Jawad Jalali for AFP

TECHNOLOGIES FOR DATA COLLECTION, PROCESSING AND COMMUNICATION IN EDUCATION IN EMERGENCIES

Mapping Practices and Opportunities in the MENA region and Globally

TABLE OF CONTENTS

ACRONYMS

AENN	Addressing Education in Northeast Nigeria
CAQDAS	Computer-assisted qualitative data analysis software
CAPI	Computer-assisted personal interviews
CCTV	Closed-circuit television
DDDM	Data-driven Decision Making
DNA	Deoxyribonucleic acid
DSEG	Data Science and Ethics Group
EGRA	Early Grade Reading Assessment
EiE	Education in Emergencies
EMIS	Education Management Information System
GIS	Geographic information system
GDPR	General Data Protection Regulation
INEE	Inter-Agency Network for Education in Emergencies
INGO	International non-governmental organization
IVR	Interactive Voice Response
LMS	Learning Management System
MEERS	Middle East Education Research, Training and Support program
MENA	Middle East and North Africa
MEQA	Measuring Evidence of Quality Achieved
ML	Machine Learning
NWOW	New Way of Working
OCHA	Office for the Coordination of Humanitarian Affairs
ODK	Open Data Kit
OMR	Optical mark recognition
PAPI	Pen-and-paper assisted personal interviews
PII	Personally identifiable information
RA	Rapid Application (UNHCR)
RQDA	R Qualitative Data Analysis
SEL	Social and emotional learning
SMS	Short message service
SSSAMS	South Sudan Schools Attendance Monitoring System
UAV	Unmanned aerial vehicle
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNHCR	United Nations High Commissioner for Refugees
UIS	UNESCO Institute for Statistics
UNRWA	United Nations Relief and Works Agency for Palestine Refugees in the Near East
USAID	United States Agency for International Development
USGS	United States Geological Survey

1. EXECUTIVE SUMMARY

The effective integration of technology in the Education in Emergencies (EiE) sector has great potential to improve the availability, quality and use of data for stakeholders globally, but the sector faces unique challenges in introducing these new technologies. Particularly, EiE interventions serve populations burdened by poverty, displacement, and violence and, therefore, have unique data privacy and protection protocol requirements. How, then, can technology best serve the needs of EiE practitioners? And what specific principles should practitioners follow when adopting a new technology? The study team answers these questions via an analysis of interviews with 35 EiE, education, and technology professionals, classifying data technologies currently in use among the three main phases of the data life cycle: data collection, processing, and communication. Based on the results of these interviews and review of the literature, the study team proposes **eight guiding principles for practitioners** as it relates to technology use across the data life cycle. Two of these principles should act as a framework for all decisions: (1) Do no harm, and (2) Follow General Data Protection Regulation (GDPR)[1] and similar standards. Three of these principles help practitioners to identify when a new technology is appropriate: (3) New technology does not mean better technology, (4) Coordinate among stakeholders, adapting technologies using a systems-thinking approach and (5) Develop innovations in collaboration with local organizations and end-users. The final three principles help practitioners in deploying the technology successfully: (6) Introduce new technology through the lens of social and behavior change, (7) View technology as a long-term investment, and (8) Nurture a culture of data feedback loops and data-driven decision making.

The study team also shares recommended practices related to each phase of the data life cycle. For data collection, the study team recommends highlighting duty of care, the need to deploy multiple modes of data collection, the importance of design, and the need to tailor technological solutions to the culture in which it is being deployed. For data processing, the study team recommends investing in strengthening analytical support, building capacity for staff to better leverage data, building expertise in multiple platforms, automating data validation processes, and developing processes to share analyses. Finally, for data communication, the study team recommends identifying data champions within the organization, creating a broad culture of data sharing within the organization, providing data in a timely fashion, following data visualization best practices, simplifying reporting, ensuring system maintenance, and using row-level security to limit access to data.

Aligned with these principles and best practices, the greatest opportunities for technology in EiE that can further strengthen existing technological systems and equip decision makers with key information on learners, teachers, schools, and systems are identified. These opportunities include expanding use of existing platforms to collect data, empowering beneficiaries, field staff, schools, and teachers with data through feedback loops, and exploring new promising technologies, particularly data collection through learning management systems, automation of data cleaning and analysis, sector-appropriate machine learning techniques, and chat-bots and enhanced information access.

[1] GDPR is the key regulation on data protection and privacy within the European Union. Under the GDPR guidelines, data controllers must design information systems with privacy in mind and no personal data can be processed unless a data subject has provided informed consent or if is used under six lawful purposes, including fulfilling contractual obligations with a data subject and protecting the vital interests of a data subject or individual. More information on GDPR can be found at https://gdpr.eu/.

2. BACKGROUND

The world is more interconnected through technology than ever before, enabling data collection at a scale previously unimaginable. [1] [2] While the potential of technology to support the humanitarian sector has been widely recognized [3] [4] [5], technological innovations in education in emergencies (EiE) remain modest, including in the data space. [6] Many humanitarian applications of technology to data collection and use have occurred within sectors other than EiE. [7] [8]

One potentially analogous sector is development, where technology has made it possible to collect and analyze valuable data that can inform and improve education interventions in some of the world's most resource-strained environments. It is therefore tempting to approach the use of technology for education in emergencies (EiE) as an extension of technology for development. After all, both international development and EiE fields require technologies that are cheap, appropriate for settings with poor connectivity and low levels of technical expertise, and support sustainability. However, EiE settings also involve unique challenges related to technology's use for data collection and analysis.

The rapidly changing nature of emergency settings particularly encourages the development of tools and platforms that can provide real-time information. Although precise longitudinal information systems are ideal in both EiE and international development interventions, this is much harder to achieve in EiE settings, where actors require flexible tools to collect data from highly mobile populations and hard-to-access locations, while also dealing with the high rates of staff turnover that characterize humanitarian programs. [9] [10] Moreover, EiE actors have often prioritized the use of technology for the monitoring of key outputs during service provision, given the difficulties associated with assessing quality of services and learning environments. [11]

This need for rapid and flexible responses has made international non-governmental organizations (INGOs) a dominant presence in the EiE space, which has clear consequences for the use of technologies for data collection and use. Despite existing coordination across INGOs and other implementing partners through the cluster system, project-based approaches lead to situations in which technology for EiE has consisted of the creation of ad hoc monitoring and information management tools, rather than more sustainable system-wide approaches for data collection and sharing for a particular emergency context. [12] Moreover, adoption of tools can respond to an organization's general attitude towards technology, often with changing preferences across country and regional offices.

The EiE sector is also unique in its inherent security risks. Scholars have long argued that advances in mobile technology make it suitable for hard-to-reach contexts [13], particularly in parts of the world that are already well-connected, such as in the Middle East and North Africa (MENA) region. Despite the increased availability of mobile phones around the world, the use of often expensive devices for data collection—such as smartphones and tablets—remains a sensitive issue in low-resource areas, as it highlights the existing inequality and power dynamics between target populations and actors collecting information. In emergency settings, the use of devices can also involve security risks to data collectors and may provoke additional suspicion and mistrust from governments and communities. Collected data also requires stronger data privacy and protection protocols. Populations served by EiE interventions often face not only the burden of poverty, but that of displacement and violence. Educational data on IDPs and refugees could become available to individuals or organizations responsible for their displacement or expose the identities of vulnerable individuals to other local actors. Moreover, even in cases in which vulnerable populations reach areas where their physical safety is guaranteed, their data is vulnerable to misuse by government authorities seeking to restrict the population's activities in their territory. [14]

It is within this context that the study team seeks to address the growing interest in understanding how the EiE field can more effectively leverage recent technological innovations to create more responsive, targeted, and high-quality monitoring and evaluation to improve educational programs, particularly for NGOs working in this space.

Barriers to the widespread implementation of technology-based data collection solutions for EiE, both in the MENA region and globally, include uneven implementation of platforms across localities and partners, technology penetration, literacy (both traditional and digital), upfront costs, poor ability to handle qualitative data by technology users, and concerns around digital harm. [6] [4] [15] [16] [17] [18] Thus, there is a need to explore issues, opportunities, and challenges related to the development and utilization [2] of digital information systems for education in emergencies.

Existing research on the use of technology in the EiE sector has focused on EdTech's potential to deliver content, whether that is learning in emergencies [16], distance learning approaches [19], the use of radio [20], or the development of intelligent tutoring systems [21]. Reviews of technologies, including mobile phone data collection [22], tablets [23], or SMS technologies [15] as well as reports on key processes where technology adds value, such as digitizing data collection for digital school censuses [24], building EMIS systems [17], and rapid data collection during an emergency [25] provide insights in successful approaches to the use of *specific* technologies in development and EiE. However, the literature lacks clear guidance on the use of data technologies more generally within the constraints of the EiE context in the MENA region.

This paper, which was developed under the USAID Middle East Education Research, Training and Support (MEERS) program, attempts to fill the literature gap by identifying data technologies relevant to MENA and global EiE contexts and characterizing the advantages and disadvantages of using these technologies across the data life cycle. Based on experiences of the experts in these contexts, the study team proposes guiding principles to the continued use of technology in the EiE sector and develops recommended practices in data collection, processing, and communication.

2.1 RESEARCH QUESTIONS

The MEERS research team set out to answer the following questions, with a focus on the needs and practices of EiE practitioners.

1. What technologies are currently used by EiE practitioners in the MENA region and globally for data collection, processing, and use?
2. What are the requirements to use these technologies (particularly as it relates to infrastructure and professional capacity)?
3. What are the advantages and disadvantages to these data collection, processing, and communication technologies in relation to general utility, costs, and safety and security?

