Big Data to Data Science

Moving from “What” to “How” in the MERL Tech Space

Kecia Bertermann, Alexandra Robinson, Michael Bamberger, Grace Lyn Higdon, and Linda Raftree

3

May 2020

Contents

Big Data to Data Science: Moving from "What" to "How" in the MERL Tech Space

Authors: Kecia Bertermann, Alexandra Robinson, Michael Bamberger, Grace Lyn Higdon, and Linda Raftree

Abstract

This paper probes trends in the use of big data by a community of early adopters working in monitoring, evaluation, research, and learning (MERL) in the development and humanitarian sectors. Qualitative analysis was conducted on data from MERL Tech conference records and key informant interviews. Findings indicate that MERL practitioners are in a fragmented, experimental phase, with use and application of big data varying widely, accompanied by shifting terminologies. We take an in-depth look at barriers to and enablers of use of big data within MERL, as well as benefits and drawbacks. Concerns about bias, privacy, and the potential for big data to magnify existing inequalities arose frequently. The research surfaced a need for more systematic and broader sharing of big data use cases and case studies in the development sector.

Introduction

Purpose

New information, new methodologies, and new thinking periodically shift the boundaries of monitoring, evaluation, research, and learning (MERL). In recent years, the emergence of new information technologies — specifically, big data — has caused such a shift. Big data is often differentiated from other data due to its volume, velocity, and variety. Some MERL practitioners have recently begun to use big data and, increasingly, data science and other methodologies to incorporate big data into various stages in the identification, design, management, monitoring, evaluation, reporting, and dissemination of their development and humanitarian programs.[1]

This paper aims to probe trends in use of big data by monitoring and evaluation (M&E) specialists and other development practitioners who are part of the MERL Tech community[2] and to take stock of the community's use of big data. Recognizing that the MERL Tech community includes many early adopters of new tech and data analysis techniques, the authors also aim to describe the extent to which humanitarian and development organizations are beginning to use big data and data science for MERL. The paper focuses on scenarios and sectors in which big data is used, and describes the barriers and enablers of use, as well as benefits and limitations.[3]

What is MERL Tech?

MERL Tech is a platform and space for discussion, learning, and sharing experiences and challenges with the use of technologies for MERL in the social impact, humanitarian, and international development fields. MERL Tech aims to strengthen understanding of the value, impact, and risks of digital technology in MERL; support learning and discussion on new approaches and tools for MERL work; and strengthen the evidence base and learning around technology used in MERL and technology in development.

[1] *Data science* broadly refers to the use of methods and algorithms to glean insights from data sets. While big data is part of that discipline, in this paper we use the term *big data* as it is more generally understood outside the field of data science.

[2] The MERL Tech community includes a wide range of expertise, including researchers, evaluators, tech specialists, academics, and data scientists. Most are connected to development or humanitarian work through NGOs or INGOs or as independent consultants.

[3] Everett Rogers (1962). *Diffusion of Innovations* (The Free Press, New York, NY).

What is Big Data?

Big data has huge volume, is generated very quickly and often in real time, is too large to be analyzed on a single computer, and, in most cases, is collected continuously — often over long periods. It is also nonreactive. This means that, in most cases, the data has already been collected digitally for a different purpose, so the information is not affected by its use for an evaluation.

These characteristics contrast with conventional surveys, focus group discussions, and other kinds of evaluation data, where the collection process can influence the information provided. For example, if respondents are asked about their sources of income, they may inflate or reduce the information they provide depending on what they perceive to be the purpose of the study.

Big data approaches have the potential to obtain data on almost an entire population of interest rather than a relatively small sample. This increases the granularity of the data and makes it possible to compare small groups or even individuals.

Source: The authors, drawing on Salganik (2018),[4] Petersson and Breul (2017),[5] and Garcia and Kotturi (2020)[6]

[4] Matthew Salganik (2018). *Social Research in the Digital Age: Bit by Bit* (Princeton University Press, New Jersey).
[5] Gustav Petersson and Jonathan Breul, eds. (2017). *Cyber Society, Big Data and Evaluation* (Transaction Publishers, London, UK).
[6] Oscar Garcia and Prashanth Kotturi, eds. (2020). *Information and Communication Technologies for Development Evaluation* (Routledge, New York, NY).

Methods

The research team conducted qualitative analysis and limited quantitative analysis. First, we collated administrative and publicly available MERL Tech conference data from the 281 sessions accepted for presentation between 2015 and 2018. This included sessions at five conferences in Washington, D.C., two in London, and one in Johannesburg. We processed and coded the data to analyze data collection, management, analysis, and policy trends and to explore key themes such as data ethics, privacy, and inclusion. We identified 54 out of 281 sessions that mentioned big data and compared trends between sessions that did and did not mention this topic.

We also conducted a limited review of all sessions submitted to the MERL Tech conferences between 2015 and 2019 and counted the number of sessions that mentioned big data, data science, data scientist, artificial intelligence/AI, and machine learning to get a sense of frequency and trends in use of this terminology.

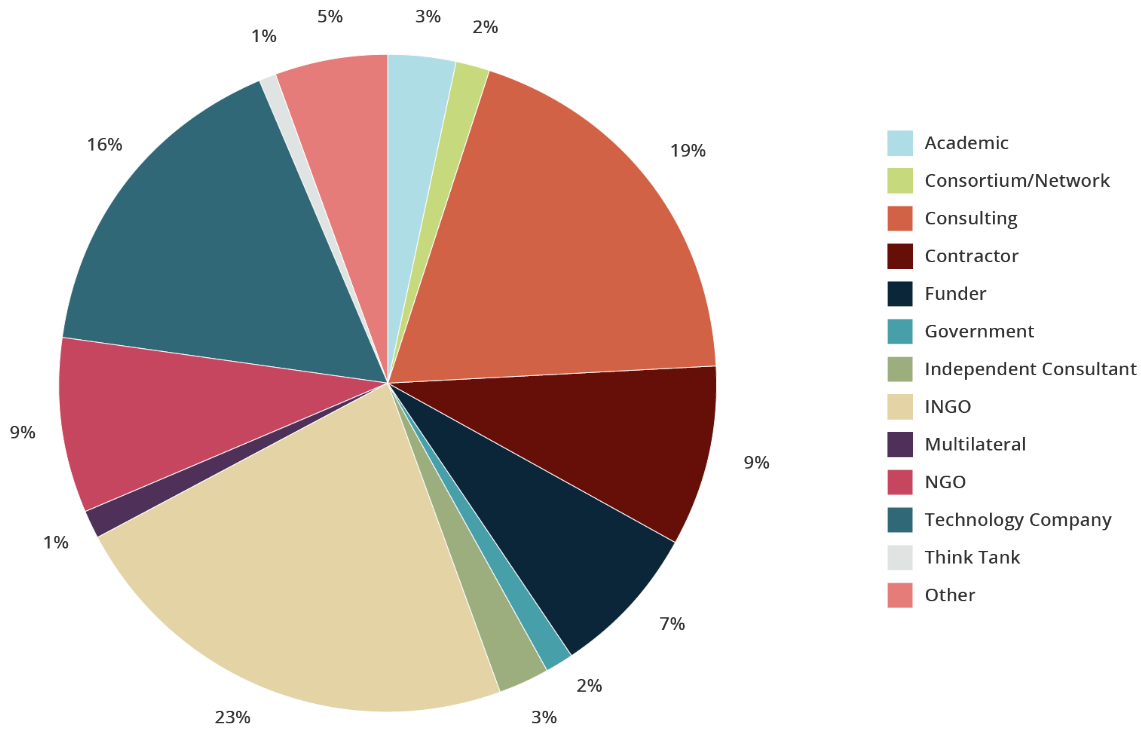
This quantitative analysis was complemented by 11 qualitative key informant interviews. We selected interviewees representing diverse viewpoints (implementers, donors, MERL specialists) and a range of subject matter expertise and backgrounds. During interviews, we explored why an interviewee chose to use big data, the benefits and challenges of using big data, reflections on the use of big data in the wider MERL tech community, and opportunities for the future.

