# Mapping open data for accountability

Transparency and Accountability Initiative and the Open Data Charter
Liz Carolan[1], May 2017

## The need for a new framework for open data for accountability:

There is growing recognition that open data is an asset for improving accountability and governance. This has driven a proliferation of work and investments in the transparency and accountability field connected to open data.

Yet despite the level of investment, **examples of impact have been slow to emerge**. There have been case studies which illustrate potential for impact and which demonstrate causal pathways, but few studies showing sustained results. This is feeding a growing skepticism, and affecting the work's sustainability.

It is likely that we need to make adjustments in the way that work in the field is organised, funded and delivered if we are to improve impact. To do this we will need a framework for understanding that work, against which we can map investments and projects to assess gaps or structural issues.

This paper offers a draft such framework. It builds on research[2] on the necessary conditions for data transparency to lead to accountability. It outlines five types of activity that need to take place from production and sharing of data, through processing, to action and an accountability response. Each stage has a set of requirements for the activity to take place, from technical capacity, to licenses, to institutional infrastructure, outlined in Table 1 below.

This framework is deliberately simple; many organisations and projects carry out activities that cut across the different types of activities, and reality is not linear. However, the framework is intended to offer a more nuanced way of understanding the field than the current "supply and demand" or "data availability and use" models[3]. This model should allow for a better analysis of the work that is already taking place to turn data into accountability, including potential project interdependencies. It should also give a better understanding of where the gaps are.

This paper concludes with recommendations for how TAI members (and other funders) could enable the mapping of projects against this framework. In doing so, it helps to inform investment and strategic planning decisions by organisations, including TAI and the ODC.

## A framework for understanding open data for accountability:

The prevailing conceptualisation of open data work uses the broad categories of data supply and demand, and work supporting impact (in this case, accountability).
   A.   Supporting data availability

---

[1] Illustrations by Brent Gorsky. Drafted with inputs from Ania Calderon, Michael Jarvis and Robert Palmer.
[2] Carolan, L. (2016). Open data, transparency and accountability: Topic guide. Birmingham, UK: GSDRC, University of Birmingham. http://www.gsdrc.org/topic-guides/open-data-transparency-and-accountability
[3] As used in, for example, recent reports by the OGP  and GovLab

B.  Supporting data use
C.  Supporting accountability

Yet this framework adds limited insights for identifying gaps and issues behind the slow emergence of evidence of impact. For this a more nuanced framework is needed, one that is based on the logic chain connecting data to accountability. Understanding this system is the best way to map the field, as insights this generated would illustrate what work is missing, or problematically siloed, and thus preventing results.

A recent guide[4] drawing on open data and transparency and accountability research, outlines that for open data and transparency initiatives to lead to accountability the following conditions are needed:
1.  The right information is published in the right way at the right time
2.  People and organisations are able to find, access and use this data
3.  There is space to generate and share insights, and demand a response
4.  Functional response systems are in place, to impose sanctions or introduce other changes; or citizens have choice or support from public officials

Taking this understanding as a starting point, it is possible to develop a practical framework, with one addition - "Data production" - in line with the challenge of missing data highlighted as part of the "Data Revolution" conversations[5].

The new framework outlines five types of activity that need to take place from production and sharing of data, through processing, to action and an accountability response.. Each has a set of requirements for the activity to take place, from technical capacity, to licenses, to institutional infrastructure, outlined in the table below.