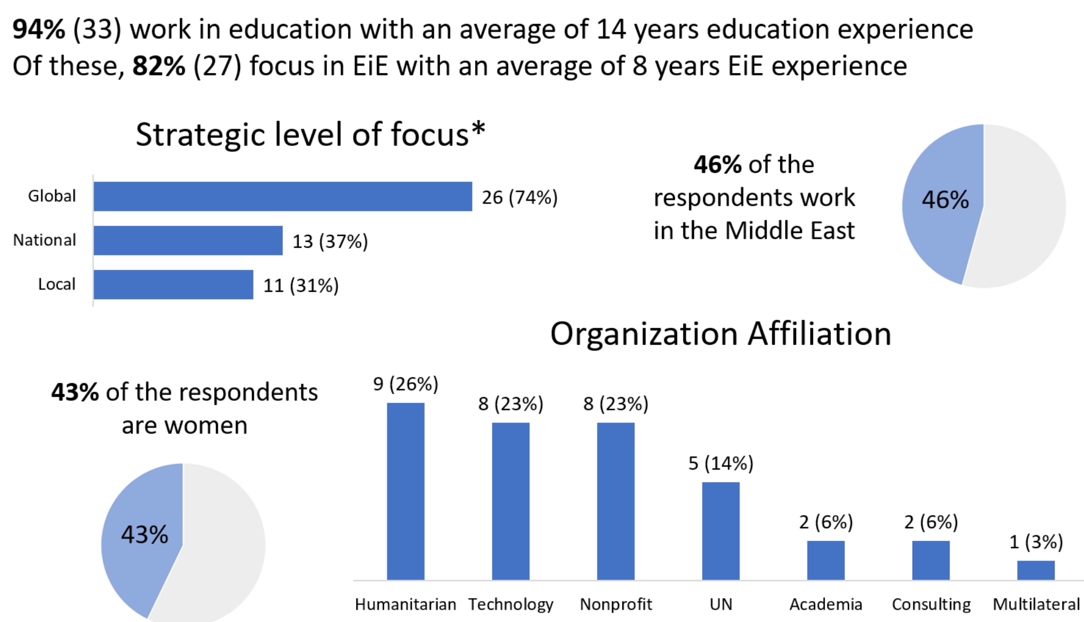
Since 2019, global actors led by the Inter-Agency Network for Education in Emergencies (INEE), UNESCO Institute for Statistics, and Education Cannot Wait have been moving towards the standardization of a shared global data infrastructure and ecosystem. The INEE EiE Data Summit in June 2019 led to the development of a global action agenda around EiE data and the formation of an INEE EiE Data Reference Group in 2020. In keeping with this global agenda, the study team argues that the effective integration of technology into the broader EiE data ecosystem has great potential to systematically improve the availability, quality and use of EiE data for a broad range of stakeholders globally. At the same time, there are always risks to the introduction of new technologies into any sector; this paper sets guiding principles for how to manage this process and recommended practices for each process in the data life cycle.

3. METHODS

This study adopted an iterative, qualitative approach combining a desk review with semi-structured interviews to map the application of technologies for data collection, analysis, and use in EiE contexts. The goals of this mapping are to develop a typology or organizing framework for the different kinds of technologies that are being used—or could possibly be used—for data collection, analysis, and use in EiE contexts and to highlight lessons learned and good practices to assist those working in the sector to improve the use of technology. Due to the combined regional and global focus, the study team reflects on the relevance of the findings for both the MENA region and globally throughout the paper.

The study team reviewed both academic and grey literature. The primary purpose of the desk review was to confirm that the research questions addressed a knowledge gap regarding EiE, data, and technology and to identify initial stakeholders for interviews, though relevant insights from the desk review are integrated into the Findings section. During the review, the study team screened over 90 articles and identified over 25 of relevance, with results from the search buttressed with relevant documents shared during the interviews.

Figure 1: Who did the study team interview?



*Note that respondents may have listed multiple strategic level of focus (i.e. a focus on national and local work)

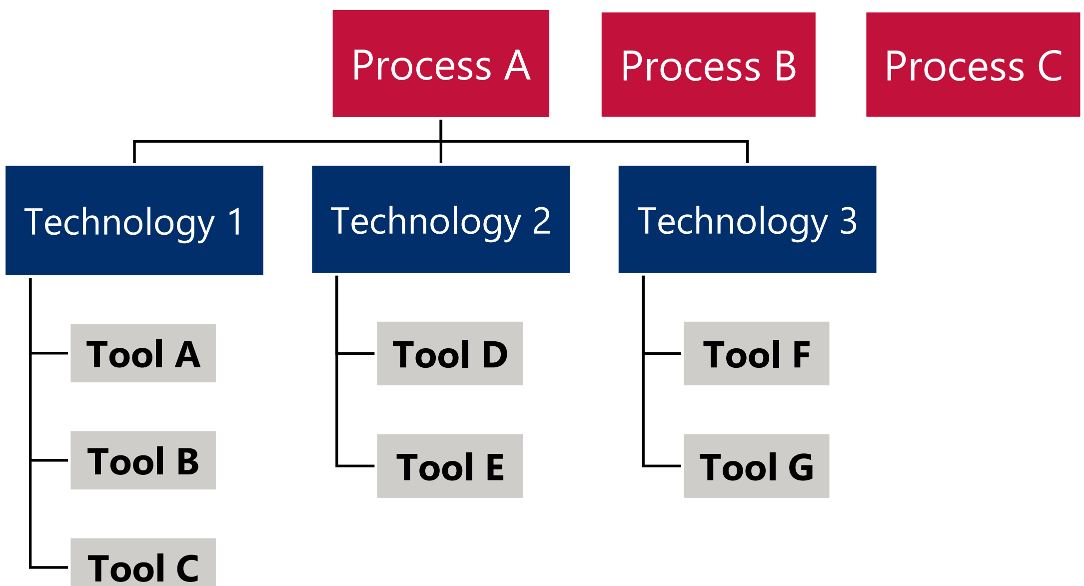
The study team conducted 35 semi-structured interviews lasting 30 to 60 minutes with professionals and researchers who have expertise related to the use of technology for data collection in the EiE sector. The interview was divided into questions on (1) demographics, (2) the types of technology used by the organization, (3) requirements to use the technology (infrastructure and professional capacity), and (4) the perceived advantages and disadvantages (general utility and costs) and recommended practices to use these technologies, and (4) potential uses of technology by the sector in the future. Interview questions can be found in Appendix *1. Interview Protocol.* The primary sampling strategy was purposeful snowball sampling, beginning with known specialists in the field to learn about challenges and opportunities as it relates to using technology in emergency settings. This purposive sampling included respondents across different levels of conflict severity, different staff roles (e.g., technical and managerial), different levels of operation (e.g., local and global), and different mandates (e.g., government, private, development and humanitarian). When additional relevant perspectives were identified during the interviews, such as the use of emerging technologies that had little coverage in the current sample, the study team adjusted the sampling, accordingly, snowballing as needed. While interviews focused on individuals directly involved in EiE, due to the cross-cutting nature of technology, technological generalists and specialists in related sectors were interviewed.

Respondent summary statistics can be found in Figure 1. The majority of interviewees had been involved with the field of education in some capacity, averaging 14 years of experience (n=33). A considerable number had also been active in the EiE space, with an average eight years of experience (n=27). Sixty-four percent of respondents identified their work as being at the global level, while 21 percent characterized it as being at both the local and national level. Slightly over a third of all respondents (37%) had ties to organizations working in the MENA region. Participants were balanced in terms of gender, with 43 percent female respondents. Twenty-three percent of respondents were affiliated with firms providing technology services for data collection and analysis; 26 percent were affiliated with humanitarian NGOs; 23 percent ranged from nonprofit organizations involved in the design and implementation of education interventions; 14% were employees of a United Nations (UN) agency, and the remaining (14%) came from academia, consulting, and non-UN multilateral organizations.

3.1 ANALYSIS ANALYTICAL FRAMEWORK

3.1.1 ANALYTICAL FRAMEWORK

Figure 2: Analytical Framework



The study team used an iterative framework to categorize and understand the technologies used in EiE settings. The team first identified existing technologies used in the EiE context by searching both academic and grey literature for papers related to information and communications technology (ICT) and data. After removing technologies that were solely used for content delivery, the study team classified these technologies as living within larger processes within the data life cycle. Through continued review of the literature and interviews, the study team identified other technologies as well as specific tools (proprietary and non-proprietary software solutions to implement these technologies) living within these larger processes. This framework of processes, technologies, and tools is shown in Figure 2.

3.1.2 ANALYSIS METHODS

The study team iteratively analyzes the data, moving from tool-specific responses, such as information related to the advantages of using PowerBI, to generalized insights. As described above, data is applied to a two-tiered typology: (1) the underlying **technology** the insight is attached to and (2) where that technology belongs within the main **processes** of the data life cycle. To do this, data is coded to identify tool-level insights in relation to requirements, advantages, and disadvantages as well as the specific technology type (i.e., categorizing PowerBI as a dynamic visualization technology) using a computer-assisted qualitative data analysis system (CAQDAS). Through iterative analysis of the insights related to specific technologies, broader themes in relation to requirements, advantages, and disadvantages and recommended practices for the larger process are identified. Finally, the study team continues the process of abstracting insights from each process within the data life cycle to guiding principles *across* data processes more generally within the EiE sector.

3.2 RESEARCH LIMITATIONS

Due to the sheer number of technologies used in the sector and space considerations within the paper, the study team generalized the results to data collection, processing, and use technologies broadly with the goal of providing insights into how technology can support data use in the EiE sector. For comparisons across technologies (for example looking at advantages and disadvantages of SMS technologies relative to IVR technologies), please see *Comparing Data Technologies* in the paper's companion dashboard.[2]

This study does not focus on the use of technology for national education systems in crisis settings, which has been researched via parallel initiatives, although some of the findings of this paper may be useful for government actors. The focus of this paper is primarily on non-governmental EiE practitioners, both large and small, global, and local. The hope is that the findings will improve the efficiency, quality, and usefulness of the overall EiE data ecosystem and ultimately make more data, both governmental and non-governmental, increasingly inter-operable and coherent. An end goal is to effectively bridge humanitarian and development data divides and move towards *A New Way of Working* in the EiE sector defined by common ways of working and collective outcomes.