Limitations

This study is limited to the use of big data in the MERL Tech community, which includes people and organizations that self-selected to join the community due to their interest in the role of technology and digital data in MERL (Figure 1 is a breakdown of the types of organizations present at MERL Tech conferences). MERL Tech conference participants include individuals and representatives of think tanks, technology companies, national and international non-governmental organizations, multilateral and bilateral agencies, independent consultants, governments, funders, contractors, consulting agencies, consortia and networks, and academic institutions.[7] Because participants are self-selected, this research does not necessarily represent the interests and viewpoints of the wider MERL community with regard to big data and its role in monitoring, evaluation, research and learning.

[7] The conferences in Washington, D.C., and London mainly attracted representatives from North America and Europe, while most participants in the Johannesburg conference were working in South Africa.

Figure 1: Types of Organizations Present at MERL Tech Conferences, 2014–2019



The analysis of session content is limited to conference proposals that the selection committee selected to reflect themes chosen by the conference organizers according to trends observed in the MERL Tech space. The committee also selected sessions to address diversity, equity, and inclusion considerations. The conference sessions do not necessarily reflect the content of all submissions.

This study is also limited to MERL Tech conference records, which improved in quality and completeness as the MERL Tech community grew and records were gathered more systematically. The quantitative analysis is limited to speakers' prior descriptions of their sessions and does not necessarily capture the depth and breadth of live conversations about big data in all sessions at the MERL Tech conferences. Finally, only the sessions included in the conference program were included in analysis of session content. Given these limitations, we did not conduct statistical analysis to compare trends by year or conference location. Instead, descriptive statistics provide a general picture of the MERL Tech community's evolving use of and discussions about technology over time.

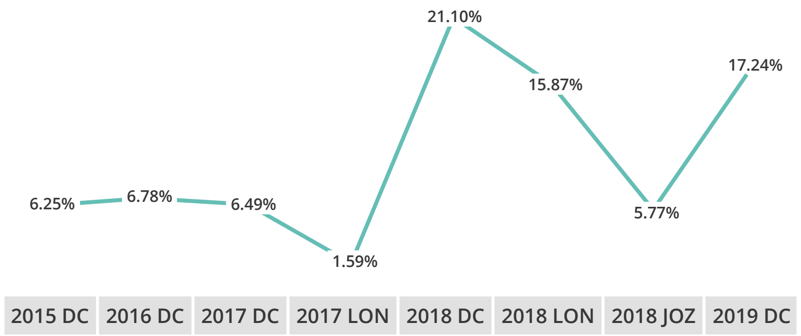
Findings

Frequency of Big Data–Related Terms Mentioned in Submitted MERL Tech Conference Sessions

Submitted sessions.

Our analysis of the terms big data, data science, data scientist, artificial intelligence/AI, and machine learning in sessions submitted to the MERL Tech conferences between 2015 and 2019 (see Figure 2) showed a large jump in 2018 in the conferences in Washington, D.C. (from 6.49 percent of submissions in 2017 to 21.10 percent in 2018 and 17.24 percent in 2019). There was also a spike between 2017 and 2018 at conferences in London (where 1.59 percent of sessions mentioned big data or related terms in 2017, compared to 15.87 percent in 2018).

Figure 2: Percentage of MERL Tech Conference Submissions that Mentioned *Big Data, Data Science, Data Scientist, Artificial Intelligence/AI, or Machine Learning*



Accepted sessions.

In any given year from 2015 to 2019, 16 percent to 26 percent of sessions at MERL Tech conferences were related to the topic of big data. As noted below in our discussion of methodology, this number is fairly consistent across the years because the conference steering committee tried to balance sessions across various themes, including that of big data.

One key informant observed that the term *big data* has gone out of fashion and that many data scientists now prefer to use *artificial intelligence* or *data mining* in their presentation titles. Analysis of the 64 sessions at the 2019 MERL Tech Conference in Washington, D.C., found that seven session titles included *big data*, seven included *machine learning*, and two included *artificial intelligence*. While quite small, these numbers suggest a shift in use of terms relating to data science and big data over time.

Types of Big Data and Data Science Approaches Used for MERL

Key informant interviews and conference session analysis identified four main types of technologies used to collect big data: satellites, remote sensors, mobile technology, and M&E platforms, as well as a number of other tools and methods.[8] M&E platforms were discussed most frequently in the earliest years of the MERL Tech Conference but steadily dropped off over the five-year period we examined. Mobile technology was referenced most often (32 percent) and most consistently in conference descriptions between 2015 and 2018. Satellites used to collect big data were mentioned most frequently, with the steady influx of data during collection cited as a major benefit. Conference data reflected this trend, with satellites the most frequently mentioned data collection tool. Interestingly, 20 percent of MERL Tech Conference sessions involving big data included mentions of satellite data, compared to only 4 percent of sessions that did not involve big data.

Types of Big Data

Big data is often differentiated from other data by its volume, velocity, and variety. However, there is no widely accepted classification of types of big data. One approach is to classify big data into:

Human-generated data

deriving from consumer-focused applications and platforms, sensors, the Internet of Things (e.g., consumption of water, use of public toilets, people entering and leaving community buildings), and surveys.

Administrative (transactional) data

deriving from program administrative data, government and other public records, and monitoring data from service agencies. It is now possible to analyze the huge volume of PDF files that most agencies generate.

Geospatial data and remote sensors

from satellites and drones, and from remote sensor data from non-human sources on things like soil moisture levels and crop health.

Source: Adapted from York, Bamberger and Olazabal 2020.[9]

[8] Other data collection methods mentioned included blockchain, call-in radio, drones, specialized data collection applications, and websites. Other data analysis methods mentioned included open source programming, M&E packages, and cloud storage.

[9] Pete York, Michael Bamberger and Veronica Olazabal (2020). Measuring Results and Impact in the Age of Big Data (Rockefeller Foundation, New York, NY.)

Conference session analysis and key informant interviews revealed six main types of tools used to analyze big data: artificial intelligence and machine learning, geospatial analysis, data mining, data visualization, data analysis software packages, and social network analysis. Artificial intelligence and machine learning did not appear in MERL Tech Conference descriptions until 2017, but these terms were commonly used at MERL Tech 2019. Interviewees also underscored the predominance of artificial intelligence and machine learning in the current state of the field. One data scientist commented, "When we say 'emerging technology,' the lion's share of that at the moment is looking at artificial intelligence and machine learning." Another participant suggested that the shift in terminology implied rebranding rather than changing analytic techniques:

"... Big data is sort of an outmoded buzzword that everyone was using in 2014 and 2015. Now it's sort of dropped off the radar, and people who are actively working in the field don't use that term very much anymore. It's kind of a matter of rebranding. People who used to be big data people are now AI specialists and in a couple of years will be calling themselves something else."

