



Building Evidence in Education Unlocking data to tell the story of Education in Africa: Moving forward together

14th October 2020 Deep Dive Summary Report

1. Overview

On the 14th of October, as part of the Building Evidence Into Education (BE2) annual meeting programme, ZiziAfrique, ESSA and the EdTech Hub ran a Deep Dive on “Unlocking data to tell the story of Education in Africa: Moving forward together”. As members of the BE2 donor working group, Dubai Cares endorsed and hosted this session which convened funders, practitioners, and researchers to discuss challenges and opportunities for increasing the accessibility and usability of data in research and programming in Africa.

This event followed on from two previous workshops. One on the 21st July 2020, to gather input on how we can work together to unlock data about education in Africa. Following this, a second workshop was held on the 8th of October 2020 to discuss the formation of a community of practice to ‘unlock’ education data in Africa.

For the Deep Dive, we had four objectives:

1. Present findings from the July workshop and analysis of data use by African education researchers.
2. Share the experiences of members on challenges and successes in increasing access to data.
3. Identify meaningful actions that funders can take to increase access to education data.
4. Explore links to other initiatives.

Key takeaways

From the many perspectives offered during the Deep Dive session, three key takeaways emerged across the conversations taking place:

1. **Advocacy for unlocking data** is a necessary effort that requires making the case for open data by synthesizing the latest evidence about why it is needed.
2. A lot of data is actually accessible, however there is little awareness about its availability and how to go about finding it. It would be helpful to **conduct a mapping exercise** of funders' existing available datasets.
3. It is important to **differentiate between different kinds of available data** (e.g. evaluation data, implementation data, etc.) in order to understand how different data types can be helpful to different stakeholders, and to ensure that data is made available in ways that suit their needs.

2. Summary of Presentation

Findings from the previous workshops were presented at BE2 as an input to stimulate discussion.

2.1. Understanding the data ecosystem

One important barrier to action is that we don't have a full picture of the data ecosystem as it stands now. Before deciding where to go, we need to properly take stock of where we are. This means fulfilling a number of key tasks:



Funders are a critical part of building this ecosystem: what funders do directly affects how data is accessed and how it is used. In order to understand and engage in effective advocacy with funders, we need to understand:

- Who is funded to collect data;
- The rules funders have around data sharing;
- How funders judge the quality of research outputs;
- How funders promote capacity exchange for local researchers; and
- How funders support ministries to collect administrative data.

2.2. Good and less-good practices

At the previous workshops, participants were invited to offer examples of good and bad practices that they had seen or experienced from funders in relation to data sharing.

A resoundingly positive example given was in the use of Demographic and Health Surveys (DHS). The data produced by these surveys is available on an open platform and is being used widely by scholars, national governments, and international agencies to carry out country-level and multi country analyses. Another positive example was the support provided by the Hewlett Foundation to Uwezo with learning assessments in East Africa.

Examples of less-good practices included weak support to cross-national learning assessments such as PASEC and SACMEQ, as well as fragmented support to project-based initiatives around data, turning data production and sharing into an addendum to research, rather than an objective. A further concern raised was an 'extractive' relationship with local researchers: the engagement of these researchers to produce data which was then analysed elsewhere was perceived to be breeding mistrust between them and international researchers and funders.

2.3. Potential solutions

At the previous workshops, ways in which funders could improve access to and use of data was discussed. Four solutions to this problem were presented:

- **Invest in making your own data accessible, findable, and usable.** A lot of funders have policies which stipulate data sharing, but these are frequently not adhered to, and researchers are rarely funded towards the tail end of a programme to ensure that the data is accessible. This should not be the norm.
- **Connect with local scholars and support them to use your data.** Go beyond the usual suspects and build wider communities of data sharing and use.

- **Incentivise open access in others.** For example, include open data sets as quality research outputs and Key Performance Indicators alongside publications.
- **Support greater capacity exchange on data use and access.** This needs to be done in a systematic way, rather than a piecemeal, project-by-project approach.

3. Break-out groups

The session participants then divided into groups, to discuss two questions.

1. How open would you like the data you have funded the collection of to be?
2. What help do you need to increase access to this data?

A final question was also asked of all participants: would you take up any of the solutions proposed in the introduction, and why?

3.1. How open would you like the data you have funded the collection of to be?

In some instances, data openness is provided on a platform (e.g. USAID's Development Experience Clearinghouse platform, or EduLinks). However, even within these platforms it can be difficult to find specific datasets without very specific search language. Even if files are available, they are frequently seen as a box-ticking contractual requirement, rather than a genuine opportunity to make data accessible.