The study is also limited by the chosen research methodology; through snowball sampling, the research team identified and interviewed experts in the field, but the identified network may exclude major players in the EiE sector. The study team particularly was unsuccessful in identifying state-level actors to discuss the use of technology for national education systems and may have inadvertently excluded non-English speaking participants. The reliance on interviews as the main source of data (rather than collecting information more systematically through a survey, for example), also results in a reduction in the breadth of information captured in favor of depth. Finally, while respondents were assured that their responses would not be tied to their name or their organization, some respondents may have self-censored due to a perceived lack of anonymity in responding to a recorded interview. De-briefs at the end of each interview helped assure potential interviewer biases (such as errors made by altering the questionnaire, irrelevant probing, or recording errors) were minimized.

[2] https://tinyurl.com/usyyjxyc

FINDINGS

This section provides the findings from the study organized into four parts.

1. Section 4.1 provides an overview of the key technologies identified in the literature and by interview participants and how these technologies map to the data life cycle (collection, processing, and use). (Research Question 1)
2. Section 4.2 addresses the requirements, advantages, and disadvantages of these different technologies as they relate to data collection in the EiE sector (Research Question 2 & 3).
3. Section 4.3 addresses the requirements, advantages, and disadvantages of these different technologies as they relate to data processing in the EiE sector (Research Question 2 & 3).
4. Finally, Section 4.4 addresses the advantages and disadvantages of these different technologies as they relate to data communication in the EiE sector (Research Question 2 & 3).

4.1 A TYPOLOGY OF DATA TECHNOLOGIES USED BY MENA AND GLOBAL EIE PRACTITIONERS

4.1.1 DEFINING THE DATA LIFE CYCLE AND THE IMPACT OF TECHNOLOGY

Below we define the different phases of the data life cycle. To elucidate the three main categories of the data life cycle (collection, processing, and use), we provide an example of the infusion of technology into collecting social and emotional learning (SEL) and well-being data for students in an EiE context. For a more in-depth look of the use of data technologies in the field, please refer to Appendix 2.

1. Data collection is the process of capturing information on targeted variables. During a phase of data collection, enumerators may collect results of a one-on-one assessment with students in a settlement and then enter this data into a tablet via KoboCollect. This data collection captures not only the main variable of interest, but also secondary information such as the location of the data collection and the amount of time spent on the assessment. This data is then synced from the tablet to a server located on the cloud that only specific individuals within the organization have access to.
2. Data processing is a broad category that includes validating, cleaning, exploratory data analysis, and modeling. For the SEL and well-being data, an analyst may take data collection efforts from several sites and append the data into a singular database using proprietary software such as Stata. They may then merge it with a previous data collection and analyze how much students learned between baseline and endline.
3. Data communication is the process of communicating data back to stakeholders, primarily through written results and data visualization. For our SEL data, an analyst may develop a Sankey graph using PowerBI to show how students' social and emotional skills have progressed across the year and show differential progress as it relates to different geographic locations. This is presented to key stakeholders and impacts how resources are allocated in the next quarter.

Table 1 provides a list of technologies identified for each process with definitions, examples of use, and software or companies that provide this technological solution. While the study team believes this categorization provides clarity on the technologies used within the data collection cycle, these categories are *not* mutually exclusive or collectively exhaustive. SMS technologies, for example, are categorized as a *data collection* technique because SMS was mostly described in that way during interviews (such as the South Sudan Schools Attendance Monitoring System). However, SMS could also be categorized as *data communication* if results are sent back to users via text messages. In the mapping below, the study team restricted categorization to the most commonly noted use of the technology.

Table 1: Technologies Used Across the EiE Data Life Cycle

Process	Technology	Definition	Examples of Use	Illustrative Software
Collecting Data	Offline mobile surveys	Assessments or surveys implemented through applications that do not require online access.	Surveys led by an enumerator to capture beneficiary data such as a school environment survey.	ODK, ActivityInfo, KoboFormScanner, Moodle, Survey123, Magpi, DeviceMagic, Commcare, Ona, iForm Builder
	Online assessments	Assessments or surveys implemented through forms that are accessible via the web.	Surveys deployed to respondents online such as sending parents a link to a survey on child welfare via social media.	Google forms, Survey Monkey, Microsoft Forms
	SMS	The text messaging service component for mobile devices.	SMS software is used directly to collect data from beneficiaries through SMS based forms or qualitative responses.	Commcare, Echo Mobile, FrontlineSMS, Magpi, Telerivet, Tera, Textit
	Interactive Voice Response	Interactive Voice Response (IVR) allows humans to interact with a computer-aided phone system through voice and input.	A deployed IVR system calls target respondents with the option to respond via numerical responses on their keypad.	Voto Mobile, Twilio
	Social Media	Interactive digital platforms that facilitate the sharing of information.	Data on school closures due to violence shared on social media.	WhatsApp, Facebook, Instagram
	Apps (Applications)	A program designed for end-users. These can be proprietary or open source. Apps for mobile platforms are called mobile apps.	User downloads and uses a learning application with integrated learning assessments on a mobile device; that data is sent back to the developer.	Learning Management Systems, WorldReader
	Audio & Video Recording	Digital inscription and creation of sound and/or video.	Images or videos of classrooms and schools. Described as a potential method to collect data on teacher attendance (such as through CCTV).	Digital sound recorders, video cameras
	Remote Sensing	Images of earth collected by imaging satellites, planes, or UAVs.	Identifying school buildings (i.e., cross referencing data collection locations with satellite images).	GoogleEarth, Sentinel Hub, USGS

	Biometric data collection	Use of biometrics (fingerprints, iris, DNA) for identifying individuals	Digital identities used to track individuals, the services they qualify for, and education qualifications.	UNHCR's Rapid Application (RApp)
	Optical mark recognition	Optical mark recognition (OMR) is the process of capturing human-marked data from document forms.	Assessments being scanned to get item-level information on student performance.	Scantron, FormScanner, Moodle
Processing Data	Spreadsheets	Computer application for organizing and analyzing data in tabular form.	Data stored on the number of attacks in a region across time.	Microsoft Excel, Google Sheets, Numbers
	Data Analysis Software	Software designed to support quantitative data analysis.	Creating maps that provide beneficiaries with information on the nearest service provision locations.	SPSS, ArcGIS, QGIS, Stata
	Computer-assisted qualitative data analysis software	Computer-assisted qualitative data analysis software (CAQDAS) software assists with transcription analysis, coding, text interpretation.	Software systems that help collate qualitative data and assist in deriving insights from these sources.	Dedoose, RQDA, MAXQDA, Nvivo
	General purpose programming languages	Programming language dedicated to general purposes.	Developing a continued process of checking on a collected dataset's data quality issues through a scripted analysis.	R, Python, JavaScript, Julia
	Machine learning techniques	Artificial intelligence technique using computer algorithms that improve through experience.	Adaptive learning management systems that provide learning materials to students based on their performance on previous modules.	TensorFlow, Apache MXNet, PyTorch
Using Data	Static Visualization	Visualization that captures a specific story of the data.	Static images showing student performance on an assessment.	Adobe Illustrator, ArcGIS, Quantum GIS, Microsoft Excel
	Dynamic data visualization	Visualization that can be manipulated in real-time by a user.	A dashboard that shows the number of students in conflict-impacted regions in a country with drill down features to see schools at a local level. Qualitative data on the result of a parent survey visualized with a word cloud.	Microsoft PowerBI, Tableau, Google Data Studio, shinydashboard, Salesforce

4.2 DATA COLLECTION TECHNOLOGIES

All interview respondents discussed the use of technological tools in the first phase of the data life cycle: data collection. Data collection technologies used in EiE are as varied as the type of data collected. Technology is used to collect data on beneficiaries (demographic and historical data) across a variety of topics: program implementation, classroom observations, standard indicators for EMISs, school environment surveys, parent engagement, school closures, attacks in the region, and much more.

The most common data collection approach discussed was in-person mobile data collection led by trained enumerators, though other means of data collection were mentioned, some of which are more nascent. An example of a newer approach is data collection directly through software applications. For example, if an organization uses a Learning Management System[3] (LMS) app within a refugee camp, the app captures data on the use of that system (e.g., what module was accessed, how long was spent on the module, how the student scored on the quizzes within the module, how much time is spent on the app at a time). While collection of such data was mentioned by only a few respondents, it will likely become more common as LMSs proliferate, particularly in response to the COVID-19 pandemic.

In this section, the study team presents requirements for individual data collection technologies. The study team then discusses the advantages and disadvantages of these technologies, highlighting commonalities and differences across different technologies, as it relates to general utility, costs, and safety and security. Later on, in Section 5, the study team presents recommended practices based on these insights, including mitigation strategies for identified risks. For detailed information on specific technologies as well as potential software solutions and use case scenarios, please see *Zooming into Data Technologies* and *Examining Examples & Use Cases of Data Technologies* in this paper's companion dashboard.[4]

4.2.1 REQUIREMENTS FOR DATA COLLECTION TECHNOLOGIES

Interviews and desk review both clearly demonstrated that the requirements for data collection vary greatly by technology. First, all technologies require the use of electricity, a requirement that should be accounted for in any data collection plan and, in the absence of reliable electricity, may necessitate power banks, solar chargers, and back-up batteries. Specific hardware requirements, costs of deployment, and human resource needs vary greatly as seen in Table 2. For an extensive discussion on offline mobile survey technologies particularly, refer to CartONG's *Benchmarking Mobile Data Collection Solutions*. Since the cost is widely variant dependent on the specific tool selected (and its associate hardware and human resource needs) as well as the scale of data collection, the study team is unable to include cost estimates in the requirement table, although the Appendix includes a table of costs for specific tools to act as a starting point for organizations in the Appendix and the accompanying dashboard provides crowd-source relative cost estimates.