Applications of Big Data for MERL

Monitoring.

This was the area most frequently mentioned in accepted sessions and in key informant interviews, as big data approaches offer ways to collect a wider range of indicators over broader geographical areas. They are now beginning to provide longitudinal data sets that may cover up to 20 years and can continue to generate data after a project closes. Data analytics also provides valuable tools for constructing and managing integrated data platforms, making it possible to identify previously undetectable patterns and relationships among different data sets.

Evaluation.

Evaluation. While key informant interviewees noted that evaluators do not widely use the term big data, several respondents provided examples of current and potential applications for evaluation. They identified four kinds of benefits:

- The ability to collect information more cheaply and quickly over longer periods of time and in ways that include difficult-to-reach groups.
- Increased sample sizes (often covering a total target population), making it possible to conduct more disaggregated and granular analysis.
- New kinds of experimental and quasi-experimental designs that are now possible.
- The possibility of using more sophisticated analytical tools.

Research.

The reasons for applying big data to research mirrored those cited for using big data for evaluation. References to how big data can strengthen learning were rare, although some of the benefits listed above are likely to result from strengthened M&E. To illustrate how big data is being used for MERL, the authors selected several sample cases that represent key technologies and analysis methods across diverse sectors. These appear at the end of this paper, beginning on page 24.

Learning.

There were a number of ways in which big data, particularly data analytics, can strengthen learning. These include data visualization, on-line tutorials and interactive discussions, personalizing learning, and gathering real time feedback from remote areas.

Operations.

Respondents cited a number of ways in which big data can support program operations (for example, Case Studies 5 and 9). These do not fall exclusively into any of the four MERL areas but are relevant to this discussion because one of the objectives of MERL is to strengthen program operations. Examples of operations support fall into several categories: 1) developing, testing, and helping implement new technologies; 2) identifying the best locations for health centers and other services; and 3) using integrated data platforms that make program information easily accessible to all stakeholders and that provide a centralized point for obtaining information on a wide range of services. A number of agencies are also developing algorithms for automated decision-making on, for example, the most effective strategies to respond to infectious diseases (for example, Case Study 2).

In particular, interviewees discussed applications of big data in MERL and program operations in the agriculture, environmental, and health sectors. In one example from agriculture, remote sensors and satellites were used in the design stage to conduct a scoping assessment that identified fertilizer needs. Remote sensors and satellites have also been used to monitor land use and crop production. The environmental sector has seen big data applied to monitor crop health and water conditions and to model and predict climate outcomes. One project used earth observation monitoring data to track where deforestation occurs and combined the information with earth observation data on land use patterns to identify areas at highest risk of deforestation (Case Study 1). In the health sector, call data records, electronic medical records, and other data collected via mobile technologies have been used to monitor health crises (Case Study 2), conduct longitudinal health monitoring, and inform national health planning. One interviewee cited a case in which 4 million emergency medical records were digitized for a health program evaluation. Overall, respondents believed that big data was most effectively applied in these three sectors, which have more traditionally relied on quantitative data and, as one data scientist commented, “where it’s easy to come up with a fairly simple automated decision process in a way that makes sense.”

In contrast, sectors such as democracy, governance, human rights, and cross-cutting themes like gender may be underrepresented in big data applications due to the complex problems that arise in those sectors and the risks of using big data to address such issues. These risks include automating or entrenching bias into government decision-making around issues that might include social benefits or surveillance systems. One interviewee argued that, outside of technical sectors such as health or agriculture “there are a lot of potential danger spots” in these less technical arenas that “we haven’t really figured out.”

As shown in Figure 3, there are a number of potential applications of big data for MERL.

Figure 3. Potential Applications of Big Data for MERL[10,11]

Monitoring	• Developing and operationalizing performance metrics • Managing large data sets • Using satellites and telephone records to monitor things that could not previously be observed • Developing new tools to monitor difficult-to-research areas such as democracy, governance, and human rights • Increasing the range and granularity of indicators • Introducing platforms that enable agencies to collect and organize all the information on their programs (for example, through data maps that gives partners easy access to all the data they collect) • Helping identify excluded groups • Tracking migration and food insecurity through use of real-time data • Reducing human bias • Deriving reliable time-series data • Monitoring land usage and crop production

[10] See also Petersson and Breul, op.cit.; “Evaluation in the age of big data” in Michael Bamberger and Linda Mabry, eds. (2019) in Realworld Evaluation: Working Under Budget, Time, Data and Political Constraints (Sage Publications, Thousand Oaks, California).

[11] These approaches were discussed in the panel presented by Jos Vaessen, Daniel de Garcia, and Michael Bamberger on using big data to strengthen the evaluation of complex programs at the 2019 MERL Tech Conference in Washington, D.C.

Evaluation	• Conducting larger and less costly surveys • Looking at results across time and across multiple organizations • Creating new kinds of experimental and quasi-experimental designs, including applying propensity score matching • Developing evaluation algorithms that make it possible to work with much larger data sets • Creating complexity-responsive evaluation designs that can apply systems analysis (systems mapping, systems dynamics, social network analysis, etc.) and other tools for analyzing complex systems • Evaluating the effectiveness of new technologies • Answering the question, “Did the program work?” • Assessing the value added by the new technology • Taking advantage of potential applications for artificial intelligence and machine learning • Addressing Sustainable Development Goal challenges across countries • In combination with other methods, strengthening validity by ground-truthing and bolstering the ease of conducting and validating findings of surveys
Research	• Sourcing innovative ideas to address health crises • Studying mobility and migration by identifying where people are located, and tracking movements • Predicting crop yields
Learning	• Using new data visualization technologies to communicate education material • Distributing online tutorials and fostering interactive discussion • Personalizing learning • Gathering real-time feedback from beneficiaries in remote areas
Cross-Cutting Support (general support for program operations that combine various MERL Tech areas)	• Identifying and testing new technologies to support decision-making and address health issues • Collecting and organizing program information (for example, through data maps that give partners access to information) • Mapping networks of people and organizations involved with a program and how they interact with one another • Selecting best locations for health centers and other services • Using data for decision-making and program or intervention design • Producing integrated digital platforms that provide households with information on different services

Perceived Benefits of Big Data for MERL

Analysis.

All respondents identified enhanced analytical possibilities as a benefit to use of big data for MERL. The ability to model impact with more sophisticated estimation, mine unstructured data, conduct continuous and longitudinal analysis, and process and analyze national- and state-level data all were seen as beneficial — especially when gaps in data sets prevent the use of traditional statistical methods. One interviewed said,

"The bigger the data, the more accurate. Bigger robust data produces better results. Social science is based on the understanding that we can't have data about everyone. Sampling is at the heart of their approach. With big data, that's less and less true … It's a profound change."

Key informants noted other benefits of big data — for example, its potential for improving efficiency and targeting and gaining remote insights in situations where an evaluator is not able to physically visit an evaluation site, as discussed below.

Improved efficiency.

