



EVALSDGs INSIGHT # 9: Using Big Data to Strengthen Evaluation of the SDGs



PURPOSE: This Insight EI#9, builds upon EVALSDGs Insight #8, which introduced key challenges for evaluating the SDGs concluding that most conventional evaluation methods are not well-suited to address the economic, logistical, methodological, and organizational challenges associated with most SDG evaluations. A key challenge, particularly for many poor countries, is that most SDGs require larger volumes, and more varied types of data than many national and local evaluation agencies have the capacity to collect and analyze.

In this Insight (EI #9), we take a look at four examples shown in the table below that illustrate how big data is being used to address some of these methodological challenges and strengthen SDG evaluation. We conclude by suggesting some next steps for assessing the wider potential of big data for SDG and other kinds of complex development evaluations.

EXAMPLES OF SOME BIG DATA TOOLS AND TECHNIQUES USED IN FOUR CASE STUDIES

Case study	Big data tools and techniques
SDG 4: USAID “Let’s Read” education program, Zambia.	Using dashboards to integrate and present large volumes of data from multiple sources
SDG 5: Assessing the effectiveness of UN Women’s programs to promote political participation of women in Mexico and Pakistan.	Using social media analytics to analyze Twitter, Facebook and radio call-in programs
SDG 8: Harambee’s employment programs for vulnerable youth in South Africa.	Using machine learning and artificial intelligence to integrate data from multiple agencies. Using mapping technology to track taxi routes.
SDG 15: GEF evaluations of the impact of forest protection programs.	Satellites, drones and remote sensors to collect information covering large areas and over long periods of time. Propensity score matching to create quasi-experimental evaluation designs

[Example for SDG 4 - Ensure inclusive and equitable quality education](#)

The **USAID “Let’s Read”** performance tracking dashboard systems in Zambia illustrate the use of dashboards to report on a wide range of large educational data sets in a format easily accessible to education policy makers, managers and teachers. The Zambian Ministry of Education collects information on almost 850,000 learners in nearly 5,000 schools, with the goal of improving the performance of the education system through more efficient school examination and grading systems. However, the data was previously underutilized because its volume was too great to efficiently handle with conventional manual analysis. Consequently, the Ministry was not achieving the goal of improving the education system through examination and grading systems. In response, **dashboards and data visualization systems** (presenting complex data in visually attractive and easily understood graphs and maps), were utilized to help organize, analyze, present and (most importantly) use

previously under-utilized data. The dashboard system became a valuable tool to detect missing data and inconsistencies and to improve the efficiency in the examination system.

Example for SDG 5 - Achieve gender equality and empower all women and girls

The electoral processes in Mexico and Pakistan illustrates the use of social media analysis to assess the effectiveness of UN Women's programs promoting women's political participation in both countries. Social media analysis was applied to Twitter (Mexico), Facebook and Radio call-in programs (Pakistan) to identify women's opinions on, and registration to vote in the Presidential elections in these two countries. Social media analysis makes it possible to examine posts and tweets on the election process of very large samples of women throughout the country and to track changes in opinions and participation rates over the electoral process. Both **Twitter and Facebooks (as well as other social media sites)** provide free use of a wide range of analytical tools that make this a powerful and relatively economic research tool for studying a wide range of social, political and economic issues and trends. It is also possible to conduct organizational analysis through techniques such as social network analysis. However, social media analysis is very different from conventional evaluation methods and it is necessary to fully understand some of the challenges such as selection bias (only certain groups use these social media outlets), and to understand how to interpret the "meaning" of different comments.