[3] A Learning Management System, such as Google Classroom, Moodle, or Canvas, is a software application to administer, document and track the delivery of educational programs. Data generated through tracking lessons or courses can be used to help tailor and curate resources to learner needs.

[4] https://tinyurl.com/usyyjxyc

Table 2: Data Collection Technology Requirements

Technology	Illustrative Software	Internet/network required	Hardware	Human resources
Offline mobile surveys	ODK, ActivityInfo, KoboFormScanner, Moodle, Survey123, Magpi, DeviceMagic, Commcare, Ona, iForm Builder	Internet required for sync.	Smartphones or tablets	Survey developers
Online assessments	Google forms, Survey Monkey, Microsoft Forms	Internet required during collection.	Smartphones, tablets, or personal computer	Survey developers
SMS	Commcare, Echo Mobile, FrontlineSMS, Magpi, Telerivet, Tera, Textit	Network required during collection.	Feature phones or smartphones	Survey developers
Interactive Voice Response	Voto Mobile, Twilio	Network required during collection.	Feature phones or smartphones	Survey developers
Social Media	WhatsApp, Facebook, Instagram	During data collection.	Smartphones, tablets, or personal computer	Survey developers, Data analysts
Apps (Applications)	Learning Management Systems, WorldReader	Internet may be required for application use.	Smartphones, tablets, or personal computer	Data analysts
Audio & Video Recording	Digital sound recorders, video cameras	No internet or network required.	Recording devices	Media capture
Biometrics	UNHCR's Rapid Application (RApp)	Internet required for sync.	Fingerprint scanner, facial scanner, etc.	Data analysts
Optical mark recognition	Scantron, FormScanner, Moodle	No internet or network required.	Smartphones or tablets	Survey developers

4.2.2 DATA COLLECTION TECHNOLOGIES – ADVANTAGES & DISADVANTAGES

Across all data collection technologies, respondents identified themes that highlight both the advantages and challenges of using data collection technologies relative to traditional modes (such as paper-and-pencil collection). These insights are categorized as they relate to general utility, costs, and safety and security.

4.2.3 GENERAL UTILITY OF DATA COLLECTION TECHNOLOGIES

Broadly, data collection technologies improve general utility through enhanced data quality, improved efficiency, strengthening information ecosystems, the general ease of adoption, and allow the collection of richer qualitative data. These advantages are offset by the exclusionary aspects of some data collection technologies, low response rates and access issues, and the ever-present danger of survey creep. Mitigation strategies for the identified disadvantages are discussed in the recommended practices in *Section 5.1.2 Considerations for Donors*.

General Utility: Advantages

Data quality: Respondents noted that one of the inherent advantages of using technology for data collection is the impact on data quality. Most data collection methodologies allow data validation rules to restrict responses, geolocation capture to confirm enumerator presence at specific locations, the implementation of skip logic to ensure only relevant questions are answered, and calculation rules to ensure internal consistency of the data captured. Real-time data collection allows managers to oversee and address quality concerns closer to the point of entry. Finally, removing the need to move paper data records into an electronic system eliminates transcription errors.

Individual technologies also have specific features that enhance quality. Data capture in learning management systems provides more precise estimates on individual students' competencies. Some data capture technologies provide a less filtered view of ground realities as information is not sent through an intermediary. Through data exhaust (data generated as a result of digital or online activities), an LMS or app can provide insights on how an individual is using technology. An example of this is WorldReader's use of app data to identify what content is being read and how long individuals are engaging in the content [26]. Similarly, analysis of social media provides a path to perspectives that might otherwise be self-censored during an interview, though privacy concerns, particularly around consent of the user, must be addressed when designing an approach that leverages data published on social media. Frameworks for addressing these issues are discussed in *5.1.1 Principles for Practitioners*.

Improved efficiency: Respondents described mechanisms that improve the efficiency of data capture including eliminating the need to enter paper data records into an electronic system, speeding up the data cleaning process by having more up-front data validation, creating immediate access to data, localizing and internationalizing surveys, and providing instant data analysis. These technologies also support scale. OMR technology, for example, requires a similar amount of work to scan 1,000 tests as it does to scan 10. Similarly, after developing the survey logic for an IVR collection system, there is no marginal labor from implementing 100 surveys vs 10,000 surveys (though there may be more labor needed for data cleaning!)

Potential to strengthen information ecosystems: The use of technology strengthens information ecosystems, as relationships between datasets can be developed within tools. Some technologies are particularly adept at strengthening information ecosystems. For example, technologies that have case management features, such as Commcare, Taro Works or ActivityInfo, provide access to other information about a recipient, which in turn guide the type of information collected during a subsequent visit[5]. Due to the uniqueness and immutability of data collected, biometric scanning is another example of how technology strengthens information ecosystems: information attached to an individual's biomarker is maintained even if a refugee faces secondary displacement, enhancing interoperability across systems. This requires coordination. As one respondent stated, "Usually [national education systems] have one to one contracts and partnerships with different technology providers who might have different systems [which have their own purposes] but at the end of the day, you end up with … literally over 50 systems running at the same time with different silos of information [that] don't talk to each."

Ease of adoption: The proliferation of technology, particularly feature phones and, increasingly, smart phones, means that enumerators are provided with a tool they are already proficient in. This is, naturally, highly context-specific: the recency of displacement among refugees and IDPs and pre-conflict material conditions and infrastructure, among other factors, can determine the prevalence of specific devices,

[5] For a more complete discussion on features of specific mobile tools, refer to CartONG's Benchmarking of Mobile Data Collection Solutions. [15]

leading an implementing partner to use different technologies in Lebanon than they would in Syria for example. This high user competence contributes to easier adaptation of a tool and, in turn, improved data quality.

Richer qualitative data: Using audio and video recordings can enhance qualitative data collection in contexts where qualitative researchers cannot spend significant amounts of time within a school or community. Similarly, social media can be used to collect a large, unfiltered set of data either directly through posts on social media or by crowdsourcing information through social media. A UN representative shared an example where they used WhatsApp to take pictures of illicit fee requests to help identify and address corruption.

> "In some communities, technology makes people very nervous. When you're working in camp locations people feel very nervous about the fact that you're standing there putting information into a phone. You live in a country like Syria, you absolutely do not trust anybody."
>
> -Humanitarian Organization Representative, MENA

General Utility: Disadvantages

Exclusionary aspects and survey bias: Not all technologies are appropriate for all contexts, and though technology is becoming more ubiquitous, some forms of technology will not reach the targeted EiE population. SMS or IVR approaches are only appropriate when the target respondent has access to a phone and to a network. Social media posts are written by those who have access to devices and are able to post within prevalent social norms. Similarly, app data will be solely collected from individuals who have access to that app.

When an enumerator collects data, the inherent power dynamic (with a device being one physical representation of that power dynamic) may exclude respondents or result in biased information. As one respondent from a humanitarian organization in the MENA region put it, "People aren't so keen to have a person walking around a camp with a tablet in their hand [when they are] living in an environment where it's 54 degrees Celsius in the sun in the middle of a desert, there's no running water, and there's no internet because the authorities … don't want people to be connected with the outside world." Individuals may be wary about providing identifying data, especially biometric data, due to cultural, gender, or power imbalances.

> "Women, youth and certainly displaced people really struggle to have their voices heard on [social media]... We need to think about … what the unintended consequences are if we if we're not hearing those voices and we're still designing our programs without those voices … So that's why I guess [data collection technologies are] one tool in your toolbox. You need to talk to some real humans human-to-human to find out what needs are, especially in those under underrepresented populations."
>
> -Global Non-profit Representative

Low response rates and access issues: Technologies that do not rely on enumerators are likely to suffer from low response rates. For IVR, respondents become used to unsolicited calls and may not respond. SMS surveys generally have low response rates, especially if there is no incentive to participate. Typing in long responses to a phone survey or SMS survey is tedious, which may negatively impact response rates and data quality. Finally, using app data may be challenging as data exhaust is not always accessible to analysts [27].

Survey creep: The marginal cost of adding extra questions to a digital survey appears very low. Decision makers may add more and more questions beyond the scope of the survey resulting in the collection of un-needed data and wasting the time of respondents. As one global humanitarian representative stated, *"Data is like water. If you don't have enough, you'll die of thirst. If you have too much, you'll drown."*

4.2.4 COSTS OF DATA COLLECTION TECHNOLOGIES

Respondents noted that while collecting data via technology may require more upfront costs, it is often cheaper in the long run, and sometimes in the short-term as well. With mobile data collection, enumerators enter data directly into a tablet rather than the expensive, time-consuming task of collecting data via paper and then transferring into a database. Some forms of data collection, such as IVR or SMS, do not require costs associated with an enumerator team as surveys are deployed directly to beneficiaries via their phones. Since many survey development software are open source, there are clear pathways to financial sustainability, though this may require substantial investments in open-source software.

To better guide practitioners, the study team developed a list of cost estimates for specific tools, which can be found in Appendix 3. Pricing Guide for Data Tools. This is further supplemented with crowd-sourced data on *relative* costs of specific tools within the companion dashboard.[6]

4.2.5 SAFETY & SECURITY & DATA COLLECTION TECHNOLOGIES

Technology has both positive and negative implications as it relates to safety and security.