Participants appreciated the use of big data to increase efficiency, scale, data quality, accuracy, and cost-effectiveness. For example, it is possible to merge different kinds of data (geospatial, PDF records, audio and photographic, phone records, digital financial records, and others) into integrated platforms to analyze patterns of association that could not previously be detected. Other examples of enhanced efficiency include the provision of continuous data in real time, the ability to fill in gaps in data sets or triangulate data points, reduced costs of data processing, and increased volumes of big data compared to traditional data sets. An evaluator explained,

"For us the biggest benefit is just the time saving and enabling us to do things we wouldn't otherwise be able to do with large-scale data … In some cases, we wouldn't be able to do it [manually] and in others it just made it faster."

Improved targeting.

Interviewees noted the potential for using big data to inform service delivery through improved targeting and value for money. “Aside from standard benefits of successful monitoring, research, evaluation, learning, [you have] more efficient services, more value back to the beneficiary, more value up to the funders,” one data scientist explained. However, because respondents did not provide specific examples of such benefits, it is unclear whether these potentials are being realized.

Remote insights.

Interviewees described the practical benefits that big data offers through remote monitoring of locations that are difficult to access due to conflict or topography. This kind of remote monitoring data increases teams’ insight into context when they cannot be on the ground. In one example, a respondent spoke of tracking population movements within and across countries. Participants repeatedly mentioned national- and state-level analysis as an important benefit of big data, particularly in the health sector.

Reduced bias.

Although many participants shared concerns about the potential for bias and its consequences in algorithmic modeling, some argued that big data can also mitigate human bias, particularly during data generation and analysis. Interviewees focused in particular on big data (in this case, satellite data) as an objective measure. Going further, one participant referred to “biases in the collection of data. People talk about biases in big data, but there are other biases in positionality. [Big data brings] objective measurement, which can be replicated and repeated.” (It should be noted that literature examining bias in big data in other sectors contests this claim.)[12] Case Study 4 discusses the challenges of selection bias in analyzing social media (Twitter, Facebook, and others) when only a portion of the population uses the platforms and those users tend not to represent the total population of interest.

[12] For example, Cathy O’Neil (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Broadway Books, New York, NY) and Virginia Eubanks (2017). *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor* (St. Martins Press, London).

Drawbacks and Limitations of Big Data in MERL

Key informants identified a number of challenges to use of big data, including lack of trust, misalignment of goals, ethical concerns, and barriers to adoption.

Lack of trust.

Several respondents working with NGOs or agencies with a grassroots focus stressed the importance of building and maintaining trust with local communities and other stakeholders. They noted that big data can erode this trust, as many communities do not understand how data is collected and used, who will see it, and, importantly, whether and how they know what happens to their personal data. Respondents considered this an important concern affecting communities' potential use of big data. One interviewee described a project's provision of "forecasts [generated by big data] to agronomists in East Africa, but agronomists … don't trust the forecasts because the project teams aren't on the ground." Another interviewee raised questions around national governments trusting big data sources:

> **"Often, governments want to use their own data to answer their questions. One of my questions would be, do governments trust big data? Do they trust insights they can glean from big data? Or is there more work that we would need to do to make sure that big data becomes a trusted source?"**

Such issues around trust often reflect the lack of direct connections with others that give data a human touch. Participants observed that careless management of big data with stakeholders may "become a barrier to connecting with people."

Participants acknowledged that the detachment of human subject matter experts can hinder the utility of big data, particularly in analyzing qualitative and unstructured data:

> **"Working with qualitative data — reading and coding transcripts of interviews and understanding the nuances of what is being said — still takes a lot of human interaction. When you start looking at the scale of big data, this becomes prohibitive. We have not overcome those challenges with technology yet. As great as [natural language processing] and large text analytics is, it's still not always as good for qualitative data as a human subject matter expert that understands how to work with qualitative data. When you start working at that scale, you lose effectiveness."**

Misalignment of goals.

Some interviewees described mismatches between the capacity and resources of the private and development sectors to generate and use big data, as well as differences in purpose — potential misalignments between what an organization wants to learn and services offered by the private sector. One interviewee articulated the tensions that can exist between services from the private sector and the need to answer pressing programmatic questions:

"Often the model for using big data to date has been private sector companies that have greater capacity and greater resources approaching folks in the development sector, particularly partner country governments saying, 'Hey, we have all this data for you. Let us analyze it and provide you with insights,' when ideally it should be the other way around. It should be country governments or development partners or implementers saying, 'Okay, these are the types of questions we have. Can we answer these questions through data that already exists? Or could big data provide some type of insight?' And have it be a question-driven approach instead of just shock and awe at times with the different types of insights that big data can provide. I do worry that sometimes big data insights don't answer real and pressing priority questions."

Ethical concerns.

Several respondents discussed data ethics related to big data, noting in particular concerns about bias, privacy, and the potential for big data to magnify existing inequalities. Some expressed anxiety that the power dynamics of big data will further marginalize vulnerable groups. One interviewee noted,

"[Big data] should be promoted but needs to be done in a responsible and ethical way which is not always at the forefront of everyone's mind when they start thinking about access and what they can do with the data. Be conscious that we work in settings where people are high-risk and often big data is collecting [personally identifiable information], sometimes collecting incriminating information, sometimes very sensitive. When we are thinking through humanitarian data, it is almost always people-centric. Not talking about machines we're collecting data on, almost always linked to humans. It needs to always be through the lens of 'How can we also protect and secure data in a way that isn't going to do harm?'. "

Interviewees also questioned the inclusiveness of big data and argued that bias, exacerbated by harmful algorithms, may have unintended negative consequences.

"People can get a bit carried away with the complexity of big data and the complexity of big data learnings and not be aware of or appreciate the potential shortcomings both in terms of analytical rigor, the potential biases in the data, as well as potential ethical or privacy implications that some big data collection or processing methods have."

The need for caution was a consistent theme in interviews. Many respondents commented on the potential for big data to enhance program-related monitoring, evaluation, and learning. At the same time, these big data practitioners acknowledged the unique role of algorithm developers in utilizing big data:

"I've argued in a number of contexts that the problems that society generates have become so complex that we're going to need machines to help us figure out optimal decisions, not optimal disease, but help us with decision-making. With the massive, massive, massive caveat that it matters who's building those algorithms, it matters who's putting that data together, it matters who's making decisions about how to use it, if vulnerable populations and marginalized populations are not involved in that meeting, there is a very, very good chance that that will it will end up becoming a weapon used against them in the future, right?"

In terms of ethical concerns, our interviews and the literature recognized related implications:

- It is often difficult for local communities to understand how big data will be used, who will have access to it, and what control local people have over the information they receive.[13]
- Many agencies do not fully understand the challenges of protecting the privacy of data, and hackers can easily breach many data security protocols.[14]
- Selection of populations for sampling may reflect biases. For this reason, a UN Women study did not examine Facebook use in Pakistan, as poorer households lack access (see Case Study 4).[15]

[13] Mentioned in interviews with our key informants.

[14] Highlighted in interviews with key informants.