Participants felt that a useful exercise would be a larger-scale mapping exercise so that people are to be able to pursue projects or research implementation with a clearer idea of what has come before. This would allow research to genuinely build on the existing research base, rather than starting from scratch with new tools and data. It would also require efforts to ensure consistency in the collection of data to ensure its comparability.

Another point noted was that once data is collected, it is frequently returned to countries in the global north / west for analysis, before results are then redistributed to local teams for dissemination. Participants felt it necessary to bring local stakeholders into analysis plans, including building capacity to perform analysis in the location where the data is collected. Anecdotal evidence suggests that where this is done, it increases stakeholder buy-in, which makes for better dissemination of data.

3.2. What help do you need to increase access to this data?

One of the first challenges raised related to **data management within funding institutions**. In some institutions, this is very decentralised, with no requirement for implementing partners to provide the data back to a centrally-located repository. This creates problems not only for external implementing partners or at the country level, but also within the institutions themselves, with internal stakeholders struggling to access datasets. A potential solution is to get data onto a central, open-source platform, with different access levels appropriate to the required security for each data type (e.g. surveys, biometrics etc.).

From an implementation perspective, a common issue is that data is hard to get hold of, in part because of **concerns around data sensitivity from funders**. However, this drive for accessibility and user-friendliness is often set in contrast to the need to ensure that data is high-quality and collected in accordance with proper standards. Tied up into these conversations are concerns about ensuring the replicability of findings, as well as protecting the reputation of the organisations funding and carrying out research.

One proposed suggestion to alleviate these interconnected issues was the adaptation of protocols regarding study registration, **creating alignment between organisations around a common 'data protocol'**. Greater uniformity among what is expected of 'high-quality' data, shared earlier in the design process, will lead to greater comparability of data outputs. In order to achieve this and increase data complementarity, more effort needs to be put into finding and working on areas of overlap for common working between donors. From the discussions it was evident that there is appetite for this, but that significant will is required to overcome the attendant bureaucratic hurdles to collaboration. Another key consideration is to ensure that researchers and practitioners have a clear understanding of who their audience is: who will be accessing and using their data? Data is not solely the preserve of analysts: accessibility to data should be commensurate to its intended audience. This will in turn inform the way the data is presented (summary tables, dashboards, raw datasets etc.).

One final point stressed by participants was that when collaborating on data, the primary concern should be focused on quality, not resources. We need to get the best people to contribute, not just the people with the means to contribute.

3.3. Would you take up any of the solutions proposed in the introduction, and why?

Participants were unanimous in endorsing the principle of increasing local scholarly involvement in data collection, analysis and use. However, this was not without the recognition of nuances in this approach. One participant highlighted the need for precision in distinguishing between data for results (e.g. aggregated for dashboards that doesn't allow for new analysis) and 'microdata' that could be much more useful but has more associated issues (e.g. around confidentiality etc.).

There was also agreement on the need for investment, critically for **joining up and giving a fuller picture of existing data.** Examples were noted of mapping exercises allowing for a user-friendly visualisation of what data is available, as well as key portals (such as that of IIEP) for accessing tools and policies. It was agreed that this was a funding priority to which all funders should offer greater contribution.

One important reflection was the need to acknowledge that, with greater amounts of data and greater access to that data comes the increased likelihood of conflicting messages between people providing and using data. **People need to be prepared to deal with these uncertainties surrounding data.** Another concern, expressed by one of the funders present at the meeting, involved balancing the desire for increased openness of data against legislative requirements around privacy and data protection (such as GDPR).

4. Next steps

Attendees of the session were enthusiastic about the potential benefits of investing more in making their own data more accessible and the idea of a centralised platform came up more than once. The following ideas for collaboration were put forward:

- Create a central data platform within our organizations;
- Develop an understanding of data users' needs in order to present data in a way that suits them;
- Develop a protocol for data quality;
- Develop ethics and best practices guidelines;
- Advocate for openly licensed data;
- Contribute to user-friendly maps of available datasets;

- Connect and work with local scholars; and
- Review contract templates to ensure they support openly licensed data.

The next step is to follow up with the most active funders in the workshop and assess their appetite for piloting one of the approaches above.

If you are interested to learn more or pilot one of these approaches please contact Lucy Heady (lucy@essa-africa.org) Björn Haßler (bjoern@edtechhub.org) or John Mugo (jmugo@ziziafrique.org).

5. Attendance

The BE2 session was attended by several representatives from USAID, the Lego Foundation, UNESCO, UNICEF, Dubai Cares, INEE, and the Agha Khan Development Network.