Safety & Security: Advantages

Enhanced data security: Information collected via technology reduces the need to store information on paper. Potentially dangerous processes are bypassed, such as shipping surveys to a field office for data entry. Many mobile data collection apps delete data from devices as soon as information is synced to the server. UNHCR's experience with cash delivery for Syrian refugees in Jordan through iris-enabled ATMs could provide clues to the education sector on how to use biometrics to safely improve targeting of services and tracking students, though using this digital identity has inherent safety risks that will be discussed later [28].

> "If you need to ship [paper records to a] field office and do the data entry, there is a risk associated with safeguarding of children and anyone who entered this information. There is a risk associated with anybody who's holding this information."
>
> -NGO Representative, Syria

Safety & Security: Disadvantages

Enumerator and Respondent Safety risks: Safety risks exist for both enumerators and respondents. As in other low-resource settings, traveling with technological hardware for data collection on EiE activities can be a safety risk as the device itself has an inherent value and can turn enumerators into targets. Furthermore, a device that is used to collect information may also store information that could put its owner at risk; phones may be checked by local authorities to see where enumerators were coming from and enumerators may face interrogation about the contacts in their phones. Data that contains PPI needs to be handled with care especially if collected data contains information of respondents' political views. Particularly concerning is data that is connected to a geolocation in regimes where the government controls access to data transferred over mobile networks. [29] Although using technologies (especially large devices) in low-resource or conflict-affected settings can pose a risk for enumerators, some data collection technologies can reduce safety hazards. Surveys deployed either online or via SMS or IVR can be taken directly at home without the need of an enumerator. Data collected from social media activity

[6] https://tinyurl.com/usyyjxyc

is another way to collect information about a community without exposing enumerators or respondents to direct risk.

> "Some people won't want to use these technologies especially in … conflict settings because of the sensitivity around personal data … you can't always know what's going to cause harm ... What seems to you a very simple data collection exercise could be completely usurped by … somebody who views it with a very different lens. You had great intentions, I'm just collecting EGRA scores for primary school kids and somebody else has a nefarious use for those data."
>
> -Global Non-profit Representative (33)

Interviewees involved in the education response in Yemen emphasized how in this and other areas with active armed conflict natural questions from authorities and respondents around who is recording information, who this information is going to, and what this information will be used for are amplified by the presence of technology and its connection to surveillance. As a result, beneficiaries may be reluctant to sign into a device using any sort of identifiable information or refuse to provide biometric information. More invasive surveillance technologies, such as the use of CCTV, must be approached with a risk-benefit analysis; while it may mitigate on-site safety concerns and improve emergency response, it also risks infringing on the privacy of teachers and learners. For more detailed mitigation strategies to address safety issues, please see Dette et al.'s *2016 Toolkit on Technologies for Monitoring in Insecure Environments.* [30]

Risk of data co-option: The use of app data has particular ethical considerations around whether users of applications have a transparent view of how their data is being used and assurances that private data remains private. Any data collection faces the danger of co-option by host countries, foreign governments, or other organizations. Data could be abused for law enforcement or national security. Biometric data particularly could lead to invasive types of profiling to identify ethnic or racial groups; Oxfam's *Biometrics in the Humanitarian Sector* [31] provides a more detailed exploration of the inherent concerns around biometric data.

> "There are some things done in the Global South … which are clearly GDPR incompatible. And technically that's legal because you're not doing them in the EU, but it's not really a good idea … GDPR is a good standard and you'd be nuts not to try and stick to it."
>
> -Private Sector, Regional (9)

4.3 DATA PROCESSING TECHNOLOGIES

The processing phase includes data validating, cleaning, exploratory data analysis, and modeling. Of the three phases of the data life cycle, data processing was the least mentioned. The study team believes that this is true for two reasons. First, most processing and analysis happens at a regional, country, or global office and therefore does not face the same technological restrictions as data collection. The main restriction unique to education in emergencies data beyond the challenges of collecting it, is the availability of certain software programs. This may be due to government sanctions, local human capacity, or infrastructure. For example, Microsoft services cannot be used in Syria which reduces the availability of deploying PowerBI. Second, improved data validation techniques (such as enhanced data validation through programmed data quality checks) and advanced analytical techniques (such as deploying machine learning algorithms or other big-data solutions) are hampered by poor resource allocation to such endeavors, including insufficient human capacity to support and use software effectively.

> "My dream would be to have a team full of programmers to kind of do all the sort of data checking type stuff. [However] the amount of time it takes, especially for [a non-expert], is too great for the kind of value that is given to it by clients particularly."
>
> -Global NGO (10)

Nonetheless, technology is a clear component of this phase. It can be used to validate data (i.e. individuals assigned to validation roles in Kobo who address any data quality issues and confirm the responses), cleaning datasets (i.e. a script written in R used to re-shape raw data into a file that is ready for analysis or visualization), exploring data (i.e. simple pivot tables in Excel or distilling insights through a computer-assisted qualitative data analysis software), and modeling (i.e. statistical tests through data analysis software like SPSS or Stata).

In this section the study team will first provide a typology of data processing solutions and requirements for these technologies. While the advantages and disadvantages of specific technologies vary by type and actor, the study team provides clear themes as it relates to this process within the EiE sector with more detailed information about specific technologies found in the companion dashboard.[7]

4.3.1 REQUIREMENTS FOR DATA PROCESSING TECHNOLOGIES

Requirements for data processing vary greatly dependent on the technology. While all technologies require the use of electricity and do not require the internet or mobile networks to run, specific hardware requirements, and human resource needs vary widely as shown in Table 3.

Table 3: Requirements for Data Processing Technologies

Technology	Illustrative Software	Internet/network required	Hardware	Human resources
Spreadsheets	Microsoft Excel, Google Sheets, Numbers	No	Tablets or personal computer	
Data Analysis Software	SPSS, ArcGIS, QGIS, Stata	No	Personal computer	Data analysts
Computer-assisted qualitative data analysis software	Dedoose, RQDA, MAXQDA, Nvivo	No	Smartphones, tablets, or personal computer	Data analysts
General purpose programming languages	R, Python, , JavaScript, Julia	No	Personal computer	Data analysts
Machine learning techniques	TensorFlow, Apache MXNet, PyTorch	No	Personal computer (potentially cloud computing)	Data scientists

4.3.2 DATA PROCESSING TECHNOLOGIES – ADVANTAGES & DISADVANTAGES

Across all data processing technologies, respondents identified themes that highlight both the advantages and challenges of using these technologies within EiE constraints. These insights are categorized as they

[7] https://tinyurl.com/usyyjxyc

relate to general utility and costs. Because data processing technologies are generally conducted in a home office, safety and security issues are not discussed.

4.3.3 GENERAL UTILITY OF DATA PROCESSING TECHNOLOGIES

Data processing technologies enhance efficiencies, increase analytical prowess, and improve reproducibility and auditing capabilities.

General Utility: Advantages

Enhanced efficiency: As described in nearly all interviews, technology plays a clear role in improving the efficiency of data processing and analysis. Data validation that is not incorporated ad-hoc through data validation within a survey is often deployed post-hoc at the analysis phase. Rather than relying on tabulation of results to identify data quality issues via Excel or random audits, validation techniques can be developed through scripted data checks. For example, a script in R or Stata can be written to analyze student attendance data on a daily basis to identify if specific schools or teachers are continuously reporting above or below an attendance threshold. The deployment of more advanced technologies, such as machine learning techniques, can even further enhance the efficiency of data collection, data cleaning, analysis, and visualization techniques (see *Machine Learning – a new frontier for data analysis in EiE?* on page 19 for a deeper discussion of the promise of machine learning technologies in EiE.)

Analytical prowess: Scripted analysis approaches and data analysis software provide access to more advanced statistical techniques (particularly relative to Excel) including regression analysis, sample weighing, power calculations, spatial analytics, and machine learning techniques. Predictive models (such as identifying the probability of a student dropping out of school based on attendance data) can be developed and continually applied and modified. These types of student-based early warning systems have become standard operating procedures in many school systems globally but, due to the current nature of data collected and data quality, have not yet been fully explored in the EiE sector. Extracting insights from qualitative data, such as data from social media or communication channels, is made possible through advances in natural language processing and data analysis software. Machine learning (like other statistical techniques) can be used for predictive analytics. AI has already been deployed to predict natural disasters [32], as early warning systems on education metrics [33], and to discover informal settlements [34].

> "We don't do a very good job of being able to explain exactly how we got from A to B and in our analysis. And that's something that I think is the fault of both the data science side and the social scientists not doing enough to learn each other's way of working."
>
> -Academic (34)

Improved reproducibility and auditing capabilities: All technological approaches to data analysis enhance the ability to audit and learn from data processing and analysis. Some technologies, though, are better than others. Spreadsheet software can provide insights into the results of an analysis but often lack clear documentation of data manipulation and processing. Scripted languages improve visibility into analyses by allowing readers to reproduce an analysis and read through comments. Moving analyses from scripted languages back to a more readable format (such as a PDF, html file, or word document) can provide managers with methods to review analytical work. One such example is the development of an R Markdown file to show implemented analyses, display results of the code, and highlight text and comments related to analytical steps in the code as well as aid in interpretation.

4.3.4 COSTS OF DATA PROCESSING TECHNOLOGIES

The greatest barrier facing further adoption of data processing technologies in the EiE sector are cost and resource related. Particularly, the sector faces a dearth of specialized technologies designed for the sector, restrictions to specialized software due to cost, and barriers in hiring specialized expertise.