**Diagram 2: Suggested framework for understanding the open data for the accountability field**
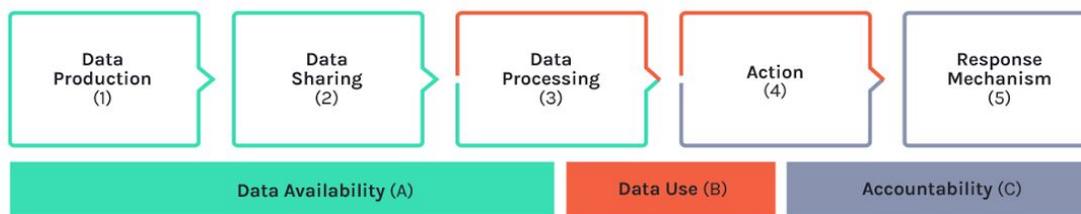


This framework intends to offer a more nuanced way of understanding the field than the current "data availability and use" models. Comparing the two models demonstrates that what we mean by each of data availability, use or accountability in fact encompasses at least 2 different sets of activities needed to get to impact, each with a range of conditions needed for success. For example, "data use" involves not just the capacity to process data to create insights, but the skills, environment and space to create and share outputs based on these insights.

---

[4] Carolan, L. (2016). Open data, transparency and accountability: Topic guide. Birmingham, UK: GSDRC, University of Birmingham. http://www.gsdrc.org/topic-guides/open-data-transparency-and-accountability/
[5] Initiatives such as the Global Partnership for Sustainable Development Data, which grew out of the Data Revolution research (http://www.undatarevolution.org/report/), highlights missing data as a key impediment to achieving the SDGs.

The table below takes each of these types of activity lays out the conditions that are necessary for each.

**Table 1: Requirements for each type of activity linking open data to accountability**

| Type of activity | Required assets, skills, systems and environmental conditions |
|---|---|
| **Data production**<br>Creating and storing data by governments and others | Governmental data collection, digitisation, storage and management mechanisms, and legal and policy framework<br>Capacity for other sources of data to be produced, inc. citizen generated data, sensors etc. |
| **Data sharing**<br>Opening access to data in ways anyone can use & share | Capacity, technology and portals for data release<br>Interactions between data producers and users to inform prioritisation, selection of formats<br>Managing risks, including around privacy<br>Licenses, standards and shared vocabularies<br>High level political will, operational level capacity and will, commitments and concrete policies |
| **Data processing**<br>Creating knowledge or insight from data, through algorithms, research, tools, data viz, investigations etc. | Capacity of data users to access, use, analyse and visualise data<br>Data literacy of decision makers, and broader citizenry<br>Existence of tools, open source code and algorithms to support processing<br>Equipped journalists, analysts, researchers, investigators with space to process data<br>*Licenses, standards and shared vocabularies* |
| **Action**<br>Sharing knowledge or insights as apps, advocacy, services, journalistic content, government action etc. | Space and capacity to share stories, insights, ideas, findings from data processing<br>Functioning market for outputs<br>Civic space and capacity to respond to stories and analysis, push for or demand change, and engage with government processes<br>Capacity for state institutions to carry out actions from improved data access |
| **Response mechanisms**<br>Changes in the world resulting from action - sanctions, policy change, citizens changing service providers etc. | Institutional infrastructure for formal accountability, for policy change or for other actions<br>Service user choice to enable "softer" accountability, such as social accountability<br>Capacity to respond within institutions<br>Knowledge production and sharing of impact (feeds into all of above) |

## Using the framework: TAI investment mapping

One potential application of this framework is in supporting TAI members to better understand their investments. To be able to understand the work that is taking place, funders and organisations would benefit from an ability to readily view, analyse and understand investments being made as they relate to this framework, ideally, in close to real time.

A sample of current investments was assessed in the drafting of this paper, and it showed that grantees were a complex and makeshift patchwork of public, private and non-state actors with different objectives. Most carry out more than one activity, ranging from building technical assets to training officials to research, campaigns and advocacy (see box at end of this section). Rather than looking at investments through either of these lenses, we can use the framework outlined above.

This would also allow  monitoring of whether investments are going towards the areas that are priorities for investment, and allow organisations to see where there are gaps presenting challenges or opportunities. A prototype of the visualisations that this would enable being built is below, using dummy data.