[15] Claudia Abreu Lopes, Savita Bailur, and Giles Barton-Owen (2018). Can big data be used for evaluation? A UN Women feasibility study. https://www.unwomen.org/en/digital-library/publications/2018/4/can-big-data-be-used-for-evaluation

- Racial and social biases are often (but generally unintentionally) built into automated algorithms that banks and other lenders, educational institutions, police departments, and public welfare programs use to determine eligibility or identify groups to target.[16]
- New regulations, including the European Union's General Data Protection Regulation, restrict use of big data in some countries (for example, how private companies may market or sell data). Because regulation generally affects commercial entities, some non-profit organizations and evaluators express confusion about how they may or may not use data legally. And, because regulations do not always keep pace with new technologies, additional ethical questions may arise.[17]

Barriers to adoption.

Interviewees discussed a range of factors that they saw as hindering the use of big data. The most commonly cited barriers related to financial cost or return on investment. Interviewees mentioned challenges for non-profits wanting to invest in big data or, conversely, their inability to justify such investment. One respondent noted, "When we have limited project budgets that are tied to very specific activities, it is hard to take that broader approach to analysis." Some also mentioned that big data is often underfunded and called for donors to increase funding so organizations can adopt it.

"I don't think the donor community has fully embraced the power of data to the degree where we can really do a lot of the great potential ideas that we may have. The funding just isn't there."

Participants also commented on a lack of human resources and skills related to big data. Many analysts in the MERL Tech community are well-trained in traditional analysis, but big data requires new skills and approaches that are not yet typical. The skills gap raises the need for investment:

"It's expensive to pay people to analyze those data sets, to store and collect the data — they cost money and time."

Resourcing is a barrier, including for the internal systems and infrastructure needed to utilize big data. The appetite for using big data is not always matched by data systems. Interviewees discussed a lack of appropriate internal technology and systems. One noted, with regard to geographic information system (GIS) data,

"In our case it is a technology and use barrier. It is complicated to use, complicated to understand, it requires a high bandwidth to be able to download imagery and specific software to be able to process it."

[16] O'Neil, op. cit. and Eubanks, op. cit.

[17] Vera Franz, Ben Hayes, and Lucy Hannah (2020). Civil Society Organizations and General Data Protection Regulation Compliance, Open Society Foundations, https://reliefweb.int/sites/reliefweb.int/files/resources/civil-society-organizations-and-gdpr-compliance-20200210.pdf

Finally, some interviewees identified potential disconnects between interest in using big data and organizational relevance, noting challenges for non-profits to justify investing in big data. For example,

"When we have limited project budgets that are tied to very specific activities it is hard to take that broader approach to analysis."

Often organizations cannot describe the tangible value they would derive from big data and do not see clear, relevant use cases. When organizations are reluctant to use big data, and its use is underfunded, the lack of available use cases further exacerbates the perception of barriers related to big data.

"I haven't seen big data use cases in development that clarify how it could be most useful or how I could use it in my own work. So the result is that it is a big unknown and is difficult to tackle."

MERL-specific challenges

Figure 4 lists some challenges affecting the four areas of MERL Tech, including an overview of challenges in use of big data for general operational support.

Monitoring can be a major challenge in many developing countries, and particularly in fragile states, due to lack of data or lack of infrastructure for data analysis. Development agencies also may not have the technical expertise for this analysis. One respondent pointed out that there is only one postgraduate data science program on the African continent, and that most development agencies cannot match the salaries offered by the private sector to attract this scarce human resource.

An important constraint to evaluation is the relatively small overlap between most conventional evaluation training programs and the training of many data scientists. Consequently, many evaluators are not yet familiar with the analytical methods that data scientists use. This issue was confirmed in a separate study that surveyed 324 evaluators' knowledge of big data. Only 10 percent of those evaluators reported having used big data in their practice.[18] However, despite a general impression that data science analytical techniques are more sophisticated than those used by most evaluators, one respondent observed that current data analytics approaches lack the sophistication required to analyze complex systems:

"Data science is useful to address important, but simple, problems."

[18] One of the few empirical studies on this question found that, out of a sample of 324 respondents from LinkedIn sites for evaluators, more than half of respondents indicated that they did not know what big data was. Of those who were able to define big data, only about 10 percent had ever used big data in one of their evaluations. About one- third reported that they did not consider big data relevant for their evaluation work (Petersson and Breul, op. cit.).

Data science usually works with much larger data sets than those used in most evaluations. A data scientist working at an innovation center reported that program evaluation was not part of his mandate, and that he was helping to introduce new technology into programs.[19,20] He expected that a few years would pass before the organization would be asked to evaluate the effectiveness of the technologies and their contribution to achieving development goals such as reduced hunger or disease.

Issues related to research were similar to those for evaluation. Some researchers were unfamiliar with data science data collection and analysis techniques, and many research data sets were too small to justify use of big data.

The lack of use cases was a major constraint for organizing learning and training programs. Other challenges included the lack of studies documenting the practical value of big data to program management (return on investment and value addition) and many trainers' lack of familiarity of many trainers with big data methodologies.

Key informants identified a number of challenges related to adopting big data, as listed in Figure 4.

Figure 4: Challenges to Adoption of Big Data for MERL

Monitoring	• Agencies lack the infrastructure to collect big data • Data is not available in fragile states or low-income countries • Technical expertise for collection and analysis is unavailable
Evaluation	• Data scientists are not familiar with evaluative approaches • Management does not consider evaluation part of the mandate of the data or innovation office • Big data works with data sets that are too large for most evaluations • Big data approaches are not yet capable of analyzing complex systems
Research	• Some researchers are not familiar with analytical techniques of data science or the different tools that are available for collecting big data
Learning	• Only one data science master's program exists in Africa • Greater focus is needed on how big data adds value and increases return on investment • More use cases would help potential users understand what big data actually is, how it is used, and its potential benefits • Trainers are unfamiliar with big data tools and techniques
Operational Support	• Agencies are not convinced of the added value (return on investment) that big data brings

[19] Michael Bamberger (2017). *Integrating Big Data into the Evaluation of Development Programs* (UN Global Pulse, New York, NY). Bamberger reported that in a number of UN and other development agencies that were interviewed, interaction between the data and innovation centers and evaluation offices within the same agency were similarly low. However, many of the data centers were still very new, and their linkages to evaluation would likely increase over time.

[20] Ibid.

Conclusions and Next Steps

Monitoring and evaluation practice shifts and flexes over time. As a case in point, the traditional term *M&E* is often augmented with *research* and *learning* and sometimes *accountability* and *adaptation*, and industry acronyms may include *MERL*, *MEL*, *MEAL*, and other iterations, all reflecting the expanding boundaries of M&E practice. During the past few years, we have also seen rapid changes in the big data arena, which has expanded beyond the use of GIS to new applications and new sectors and to explore multiple data sets. Like the evolution of terminology and practice surrounding M&E, terminology related to big data is evolving. Descriptors like *data science*, *artificial intelligence*, and *machine learning* are emerging and sometimes take the place of *big data*.

Our review of MERL Tech conference sessions and key informant interviews shows that rates of adoption and kinds of big data applications vary significantly among sectors (big data is widely adopted in agriculture, health, and education, but much less so in governance, human rights, and cross-cutting areas such as gender); regions (industrial nations compared to countries at different points on the development spectrum); and the kind and size of the organization. Therefore, caution is required when generalizing about trends in adoption and use of big data.