Access to specialized software is restricted due to cost: Proprietary software such as Stata, SPSS, or ArcGIS requires additional payment per user. In addition, access to licenses and regular software updates is often restricted in countries like Syria, which are subject to economic sanctions by the United States and the European Union. While open-source analytical tools, such as R or Python, exist these tools are designed for general purpose programming and, as such, are less intuitive and less commonly used in the social science world than proprietary software. This is exacerbated by the fact that existing technology paradigms are designed with commercial interests in mind: Increased technological efficiencies in the EiE sector have relied on advancements in the private sector. For example, while Translators without Borders is effective in using machine learning techniques to aid in the translation of languages used in humanitarian and development settings, transcription technologies still require further advancements in commercial voice to text technologies to be effective in local languages. This is hampered by the limited commercial market for the language of the communities being served. Advancements in the tech sector (such as in artificial intelligence) are designed with commercial interests in mind, which creates more costs to adaptation for the EiE context.

Specialized, sector-specific analytical support and data expertise is required: The implementation of advanced analytics requires a different set of resources, including data analysts, data scientists, and programmers: individuals who are in high demand in the private sector. The limited supply of sector-specific analytical support is further exacerbated by the fact that highly skilled experts may flee or be displaced during an emergency. Insights also need to be clearly 'translated' from the world of data scientists to the world of social scientists and practitioners; individuals with these experiences are few and far between. As one respondent stated, "Moving into ethics and data science means that we as researchers, practitioners, we need to develop the skills to interact with this field. And we don't have them at the moment."

TECH IN ACTION: IS MACHINE LEARNING A NEW FRONTIER FOR DATA ANALYSIS IN EIE?

Machine learning (ML) is an artificial intelligence technique that develops systems which automatically learn and improve from experience without being explicitly programmed. As the most common emergent technology discussed by interviewees, we briefly discuss how this technology may impact EiE in the future.

Why use machine learning?

Machine learning techniques can greatly enhance the efficiency of education data collection and analysis. An example is the work done by Translators without Borders in text translation. Techniques that would normally be done by hand can be first done through machine learning algorithms and adjusted with human revisions. These revisions can in turn improve the performance of the machine learning process. The output of this technique greatly enhances the efficiency of collecting data in one language and processing for use in another. Similarly, data cleaning and visualization processes can be enhanced and automated through machine learning techniques.

Even more promising is that AI tutoring systems (built on data collected and analyzed) can help identify exact competencies of children and present materials at the right level. While we were unable to determine actors working directly in the EiE space, companies like RoboTutor use voice recognition software to learn basic reading, writing, and arithmetic without adult assistance in developing countries. In the private sector, Squirrel AI is the first education provider in China operating at scale that uses a real-time adaptive learning system.

What are the challenges?

Human biases are built into ML models. Biases in the training dataset (such as traditional measurement, exclusion, association, or racial biases) will be amplified by the model. Machines are learning from the data that is given to them; if a model is learning off of men's speech, for example, it would be biased to understand men's voices more readily than women's voices. A model could similarly be fed data that reflects historical inequalities or biases, resulting in an output that has discriminatory features. This is exacerbated by the fact that some ML techniques are black-box processes (such as convolutional neural networks, or CNNs). Rather than a traditional input to output analysis, a convolutional neural network has a series of hidden layers that learn implicit features that are not humanly comprehensible but provide statistically useful correlations. Because ML experts can’t precisely see under the hood, it compounds the difficulty of translating their results for EiE experts. This in turn impacts the ability to elucidate assumptions and limitations of said algorithm. EiE researchers and practitioners need to learn the language of data scientists to incorporate ML results effectively.

ML also has unique privacy and transparency limitations. One of the simplest sources for data that could power a machine learning approach is data from an LMS. When a learner uses an LMS, lack of transparency on how data is used and the risk of data being re-purposed raises questions around privacy and transparency. This is further aggravated by the fact that rather than relying on internal resources, organizations can use providers such as Google and Amazon to implement Machine Learning as a Service (MLaaS) engines.

See DSEG’s *A Framework for the Ethical Use of Advanced Data Science Methods in the Humanitarian Sector* (2020) for further ethical considerations and guidelines to using ML techniques.

4.4 DATA COMMUNICATION TECHNOLOGIES

> "We still haven't quite completely evolved out of the dusty report method of sharing information and I don't know why we haven't … Because our donors require it, we still write that dusty report. But people don't use it. Don't even read the damn thing, honestly."
>
> -Global Non-profit Representative (33)

EiE data communication can roughly be divided into two categories of use: operational (planning, monitoring, coordinating) and strategic (evaluating, policy making, advocating) [35]. And while there is a clear use for technology in improving the way the EiE sector can use data, moving beyond the 'dusty report' into actionable information that supports data-driven decision making (DDDM), barriers faced in developing an organization that uses DDDM appears to be much more cultural than technical.[8] It comes as no surprise, then, that respondents highlighted *organizational culture* around data communication more regularly than discussing in-depth specific technologies. Outside of organizational culture, there were two main uses of technology that were highlighted. The first is leveraging technology to provide instant feedback to beneficiaries and stakeholders. The second is technology's use in data visualization, particularly related to dynamic dashboards.

The study team finds it particularly useful to delineate how technology is being used to provide instant feedback and in-field support data versus at the home office or for advocacy reasons. Of course, many tools are dual purposes: data collection for an EMIS system can be used for decision making at the national level, but schools and teachers can have direct access to the data they provided. The illustrative list below provides insights into how technology is transforming the way data communication technologies can be leveraged in EiE settings.

Instant feedback and in-field support

- Although primarily in use in development contexts (rather than EiE contexts), the Measuring Evidence of Quality Achieved (MEQA) dashboard by World Vision provides data on student enrollment, classroom environment, and school environment standards among other data. The MEQA app is used at the point of data collection to provide instant feedback to teachers.[9] These dashboards are used for reflection, program improvement, and reporting purposes.[10]
- FHI 360 has developed instant feedback teacher coaching tools for institutional use in Ghana and Nigeria, including conflict-affected Northeast Nigeria in the Addressing Education in Northeast Nigeria (AENN) program. This is a tailor-made ODK-based mobile data collection tool that walks a coach through a structured classroom observation form with a focus on fidelity of implementation and quality teaching, and then provides tailored instant feedback for coaches to use with teachers in a post-observation coaching session. The data aggregates up into dashboards for use within programs or government systems.
- Waliku is a teacher-facing and system-facing mobile app developed for tracking student attendance in low-resource contexts.[11] At the teacher level, the application notifies teachers about students with chronic absenteeism and in-app features provide guides for counseling these students. The school admin application provides support for school leaders to manage population records with automated reports on absences and the underlying reasons for these absences.
- UNRWA's EMIS aggregates data both at the agency-level for advocacy and strategic purposes but is also available to school administrators for decision making locally. Three modules (student, staff,

[8] See the Harvard Business Review *10 Steps to Creating a Data-Driven Culture* for tips on supporting DDDM organizationally.
[9] https://play.google.com/store/apps/details?id=org.smap.smapTask.android.meqa&hl=en_US&gl=US
[10] https://www.meqadata.com/
[11] https://www.waliku.org/index.html

and premises) operate at four levels (headquarters, field, area, and school) with reporting available to all levels relevant to access rights. With the addition of a new EiE module, UNRWA education staff can track indicators including levels of student access to learning during emergencies, the number of school days lost due to an emergency, the percentage of schools with emergency preparedness plans, the number of education staff trained on EiE topics, the percentage of students who participated in recreational activities or attended counseling sessions.

- RefugePoint and Vera Solutions developed a case management system to work offline in refugee camps in Kenya. This allows field staff to access previously recorded information on beneficiaries in the field, update this information, and sync data to be aggregated via a Salesforce dashboard.[12]

Traditional reporting and advocacy

- Humanitarian Data Exchange's Education in Emergencies dashboard aggregates over 1800 downloadable datasets, searchable by location, organization, format, and other filters, and displays basic charts and graphs previewing their contents.[13]
- OCHA's dynamic dashboards provide information on the Education Cluster's humanitarian response.
- 'Flash' reports created by IMMAP to share weekly explosive reports in Iraq via simple infographics.[14]
- Save the Children's IDELA Data Explorer provides data visualizations on results of the International Development and Early Learning Assessment across 100+ datasets.[15]
- OpenEMIS started as a monitoring system for tracking educational data on Syrian refugee children in schools and education centers in the Za'atari refugee camp and then expanded in scope to provide school-level management systems and annual reporting throughout the Kingdom of Jordan.
- Translators Without Borders language dashboards provide detailed information on the distribution of language use within a country or a region.[16]

In this section, the advantages and disadvantages of using technology to support data communication are discussed. For detailed information on specific technologies, please reference this paper's companion dashboard.[17]

4.4.1 REQUIREMENTS FOR DATA COMMUNICATION TECHNOLOGIES

Table 4: Requirements for Data Communication Technologies

Technology	Illustrative Software	Internet/network required	Hardware	Human resources
Dynamic data visualization	Microsoft PowerBI, Tableau, Google Data Studio, shinydashboard, Salesforce	Depending on function	Personal computer	Data analysts

[12] https://www.verasolutions.org/portfolio/refugepoint/?locale=en
[13] https://data.humdata.org/dashboards/education-in-emergencies
[14] https://reliefweb.int/sites/reliefweb.int/files/resources/weekly_explosive_hazards_incidents_flash_news_dec_17_-_23_2020.pdf
[15] https://data.idela-network.org/
[16] https://translatorswithoutborders.org/language-data-nigeria/
[17] https://tinyurl.com/usyyjxyc

Static visualization	Adobe Illustrator, ArcGIS, Quantum GIS, Microsoft Excel	No	Personal computer	Data analysts

The two main technologies described by respondents were dynamic and static visualization software. While neither of these technologies require the internet for developing data visualizations, depending on the approach to sharing the results these requirements would change. For example, while a dashboard can be developed in Microsoft PowerBI on a local machine not connected to the internet, sharing this information via the PowerBI app on a users' smartphone would require the internet to refresh the dashboard with new data. These requirements are seen in Table 4.