**Prototype 1b: Number of investments and amount invested by type of support (dummy data)**
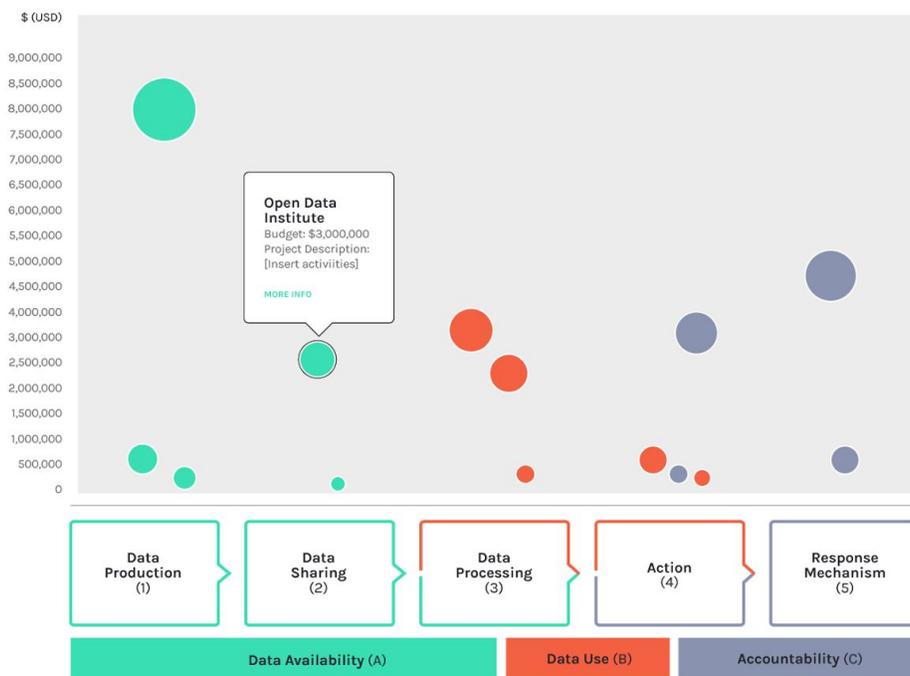


However, it was not possible to map the information on investments as it is currently collected against this framework in any systematic way. While all TAI members publish information on their investments, there is no shared system for classifying investments so they can be combined, analysed and visualised. Small changes to how TAI members (and other funders) manage information on investments could rectify this.

This paper recommends that **the new framework above be used as a building block for a system for tagging information on investments made by TAI members** (and ideally other funders). A coding

system for projects which tagged where an investment sits within the new framework would enable the TAI team to develop a dashboard or visualisations. Two options are to tag investments with a single "step":

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | **Grant recipient** | **Funder** | **Grant size** | **Impact step** | |
| 2 | Open Data Inc. | DFID | $100,000 | Data Sharing ▾ | |
| 3 | Transparency & Co. | OSF | $250,000 | | |
| 4 | | | | Data Production | |
| 5 | | | | Data Sharing | |
| 6 | | | | Data Processing | |
| 7 | | | | | |

Or to allow for an estimate of the proportion of grant amounts to be allocated to each "step".

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | **Grant recipient** | **Funder** | **Grant size** | **Data production** | **Data sharing** | **Data processing** | **Action** | **Response** |
| 2 | Open Data Inc. | DFID | $100,000 | 50% | 30% | 20% | 0% | 0% |
| 3 | Transparency & Co. | OSF | $250,000 | 0% | 10% | 50% | 20% | 20% |

Additional information could be added into any visualisation, including a breakdown of where organisations are funded by multiple donors, or where organisations that are cross cutting fit horizontally.

Of course, this framework has many more potential applications than mapping funding.  For example, a number of leading global open data organizations are already starting to map their own areas of expertise and programming (categorizing by research, advocacy, capacity building, etc.) to improve coordination and recognize respective comparative advantages. This could be another means to give granularity to their self-assessment to inform understanding of current **field-wide capacity and resources**.

Another potential application would be at **country level** to inform assessment of where the greatest gaps/needs are of accountability actors to harness data effectively. As TAI members consider the potential for collaborative engagements at country level this is one tool that can be tested.