The literature on diffusion of innovations highlights stages of innovation and types of adopters.[21] Incorporation of big data into MERL follows a typical pattern, with early adopters the first to test and trial these approaches. As with other innovation pathways, MERL practitioners' use of big data varies widely. In these early days of use, MERL practitioners are exploring new applications of big data and testing its utility for monitoring, evaluation, and learning. The analysis in this paper recognizes that we are still in the nascent stages of big data use. Practitioners are experimenting, and there is little consistency in use and application of big data at this time.

During this early stage, such fragmentation might well be appropriate. Within the boundaries of responsible and ethical data use, it could be counterproductive to propose uniform definitions and approaches during this period of testing and adoption. Best practice is still emerging, and there is not yet enough standard practice to consolidate practices to consistently apply big data, particularly as the term has a wide range of sub-specialties (geospatial, social media, and others).

[21] Rogers, op. cit.

At the same time, our research indicates that use of big data is disjointed and could benefit from more systematic sharing of use cases and lessons learned. Even key informants who were using big data pointed to the need for more sharing about its use among MERL practitioners. The community requires use cases — moving beyond abstract conceptualizations and toward case studies that clearly explain the benefits and challenges of using big data and specific uses that are most appropriate for it. Use cases are particularly important to help organizations determine whether using big data provides sufficient return on investment and whether the benefits of use outweigh the costs. Case studies should be complemented by frameworks to help guide organizations that are new to big data in deciding when and whether to use big data in their research, learning, and evaluation.

An important development is the transition from big data as a way to reduce costs and time of data collection and broaden the range of data that can be collected to integration with data analytics. Data analytics makes it possible to organize and analyze multiple data sets and to conduct more sophisticated kinds of analysis, predictions and impact assessment.

We will likely see a growing use of data platforms that bring together a range of different data sources and allow the identification of patterns of association that could not previously be detected. The analysis of these new integrated data sets will be facilitated by the rapid expansion of data mining, machine learning and artificial intelligence.

A new challenge for MERL Tech will be creating ways to bring together the specialists working on data analytics (many of whom are not directly involved in the development field particularly at the grass-roots level) with the evaluation practitioners who are starting to create and use big data, but who are not yet familiar with the multiple tools and techniques for the analysis of the data.

As use of big data is further tested and applied among MERL practitioners, standard applications and practices will begin to emerge. We are already seeing thoughtful responses from practitioners guiding this movement, who raise important considerations for responsible big data use for MERL purposes. Importantly, we are moving beyond seeing big data as a shiny new object and are beginning to think more carefully about utility and ethics.

What is next?

The MERL community should actively engage in this space and lead the way in identifying further use cases for applying big data. The community can do this in six primary ways:

1. Considering.

Big data and data science are new additions to the MERL Tech toolbox, and it might be tempting to use them with limited consideration of why to do so. As with any new tool, it is important to first consider relevant learning questions and then decide whether big data is the best tool to answer those questions — or whether another source or method could answer them just as well.

2. Pilot testing.

Practitioners should seek opportunities to pilot test different applications of big data and to assess their utility and the value they add. Pilot testing should be collaborative; for example, an organization with strong roots at the field level might work with an agency that has technical expertise in relevant areas.

3. Documenting.

As noted above, the current body of documentation is insufficient to highlight relevant use cases and identify frameworks for determining return on investment in big data for MERL work. The community should do more to document efforts, experiences, successes, and failures in academic and gray literature.

4. Sharing.

As evidenced in this report, there is a hum of activity around big data in the vibrant MERL Tech community. Many practitioners are involved, although many are not aware of new applications or developments. We encourage the MERL Tech community to engage in fora such as communities of practice, salons, events, and other convenings, and to seek less typical avenues for sharing information and learning and to avoid knowledge silos.

5. Learning.

We encourage practitioners to continue to learn. The MERL Tech space is not static; indeed, the terminology and applications of big data have shifted rapidly and will continue to change over time. The MERL Tech community should participate in new training related to big data, continuing to apply critical thinking to new applications.

6. Guiding.

Big data practitioners are crossing exciting frontiers as they apply new methods to research and learning questions. These new opportunities bring significant responsibility. MERL Tech programs serve people who are often vulnerable — but whose rights and dignity deserve respect. As we move forward with using big data, we must carefully consider, implement, and share guidance for responsible use of these new applications, always honoring the people at the heart of our interventions.

Building Bridges Between Data Scientists and MERL Practitioners

While our study focuses on MERL Tech practitioners' use of big data, it is important to identify ways that the data science community can benefit from working with MERL practitioners.[22] New conversations among evaluators and data scientists acknowledge the potential to integrate their skills sets to better answer evaluation and learning questions. Indeed, some members of the MERL Tech community identify as both evaluators and data scientists. Conversations and convergence among these communities are taking place as evaluators and data scientists explore the possibilities and limits of their approaches. Building bridges between MERL practitioners and data scientists might include the following considerations:

- Theory-based evaluation, which includes theory of change, can strengthen data mining and artificial intelligence by providing guidelines for developing evaluation questions and a framework for interpreting findings. This is a hotly debated topic, as some advocates argue that artificial intelligence can identify and address such questions without the need for theory.

- Evaluation practitioners have developed tools to assess data quality and representativity and address such issues as construct validity. As data science often works with data of questionable quality or relevance (for example, the kinds of data used in social media analytics), these tools can add value across disciplines.

- Mixed methods and triangulation can support data quality and broaden the interpretation of findings.

- Data science often works with available data, without a framework for addressing issues of social exclusion, such as identifying groups that may be excluded from data sets.

- Evaluation specialists have developed numerous methodologies for assessing causality. Data science is starting to use some of these approaches (for example, digital experimentation and creating counterfactuals through artificial intelligence). However, evaluation can help strengthen what is often considered one of the weaker areas of data science.

- Finally, many evaluations use participatory approaches that involve affected populations in design, data collection, and sometimes analysis of evaluation findings. The investment of time and effort required to build trust with these populations is considered critical both on ethical grounds and to ensure reliable and meaningful data. Data scientists often collect data remotely and may have no contact with affected populations; thus, they may see no need to build trust with them. Several key informants expressed the opinion that erosion of trust is one of the greatest barriers to use of big data.

[22] See, for example, Petersson and Breul, op. cit.; Michael Bamberger (2019) "Evaluation in the age of big data," Chapter 18 in Bamberger and Mabry, op. cit. https://edge.sagepub.com/bamberger3e/student-resources-0

Case Studies

To illustrate how big data is being used for MERL, the authors selected several sample cases that represent key technologies and analysis methods across diverse sectors. Case studies that represented common use cases (such as types of technologies or analysis methods) seen in the case's particular sector were prioritized for inclusion in this paper. The cases were selected from key informant interviews and MERL Tech Conference sessions. The authors also referenced available reports related to these case studies. When neither conference historical data nor key informant interviews featured case studies for a particular sector, the authors looked to openly available reports.

Case Study 1

Sector: Environment
Title: Impact Evaluation of GEF Support to Protected Areas and Protected Area Systems
Organization: Global Environment Facility (GEF)
Source: MERL Tech State of the Field key informant interview
Types of data collection technology: Remote sensors and GIS imagery
Types of data analytics: Analysis of GIS and remote sensing data, machine learning **Document link:** GEF Report[23]

For what purpose was big data used and by whom?