4.4.2 DATA COMMUNICATION TECHNOLOGIES – ADVANTAGES & DISADVANTAGES

Across all data communication technologies, respondents identified themes that highlight both advantages and disadvantages of adopting data communication technologies in terms of general utility and cost.

4.4.3 GENERAL UTILITY

Data communication technologies enhanced utility through supporting the automation of reporting, real-time engagement with data, the development of management and advocacy tools, and the localization of analysis. Barriers to continued adoption include (paradoxically) the difficulties in promoting internal use of a tool and once adopted, over-reliance on developed tools.

General Utility: Advantages

Report automation: Reporting can be increasingly automated through technology, saving time and resources in the long run. UNRWA, for example, ties attendance to unique IDs, which are used to calculate indicators used for regular, automatically generated reports on drop-out rates, grade repetition, etc. This practice can extend to the development of early warning systems.

Real-time engagement with data at multiple levels: Through dynamic visualization, rapid and real-time insights can be provided. This includes supporting real-time decision making, developing program and process monitoring tools, and generating instant feedback loops for a teacher or school. This is particularly important in an EiE setting. One example comes from FHI 360's instant teacher coaching system developed for education programs in Ghana and Nigeria [36] or World Vision's MEQA system. After entering data following a classroom observation into a data collection tool, a set of important feedback points are curated automatically, and the observer provides guided feedback immediately to the teacher.

> "Technology helps us make decisions about where to apply resources; so you know what students should we invest in more because we expect they might have problems because they're behind in learning, and either hitting early milestones or their educational performance isn't up to their grade level … if we weren't doing it electronically, it would take much longer, it would take more resources in the end and it probably wouldn't be used because it would be reported maybe once a semester."
>
> -Global Humanitarian Organization (11)

Advocacy tool: High-level visualizations can be used for advocacy in a number of ways. They can provide supporting evidence about the impact or reach of a program or highlighting the differences between a treatment and control group. They are also imperative for resource provision, as in the case of enrollment information from EMIS systems. Technology increases the frequency with which this data can be reviewed and acted upon. Technology can also provide a means to aggregate and conduct gap analysis across projects or can support 'zooming in' on smaller NGOs.

> [We can greatly improve by just] understanding how to represent data in a meaningful way … It doesn't necessarily require better kind of like quantitative skills or something. It just requires data literacy."
>
> -Global NGO (10)

Localization of analysis: Dashboards provide the ability to zoom in on individual contexts. This ability to dig deeper (or drill down into various datasets) can support data driven conversations with the individuals closest to program implementation. The aforementioned Translators Without Borders dashboard shows an example where a user can zoom in on language needs within specific communities.

> "We made use of the calculation capabilities of ODK so that based on the input that the form receives, feedback comes out … a monitor will be able to see at the end of [their] observation, the key strengths, the key recommendations, and the key challenges that specific teacher or that environment is facing so that the visit becomes beyond data extraction and becomes a mentorship process or a coaching process."
>
> -Humanitarian Organization, Global (7)

Management tool: A dynamic visualization tool can be developed as a management tool. This could be through providing accountability and visibility into program reach by displaying heterogeneity in various indicators across operators or by identifying students who need more support. Technology-enabled processes, such as texting out the results of an audit to a community or sharing data with school leaders via a dashboard, can further develop accountability mechanisms within a community or within a school. A global humanitarian respondent stated that dynamic visualization provides a "much more rapid and more accurate and relevant disaggregation and comparisons across groups. And we use that … to compare different locations, to compare data from primary and secondary schools, to compare data for male and female student populations, to compare according to how long we have been supporting a certain school, etc. in a way that helps us inform program management decisions."

General Utility: Disadvantages

Inherent difficulties of getting teams to adopt new tools: Technology-based tools need to be both practical and fit into existing workflows. Design thinking methods, a human-centered approach to redefine problems and create innovative solutions through continued iteration, can be used in the development of these tools to ensure they fit local needs and make existing processes more efficient. If the tool's aim is to change workflows, it needs to be developed within a larger change management process. FHI 360's work on using coaching tools within government systems in Nigeria and Ghana demonstrate how well-designed tablet-based tools can make workflows easier for school supervisors and build motivation at the same time [36].

Over-reliance on technology: While technology can help support and improve processes, there is always a risk of becoming overly dependent on a technology. A classroom observation tool, for example, may help identify coaching tips after entering data about an observed lesson. Such automatic tips can contribute to behavior change—building trust and accountability in a new system, new methods of teaching and learning—and can support populations with lower levels of overall technical qualifications. However, they are limited in supporting deeper reflection. While some people would use these tips as the springboard for a deeper conversation, using their own expertise to provide richer feedback, other individuals will deliver only the barebones guidance provided by the technology.

> "I feel like with some people because they have this tool, [they think to themselves] 'I read out the strengths. I read out the recommendations. That's it. My job is done.' Rather than seeing it as part of the bigger process."
>
> -Humanitarian Organization, Global (7)

4.4.4 COSTS

Although current costs for proprietary software (such as Tableau or PowerBI) may be too high for an NGO, most software has free or partially free versions. When trying to deploy these technological solutions at scale, such as sharing dashboards at the school level, these free versions are no longer viable. Human resources are also a challenge, as the adoption of improved data visualization software requires data visualization and design specialists. This will change, however, as technology becomes more intuitive through advancements in data visualization software and more and more non-technical staff have the requisite skills to build and maintain dashboards.

5. DISCUSSION AND RECOMMENDATIONS

In this section a set of cross-cutting guiding principles are proposed to help inform the decision to adopt a new technology and present recommended practices for EiE data technologies based on the analysis of the findings.

1. Section 5.1 lays out the general guiding principles for data collection, processing, and use. This section is divided into two parts: principles for practitioners and considerations for funders.
2. Section 5.2 discusses the recommended practices for each process within the data life cycle.
3. Finally, Section 5.3 explores potential new frontiers for technology in data collection, analysis, and use.

5.1 GENERAL PRINCIPLES FOR DATA AND TECHNOLOGY IN EIE

5.1.1 PRINCIPLES FOR PRACTITIONERS

The following principles can be applied across all phases of the data life cycle. These principles help guide practitioners' decisions to adopt new technologies in the EiE sector. The first two principles can act as a framework for all decisions. Principles 3, 4, and 5 help practitioners identify when a new technology is appropriate. Principles 6, 7, and 8 help practitioners in deploying the technology successfully.

1. **Do No Harm.** Do No Harm principles are the minimum standard of practice to avoid causing inadvertent harm. Decisions on what technologies are implemented need to be made within an understanding of the underlying context and how the technology interacts with the context, as well as implemented in a way that avoids negative impacts and maximizes positive results. Be conscious about how technology will be perceived in a given context. Ensure that any research, study, assessment, or situational analysis conducted in EiE settings, especially with minors, is reviewed by a certified Institutional Review Board (IRB) to ensure safety and ethical treatment of participants and enumerators concurrent with state and international law. Resources on child safeguarding can be found on INEE's resource collection on child protection[18].
2. **Follow GDPR and related standards.** The GDPR is a regulation in EU law that protects data privacy and establishes rules around data sharing. Although it may not be possible to follow these

[18] https://inee.org/collections/child-protection

regulations at all times, procedures related to consent for each data collection and processing instance should be established. When this consent is not possible to acquire, justifications around protecting the vital interests of a data subject or other individual or demonstrating the need to process the data to perform a task in the public interest can guide the correct use of data. OCHA's *Data Responsibility Guidelines* [37] and the Humanitarian Data Science and Ethics Group (DSEG) *A Framework for the Ethical Use of Advanced Data Science Methods in the Humanitarian Sector* [38] also provide guidelines for general data responsibility and advanced methods respectively.

3. **New technology does not mean better technology.** Practitioners should not be focused on specific tools (especially as some tools may be sanctioned in EiE countries). Instead, approaches need to be designed with the local context in mind. For example, rather than developing an application for beneficiaries to report access to provided services, existing communication channels that are prevalent in the community should be leveraged to collect this data, such as WhatsApp or SMS. Of course, the technology ecosystem within a community can change rapidly and should be assessed on a regular basis to adapt approaches to the context.
4. **Coordinate among stakeholders, adapting technologies using a systems-thinking approach.** Technology adaptation should move beyond the paradigm of project-based investments. Many activities in the EiE sector have a project-based approach that does not look at the education system as a whole or work toward cooperating with the national system. When governments have an existing system, solutions should be built for this system. When they don't, solutions should not simply replicate those of other countries, but be designed and costed for the local context to ensure that support does not start and end with the funding. Before making a major investment in a technology, reach a consensus among stakeholders on the intervention, targets, and the technologies needed to measure these targets to ensure that impact is maximized.
5. **Develop innovations in collaboration with local organizations and end-users.** Identifying technologies centrally from headquarters to solve local issues continues to be the norm. While this is not unique to the EiE sector (see the Disasters Emergency Committee response review to the 2015 Nepal earthquake as an example [39]), the overwhelming role donors and large multilaterals play in supporting education in emergency efforts can result in top-down and one-size fits all approaches dominating the conversation. However, some of the most impressive uses of technology (such as the widespread adoption of WhatsApp and its role as a new data source) were initiatives led locally and organically. Although global actors have a role to play in introducing new technologies and innovations that may have great potential but are not well known in local contexts, local engagement almost always will be more conflict-sensitive, more contextually appropriate, and will encourage local ownership of data leading to greater use.
6. **Introduce new technology as social and behavior change.** Understanding and deploying a new technology may be difficult but it is even more difficult to change human behavior.[19] When deploying a new technology, include thoughtful training, coaching, ongoing support, and smart incentive systems, and foster the leadership and government buy-in needed for success. For example, if a new data system requires constant data entry by a teacher, after an initial training, teachers will need to have ongoing coaching and support to ensure they can use the system. Incentives for teachers, school leaders, and school districts need to be aligned to promote consistent interfacing with the new technology. Finally, continuous engagement and advocacy with the government is required for the intervention to be sustainable.

