_leaders_source_data_and_use_it_to_accelerate_development.pdf

18 https://e-estonia.com/solutions/interoperability-services/x-road/; https://e-estonia.com/solutions/interoperability-services/x-road/ ; https://x-road.global/xroad-history" https://x-road.global/xroad-history

different information systems, search through a unified database, and send data sets. X-Road ensures privacy by authenticating and logging new data while encrypting and digitally signing outgoing data. Today, 99% of Estonian State Services as well as 52,000 private organizations use the platform, and it is estimated that X-Road has saved 844 years of working time for its clients. X-Road demonstrates that interoperable data sharing systems between different agencies can translate to substantial efficiency gains across the ecosystem.

- **Mismatch between donor priorities and domestic demand:** Several studies suggest that foreign aid often influences the types of data that is being collected, incentivising data that appeals to external actors over domestic policy makers.[19] Limited investments in data systems (beyond on-off project planning or monitoring) that are representative, dynamic, and useful to local actors leads to creation of inefficient data operations and systems.

AID TRANSPARENCY PORTAL (TIMOR L'ESTE) [20]

As stated on its website, the Aid Transparency Portal (ATP) was developed as a "central repository for all aid information in Timor-Leste, and aims to improve aid transparency, accuracy and predictability and to ensure assistance provided is efficient and effective."[21] While the ATP was created with the intent of increasing transparency and collaboration, a 2017 study found limited use of the data portal. Unreliable internet access, technical issues with the portal, and limited utility of the data formats were some of the problems that led partners and bureaucrats to conduct their own analysis. Development data was often published in English and Portuguese, rather than the local Tetum, making it difficult for many local officials to understand the data. Locally relevant data delivered by fostering relationships with the government officials was valued more. Thus, despite widespread demand for data amongst different actors, the ATP met with many challenges.

D. Restrictive data cultures: The cultural and behavioural norms of institutions regarding data may prevent or disincentivize an actor from using data.

- **Lack of trust:** Public distrust in government/large private sector actors includes fears that data may be used for political surveillance or economic gain of the private sector. To improve public

[19] Sandfeur and Glassman (2014), 'The Political Economy of Bad Data: Evidence from African Survey and Administrative Statistics'. Centre for Global Development.
[20] AidData (2017), 'Avoiding Data Graveyards', http://docs.aiddata.org/ad4/pdfs/avoiding_data_graveyards_full_report.pdf
[21] "Aid Transparency Portal." Accessed 20 August 2021. https://www.aidtransparency.gov.tl/portal/

trust in data use, there must be sound institutional and governance practices that advocate for data use while protecting the rights of citizens. Trust of data use within political systems can be instilled through the government and disseminated throughout smaller communities. Success in data use would require agencies to be transparent about how the data is used, receptive to the needs of the communities they serve, and to support locally driven ideas for data initiatives.[22]

- **Restricted data freedom:** Often, data is used to justify an established course of action rather than to evidence the need for change. Mistrust or misunderstanding of the value of data in decision-making could become a barrier to its use. Actors could also perceive incentives to use favourable data and ignore data that may not be viewed well by governments. To combat this bias, a culture of data freedom and inclusion must be established.[23]

    Public trust and improved data freedom can be built through progressive legislation and guidelines. For example, the 2019 Data Protection Act in Kenya outlines, for both the government as well as the general populace, how data will be collected, who will be responsible for the data, and how it will be used, as well as what safeguards are in place thereby ensuring transparency and incentivizing trust.[24]

E. Resource Barriers: The lack of investment, whether financial or time resources, may also deter data collection, analysis, and implementation. Limited financial and time resources might be insufficient for program goals or misallocated, leaving the program without capacity for growth. Often, processing money and data can take longer than anticipated, which can be difficult to align with donors' project timelines.[25] To ensure program success, data programs should have sufficient financing into appropriate areas, research into the methods and tools of the program, establishment of standards for data use, and finally, make the data accessible for both national priorities and local demand.

Conclusion

As this paper suggests, mere availability of data does not automatically lead to sustained data use. Available data might not be of sufficient quality, stakeholders might lack the capacity to use it, there might not be incentives for leveraging evidence in policy making, and the absence of a data culture might lead

---

[22] Measure Evaluation (2018), 'Assessing Barriers to Data Demand and Use in the Health Sector A toolkit'. https://www.measureevaluation.org/resources/publications/ms-18-134/at_download/document

[23] Fisher et al. (2021) 'Confronting Data Inequality' World Development Report 2021 background report https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3825724

[24] Deloitte (2021), Patel et al. 'Kenya Data Protection ',https://www2.deloitte.com/content/dam/Deloitte/ke/Documents/risk/Kenya%20Data%20Protection%20Act%20-%20Quick%20Guide%202021.pdf

[25] The World Bank (2020), 'Implementation Completion and Results Report: Modernizing Production of Statistics in Vietnam', https://documents1.worldbank.org/curated/en/705401622743777388/pdf/Implementation-Completion-and-Results-Report-ICR-Document-Modernizing-Production-of-Statistics-in-Vietnam-P166158.pdf

to underestimating the importance of data. The identification of these and other barriers has led to reframing the discussion within the data for development community, from an initial focus on making more data available to an increased debate around the conditions which determine whether data is used and how to increase its use. From this debate emerge a number of questions which underpin the [Data Values project](), alongside separate tracks on equity and inclusion, and data governance:

- What factors are most commonly perceived as barriers to data use? Are these barriers context-specific or common across communities and geographies? Are there successful strategies that can help overcome these barriers and enable data use? Which barriers are most commonly addressed by the development community, and which are more often left aside and deserve more attention?
- What are the drivers of sustained data use? What institutional incentives contribute to sustained data use by different stakeholders? How should such incentives be structured?
- How can we bridge the mismatch between data supply and demand and foster better communication between data producers and data users?
- What can be done to create sustainable and inclusive engagement around data use so that this practice is embedded in decision making? Which actors are pivotal to engage with if sustained data use is to be promoted?

Through a number of activities and discussions that draw on the research and experience of our broad partner network, we aim to gather insights on these questions and develop an advocacy agenda and practical recommendations for policy makers and practitioners to build a culture of sustained data use to drive development outcomes for all.