The analysis evaluated the relevance, attributable impact, and cost-effectiveness of GEF's portfolio of 618 long-term biodiversity projects across 137 countries and 1,292 protected areas.

Methods

This quasi-experimental impact evaluation leveraged time-series remote sensor, satellite imagery, wildlife population, and development data to track longitudinal biodiversity trends. The evaluation triangulated these findings with protected area management effectiveness data, primary qualitative data collection, and secondary project document review. A variety of complex data sets were aggregated to conduct analysis using statistical (e.g., regression) and predictive analysis (e.g., machine learning-based causal tree analysis).

Findings

GEF-funded protected areas exhibit comparatively less habitat destruction, as demonstrated by forest coverage, than areas that did not receive GEF support.

Benefits

- Access to geospatial, remote sensors, wildlife population, and other data sets expanded the scale and reach of data analysis across geography and time.
- Use of big data enabled the development of counterfactuals (using propensity score matching) in the absence and infeasibility of treatment randomization.
- Use of big data expanded the breadth of variables included in the analysis, as well as the ability to understand causal pathways and interaction effects.

Limitations

- Gaps in both time series data and programmatic data affected the ability to establish counterfactuals and make confident comparisons.
- Inconsistent project documentation data made it difficult to understand and compare intervention models.
- Using big data expanded the variety and volume of data, but on-the-ground contexts were critical to understand and verify conclusions. Triangulation is still seen as critical to understanding human-ecological interactions.

[23] Global Environmental Facility (2015). Impact Evaluation of GEF Support to Protected Areas and Protected Area Systems.

Case Study 2

Sector: Health
Title: Airtel Case Study: Harnessing Mobile Big Data to Identify Tuberculosis Hot Spots in India
Organization: Airtel, Be He@lthy, Be Mobile (a joint initiative between the World Health Organization and the International Telecommunications Union), GSMA
Source: MERL Tech State of the Field key informant interview
Types of data collection technology: Mobile network data from more than 40 million mobile users
Types of data analytics: Machine learning using Airtel algorithmic models conducted on mobile network data (and tuberculosis incidence data)
Document link: GSMA Report[24]

For what purpose was big data used and by whom?

The study aimed to develop a proof of concept for using mobile network data to identify geographical locations at risk of tuberculosis in the Indian states of Uttar Pradesh and Gujarat.

Methods

- To estimate catchment area tuberculosis incidence rates, GSMA mapped the locations of tuberculosis clinics and their catchment areas and then combined clinic-level annual tuberculosis incident data with population estimates.
- The study used aggregated and anonymized Airtel mobile data to track habitual movement (regular commuting) data within states. It layered this information with tuberculosis catchment data to identify common commuter flows between low-, medium-, and high-incidence catchments.
- Areas with low tuberculosis incident rates but high-volume movement to high-incident areas were identified as potential hot spots at risk of tuberculosis.

Findings

"Statistical analysis showed that regular population movement is a stronger indicator of tuberculosis incidence than location proximity between high- and low-tuberculosis regions" (GSMA).

Benefits

Access to large, granular, continuous, real-time data enabled the identification of potential hot spots vulnerable to the spread of tuberculosis. GMSA considered predictive analytics to offer the potential to enable data-driven and proactive public health approaches to prevent the spread of disease, target resources, and promote treatment adherence.

Limitations

Partnerships to aggregate and analyze individual-level records involve complex data governance, ethical, privacy, and regulatory considerations. Clear data sharing agreements that define responsible and compliant data processing, anonymization, retention, and usage processes are critical to protect the privacy and security of individuals and communities.

[24] Helping End Tuberculosis in India by 2025 (2018). GSMA.

Case Study 3

Sector: Humanitarian Response
Title: Utilizing Real-Time Mobile Analytics to Inform Emergency Disaster Response in Turkey
Organization: GSMA, Turkcell
Source: GSMA
Types of data collection technology: Mobile telephone
Types of data analytics: Machine learning
Document link: GSMA Report[25]

For what purpose was big data used and by whom?

Turkcell, Turkey's largest mobile network operator, developed the real-time Galata data platform to alert citizens of impending natural disasters and support the Turkish government's emergency response.

Methods

Turkcell processes more than a terabyte per day of mobile network user metadata and mobile signal data through Galata, a pseudonymized location platform that can identify the number of people affected by natural disasters across Turkey. Galata analytics staff analyze data to provide on-demand, aggregated, anonymized data to enable the Turkish government to plan evacuations and rescues and identify high-risk locations.

Findings

Ongoing improvements in processing capacity (e.g., through investments in hardware extension) are needed to manage such high volumes of mobile data.

Benefits

Use of real-time data supports efficient mobilization of rescue operations and more precise identification of disaster-affected locations and populations. It helps ensure provision of life-saving assistance and effective direction of resources. Galata supported Government of Turkey response efforts during a 2017 earthquake in Bodrum and flash flooding in Rize in 2018.

Limitations

Processing large quantities of data throughout the application development phase was challenging and will require further resourcing.

[25] Using Real Time Data Analytics to Inform Emergency Response in Turkey (2019). GSMA.

Case Study 4

Sector: Political Participation
Title: Using Social Media to Evaluate Programs to Promote Women's Political Participation
Organization: UN Women
Source: MERL Tech Conference 2018 presentation, and publication (UN Women, 2018)
Types of data collection technology: Twitter (Mexico) and Facebook and radio data (Pakistan)
Types of data analytics: Publicly available application program interface and proprietary Crimson Hexagon software
Document link: UN Women Report[26]

For what purpose was big data used and by whom?

UN Women sought to understand the feasibility of leveraging social media data to improve evaluation of gender equality and women's empowerment and the agency's contribution to women's leadership and political participation.

Methods

The research identified Twitter (Mexico) and Facebook (Pakistan) hashtags relevant to UN Women's campaigns to promote women's political participation. Radio call-in data was also analyzed in Pakistan. The analysis used application program interface and Crimson Hexagon software.

Findings

- Twitter worked well for evaluating UN Women's interventions, but only in countries such as Mexico, where high proportions of the population use social media.
- In poorer countries, such as Pakistan, where only the middle and upper classes use social media, radio call-in programs are more effective.
- Social network analysis can help reveal online networks of users and their degree of influence within those networks.
- Interpretation of findings usually requires face-to-face focus groups or individual interviews.

Benefits

- Social media can support more in-depth analysis of social processes, including those that evolve over time.
- It is possible to incorporate techniques such as social network analysis into this kind of research to gain insights into influence and information dissemination patterns within and between groups.
- It is possible to collect information from much larger samples.

Limitations

- There is a potential for selection bias.
- Selecting appropriate hashtags is time-consuming.
- It is difficult to interpret the cultural context of each site.

Case Study 5

Sector: Education
Title: USAID Let's Read Project Performance Tracking System Dashboard
Organization: USAID, Education Development Center
Source: MERL Tech Conference, Washington, D.C., 2019
Types of data collection technology: Integrated data platform combining educational sets on tests and grades disaggregated by province, district, zone, and school
Types of data analytics: Dashboard
Document link: USAID Let's Read Project: Performance Tracking System Dashboard

---

For what purpose was big data used and by whom?

The study was designed for the Zambian Ministry of Education to monitor and improve the efficiency of school exam and grading systems by organizing an educational data set too large to handle efficiently through manual analysis into an easily accessible dashboard.