Example for SDG 8 - Promote sustained, inclusive and sustainable economic growth and productive employment and decent work for all

The Harambee youth employment program in South Africa illustrates the use of **machine learning and artificial intelligence [AI]** to promote access to employment for vulnerable youth. Youth from slums and informal settlements face many barriers to employment that go beyond their job qualifications. While many agencies, (e.g. government, firms and civil society organizations), collect information on these challenges, the information is dispersed. Also, no information has been systematically collected on the major challenges for youth living in remote areas of the city resulting from a dis-functional transport system – largely relying on private taxis – which disqualify many potential job applicants due to the huge time and cost of getting to jobs. Harambee used machine learning and AI to integrate all the data sources and map travel patterns to provide guidance to youth on what jobs could be reached within a reasonable amount of time and also to work with taxis to plan more rationale routes. The analysis provided vulnerable youth with more realistic guidance on job seeking strategies, and it also helped job placement firms to better understand the real-life challenges that qualified job seekers confront.

Example for SDG 15 - Protect, restore and promote sustainable use of terrestrial ecosystems

The GEF (Global Environment Facility) programs illustrate the use of **satellites, drones and remote sensors** to evaluate the effectiveness of GEF programs to protect forest cover in protected forest areas. Many donors and governments have created protected areas of land and water to conserve forests, other fragile land areas, oceans and areas such as mangroves. These are often very large and inaccessible areas and therefore difficult to monitor and evaluate. In response, GEF has utilized satellites, drones and remote sensors to assess the effects of donor and government interventions on protected forest areas. These rapidly evolving technologies make it possible to: generate time series data (sometimes more than 20 years) to monitor changes in forest cover; and to collect data on the economic, infrastructure and ecological characteristics of large areas surrounding the protected zones. For evaluation, the large quantities of data allow selection of carefully matched comparison areas (propensity score matching) to permit the use of pretest-post-test comparison group designs – as well as more sophisticated evaluation approaches. Geospatial technology combined with the ability to economically collect and analyze huge volumes of data covering large areas and long periods of time, will be essential for tracking and evaluating climate change and the effectiveness of efforts to addresses such complex problems.

What You Can Do

Big data offers many exciting, accessible and affordable techniques to allow a wide range of development agencies, CSOs and other actors to address the scope and complexity of SDG evaluations. The examples presented here provide examples of the possibilities Big Data presents for evaluation. Evaluators have been slow to integrate big data in their evaluations and it is time to rethink conventional evaluation approaches. Some initial steps for funders, development agencies and evaluation practitioners to consider include:

- Becoming familiar with big data tools and techniques
- Reviewing how big data techniques are already employed in evaluations and research related to the intervention area
- Critically assess the weaknesses and limitations of current evaluation approaches and identify ways in which big data could fill some of these gaps.
- Identify pilot evaluation projects to test the benefits (and drawbacks) of big data and to assess the value-added of these approaches.
- Considering that big data is not a magical panacea for all evaluation challenges, it is critical (as with any methodology) to be cautious of the hype and over-ambitious claims of some big data advocates, and instead to recognize its limitations and issues and concerns surrounding its use.

About EVALSDGs INSIGHTS

EVALSDGs is a global network committed to adding value and learning to the SDGs, as well as support processes to integrate evaluation into global and national SDGs review systems. EVALSDGs Guidance Group is a sub-group of EVALSDGs made up of evaluators ready to support you in the evaluation of the 17 SDGs which is complex and requires strong collaborations and partnerships. It produces INSIGHT publications, training materials, and information on opportunities to strengthen global and national evaluation capacity. It shares good cookies around the impact and sustainability of the SDGs and suggests which competencies you need. EVALSDGs INSIGHTS are short, light and easy to digest notes on topics related to evaluation and the SDGs. They present ideas and new information and stimulate thinking to strengthen evaluation capacity. This EI has been jointly produced by UNITAR and EVALSDGs Guidance Group.

Author: Micheal Bamberger

Editorial Lead: Scott Chaplowe

Contact us to provide your feedback or for more information:

Isha Wedasinghe Miranda, Sri Lanka Evaluation Association (SLEVA): ishamiranda2011@gmail.com

Hayat Askar, Jordan Development Evaluation Association (EvalJordan Vice President): hayat.askar@gmail.com

Katinka Koke, UNITAR, Specialist, Planning, Performance M&E Unit / Division for Strategic Planning and Performance: katinka.koke@unitar.org