Methods

The study synthesized data on 848,074 learners in 4,777 schools throughout Zambia. It used dashboard and data visualization technology.

Findings

- Not all schools, teachers, and administrative units follow the school examination administration policy; data is far from complete.
- Data is not comparable because schools use different tests.
- The data does not consider sub-skills such as phonics or writing; only general literacy data could be analyzed and compared.
- Data cannot be used for planning until tests are standardized.

Benefits

The system is easy to use and understand. It is effective in identifying missing data and failure to use required tests.

Limitations

The presentation did not identify any limitations.

Case Study 6

Sector: Youth Employment
Title: Using Machine Learning to Strengthen a Youth Employment Program in South Africa
Organization: Harambee
Source: Reflecting the Past: Shaping the Future. Making Artificial Intelligence Work for International Development. Referenced in a presentation by the USAID Global Health Bureau at the 2019 MERL Tech Conference in Washington, D.C.
Types of data collection technology: Integration of multiple sources of data that Harambee had already collected from its youth employment programs
Types of data analytics: Machine learning and artificial intelligence
Document link: USAID Report

For what purpose was big data used and by whom?

Harambee, a youth employment program in South Africa, used machine learning to address structural employment barriers that unemployed youth face in entering the formal labor market.

Methods

Harambee created an integrated data platform to better leverage the information that it routinely collects from potential employers and job-seeking youth. The platform expanded lists of conventional work skills to include over 500 identified competencies classified into seven job "families." These were combined with an analysis of location and travel patterns, as poor transportation is a major barrier for youth living in remote, poorly served townships. The analysis included tracking taxi travel, predicting likely routes and travel times, and providing guidance on the most efficient routes.

Findings

The program is at an early stage. Some initial findings include:

- Inefficient transport services are a major constraint to accessing many opportunities.
- Traditional employability criteria requires significant revision to address the realities of low-income youth.

Benefits

Travel time and cost is an important job screening criteria. Harambee is starting to work with taxi companies to rationalize their routes. The program helps Harambee and potential employers use existing data more effectively.

Limitations

Some kinds of data are difficult and expensive to collect. The program is at an early stage, so there is still a need to ensure buy-in from potential employers.

Case Study 7

Sector: Agriculture and Water
Title: Using Drones in the Evaluation of an Irrigation Project
Organization: World Bank Independent Evaluation Group
Source: Presentation at the 2019 MERL Tech Conference in Washington, D.C.
Types of data collection technology: Satellites and drones
Types of data analytics: Visual and computer analysis of images

---

For what purpose was big data used and by whom?

The World Bank's Independent Evaluation Group evaluated images of an irrigation project.

Methods

The study used satellite and drone imagery to compare changes in the state of infrastructure, encroachment on farm land, and land use at different stages in the irrigation project.

Findings

The presentation did not mention the study's findings.

Benefits

Use of two technologies provides time-series comparisons over the life of the project, and often for a period before the start of the project and after its completion. The combination of different scales makes it possible to examine large geographical areas and measure detailed changes in small areas that can be checked through on-the-ground visits.

Limitations

There is a danger of misinterpreting geospatial data. Ideally, the images should be combined with on-the-ground measures, which increases costs.

Case Study 8

Sector: Democracy and Governance

Title: Big Data in Evaluation: Experiences from using Twitter Analysis to Evaluate Norway's Contribution to the Peace Process in Colombia

Organization: Global Environment Facility (GEF)

Source: Javier Fabra-Mata and Jesper Mygind (January 2019). "Big Data in Evaluation: Experiences from Using Twitter Analysis to Evaluate Norway's Contribution to the Peace Process in Colombia." Evaluation 25, No. 1: 6–22. doi:10.1177/1356389018804259.

Types of data collection technology: Twitter

Types of data analytics: Social media analytics (content analysis, trend and time series analysis, and user analysis)

Document link: https://journals.sagepub.com/doi/abs/10.1177/1356389018804259

For what purpose was big data used and by whom?

Norad's Evaluation Department evaluated Norway's contribution to Colombia's peace process (June 2010 to December 2016). Evaluators analyzed Twitter data to learn about critical moments in the peace process and Norway's actions; capture data related to the peace process from the wider public; identify key stakeholders in the peace process; and generate new insights to inform the evaluation.

Methods

Multiple data sources included key informant interviews, archival document review, and social media analysis. Social media analysis used Twitter data. Data was extracted via Gnip, using keywords related to public discourse and actor groups in a limited time period. Data was cleaned and enriched with language detection and geo-located. Data analysis included sentiment analysis (natural language processing), trend and time series analysis, and user analysis using social network analysis methods.

Findings

The findings from the Twitter analysis were of limited use to the evaluation itself, as the scope of the evaluation was narrower than the analysis of the Twitter data.

Benefits

- Twitter data is useful for data triangulation, enabling evaluators to verify data and findings from other evaluation sources.
- Twitter data helped the team verify the evaluand's Twitter activity.
- Sentiment analysis was used as a proxy for evaluating attachment among various peace players in the evaluation's model approach to trust.
- Interim findings shared with the Norwegian Ministry of Foreign Affairs might have influenced outcomes among individuals.

Limitations

• Because sentiment analysis algorithms do not incorporate context well, it is possible to misclassify statements. • When Twitter data is extracted using keywords or other criteria, coverage could be compromised or have noise. • Because only a sub-set of the population uses Twitter, the data does not necessarily reflect the opinions of an entire population. • Analyzing data for time series trends was a challenge, as Twitter activity skewed toward the end of the peace process. • Twitter location data is a challenge, as each user enters location field data. This was mitigated by using the Google Maps application programming interface to retrieve country locations. • Analysis was limited by actors' Twitter activity and limited use of hashtags.

Case Study 9

Sector: Human Rights (Counter-trafficking)
Title: Measuring Resilience and Vulnerability to Human Trafficking: A Big Data Analytics Approach Using Weak-Signal Analysis
Organization: Novometrics LLC and Social Impact
Source: Presentation at the 2019 USAID Counter-Trafficking in Persons Evidence Summit
Types of data collection technology: Integration of multiple existing open data sources
Types of data analytics: Weak signal analysis

For what purpose was big data used and by whom?

The study sought to identify key combinations of human social-cultural and behavioral characteristics that make some Filipino workers working overseas vulnerable to human trafficking and others resilient. Integrating these attributes enabled researchers to construct a vulnerability index that predicts vulnerability to human trafficking.

Methods

Researchers collated many existing open data sources, including detailed national census data, health and educational survey data, and remote sensing data. Correlation analysis and weak signal analysis were used with geospatial integration to create a vulnerability index that could estimate the risk of being trafficked.

Findings

According to the Vulnerability Index, overseas foreign workers from areas from which many people leave the country for employment were less vulnerable than those from areas that send few workers. The weak signal analysis suggests that attributes of vulnerability are more associated with individual empowerment and decision-making autonomy than with economic opportunity or established enclaves resulting from chain migration.

Benefits

This method was helpful for measuring and estimating the prevalence of an extremely rare concept that is difficult to measure. Key to implementation were the need for a calibration data set (in this case, victim data) and sufficient socio-economic data from other sources, disaggregated by sex.

Limitations

This method is initially most useful for program design but, if repeated over time, could support evaluation.