[19] Venkatesh et al. (2008) provide insights into managerial decision making on how to support the adoption of technologies, including suggestions for interventions prior to and following the deployment of a new technology as well as resource allocations to support adoption.

7. **View technology as a long-term investment.** While investing in a new process may be costly, in the long run, it will lead to efficiencies and cost saving. This is an investment not solely in the costs for software development or license procurement, but also costs related to hiring the right staff to develop, maintain, and leverage insights from the deployment of a new technology and costs related to staff capacity development and change management.
8. **Nurture a culture of data feedback loops and DDDM**. Take steps to get data back into the hands of stakeholders across the project – data will be more powerful if it is closer to the hands of those who can use it. While the EiE sector has been successful in using technology to compile and share secondary data for advocacy and strategic purposes, it has been less successful in using technology to return primary data to impact operations.

5.1.2 CONSIDERATIONS FOR DONORS

The current deployment of technology for data collection, processing, and use would benefit from stronger guidelines by funders to help support a system-wide infrastructure to best leverage data in the EiE sector.

1. **Incentivize organizations to adhere to the aforementioned guidelines.** Guidelines for practitioners must be set at the system level. Funding agreements should include recommendations or requirements to follow GDPR or OCHA's data responsibility guidelines. Organizations should be incentivized to co-develop innovations with local partners and end-users rather than proposing new technological systems and demonstrate how this data is returned to individuals in the field.
2. **Incorporate technology for data collection, processing, and use into wider coordination efforts.** The use of technology for data collection, processing, and use should be guided by global EiE coordination efforts. Data-driven coordination on EiE helps (i) avoid conflicts or duplication, (ii) identify opportunities for collaboration, and (iii) optimize the distribution of activities. Incorporating technology for EiE data into these global efforts and mandates will help harmonize how technology is used to produce meaningful data for coordinated education activities and will also help make data more accessible and interoperable.
3. **Provide proportionate funding for data collection, processing, and use during humanitarian crises.** Effective deployment of technologies to improve data collection, processing, and use can be costly, and donors may underestimate these costs. Human resources—teams to capture, process, and generate useful insights from data—are expensive. Finding local talent may be more difficult due to the impact of displacement on highly skilled talent. Certain technologies, such as implementing machine learning techniques into data analysis, can appear prohibitively expensive. The use of proprietary tools impacts the price of a technology deployment (with sanctions for specific tools in turn impacting access), as does investing in open-source software development.
4. **Develop interoperability guidelines.** The development of interoperability guidelines support research within and across education systems; data integration further supports the development of insight harvesting methods that require larger amounts of data, such as machine learning techniques. Interoperability guidelines include frameworks for data management and governance, standard classifications and vocabularies, and rules around data and metadata modeling (see Morales and Orrell (2018) for interoperability guidelines for the development sector). As an example, guidelines on data collection for an implementing partner working on capturing attendance data would ensure data is identified within the standards of the larger national EMIS system, that captured fields follow agreed-upon naming conventions and definitions, and accompanying data dictionaries are published.
5. **Support the development of new software and the continued use of existing software**. Donors can both fund open-source software and help reduce restrictions on software use. Open-source software is a type of computer software where source code is released allowing users the

rights to use, study, change, and distribute the software to anyone. The most common open-source software used in the EiE sector is the Open Data Kit (ODK) and other XForm-based data collection platforms. Users can continue to adapt open-source software to address contextual needs but creating and developing a community to iterate and improve is difficult when users in the EiE sector cannot contribute on a technical level. The development of open-source software for EiE contexts, such as developing dashboard or case management software that works offline to improve data use in the field, requires funding from donors to move beyond project-based solutions to solutions that can be adopted across the EiE sector. Donors should also be aware of the impact of sanctions on specific software, currently a particular concern in Syria. Donors should advocate for sanctions to not negatively impact EiE interventions, particularly as it relates to data collection.

5.2 WHAT ARE RECOMMENDED PRACTICES FOR USING THESE TECHNOLOGIES IN EIE?

In this section, recommended practices revealed through the interviews for each process within the data life cycle are synthesized. When possible, practices are separated between strategic and operational suggestions.

5.2.1 RECOMMENDED PRACTICES FOR DATA COLLECTION TECHNOLOGIES IN EIE

Strategic Recommended Practices

Address direct and maintenance costs upfront. Before engaging in data collection efforts, ask how it will be sustained moving forward. How much will it cost to develop up front? Who will maintain data collection? Cost-benefit analyses and the demonstration of successful widespread adoption in other organizations highlights the long-term benefits of investing in technology while realistically addressing costs. If relevant, link this work to local infrastructure, such as through the co-development of sustainable MEL systems that can be transferred to the government or other local actors upon project completion.

Consider coordination and data sharing in advance. Data collection effort duplication (such as going to the same location to collect education data for one study and sanitation data for another) is time-intensive and costly. Leverage the investment in collecting data for one purpose to collect data that is needed either internally or by partner organizations. Continue the push towards the New Way of Working (NWOW) by coordinating data efforts with humanitarian, development, and other actors to deliver collective outcomes. Consider the technical infrastructure required for data to be shared so someone else does not need to collect the same data next week. Tap into existing networks to best understand the data landscape and what has already been collected at the national and sub-national levels, thinking about the appropriate level to collect and share data while designing the data collection system. Follow guidance for global public goods in open infrastructures and other global interoperability standards as developed by collaborative bodies, such as the Inter-agency Network for Education in Emergencies (INEE) Reference Group.

Research the technological landscape. As penetration rates of individual technologies vary between and within countries, when selecting a technological solution that reaches out directly to beneficiaries (such as through WhatsApp or SMS) research access barriers related to gender, economic status, disability status, vulnerable group status, etc. before settling on a technology.

Conduct routine evaluations of technologies. Because technologies adapt quickly, establish routine evaluations of the technologies used to determine whether they are still appropriate within the current context. In particular, evaluate the tool's efficacy in terms of collecting the data required, the experience of the data collector, and the experience of the respondent.

Operational Recommended Practices

Duty of care is first and foremost. This includes both the safety of enumerators and respondents as well as ensuring data collected is kept safe and secure. Obtain approval from local and international IRB boards for all research conducted in EiE contexts. Deploy multiple modes of collection when in-person data collection may be dangerous for enumerators or respondents. Organizations have adopted GDPR standards for data protection and privacy even when this is not a requirement from a donor. Many applications (including Kobo Toolkit and ODK Collect) allow data that is stored for an individual to be encoded numerically at the point of collection and then wiped from the device as soon as the survey is synced. Protocols for destroying hardware can be developed when necessary. Row-level security is a feature of some technological solutions that filters data views depending on the user to restrict access on a need-to-know basis (i.e., as assigning data quality analysts access to solely data for their own region), which reduces database exposure to an unauthorized disclosure of personal data.

Deploy multiple modes of collection. As well as the aforementioned benefit of addressing potential safety issues, multiple modes of data collection reduce sample bias. For example, one organization assessing learners in non-formal learning centers used IVR for the majority of respondents but had a secondary arm of data collection that included in-person visits and tablet data collection in locations with poor network access.

Ensure clear design for high response rates. Surveys that are designed well, and familiar to the user, will have higher response rates. Be wary of collecting data for the sake of collecting data. As technology enables practitioners to collect and process more and more data, the sheer amount of information can be overwhelming. While technology may make it easier to add just one more question, refrain from over-saturating surveys with content. One practice to limit the collection of un-needed data is the use of a theoretical framework in justifying any new collection efforts. Limit the length of the survey as much as possible, though length is less of a factor if incentives are provided [40]. As much as possible, deploy the survey so it is free for respondents (such as paying for data when a respondent uses WhatsApp to take a survey) or provide a small incentive for sharing information.

Tailor technological solutions to the culture in which it is being deployed. Data collection may be impacted by cultural norms; technological solutions and processes need to be tailored to these norms. For example, taking photos of school buildings may not be acceptable to local military authorities. Satellite imagery or descriptive school environment surveys can be deployed instead.

Use familiar technologies when possible. For data collection directly from beneficiaries, technologies that are more familiar to the respondent are more appropriate and facilitate adoption. IVR, for example, runs on any kind of phone and, if well-designed, can be deployed to non-literate populations.

Conduct regular resource assessments. The data collection tool landscape is continuously changing, with new tools, new features, and changes in the accompanying prices and human resource requirements. Develop regular resource assessments to identify the most cost- and human-resource efficient tool within the specific context.

5.2.2 RECOMMENDED PRACTICES FOR DATA PROCESSING TECHNOLOGIES IN EIE

Strategic Recommended Practices

Invest in more analytical support. Generating high quality insights in the humanitarian sector can be challenging due to the lack of quality data: investing in resources to improve data processing leads to higher quality data and clearer demand for advanced analytics. Clear pathways to integrate advanced analytics in existing work streams and improved incentives to recruit this talent is required to strengthen analytical prowess in the EiE sector.

Obtain clarity on which data must be supplied to a client or government. At the beginning of the project, ensure that it is clear what data is collected and at what level this data will be reported at. Choose tools that ensure data can be quickly anonymized and de-identified, especially as it concerns qualitative data.

Build capacity of staff to leverage data. Capacity building investments to ensure data scientists and practitioners can speak the same language are required to ensure insights and evidence are successfully communicated and applied. For example, if a machine learning technique is used to help identify at-risk learners, a practitioner should have the knowledge to engage with the algorithm's assumptions to ensure these assumptions match ground realities.

Build expertise in multiple platforms. Costs for data processing technologies shift as do the availability of platforms in specific contexts. Over-reliance on specific tools can negatively impact operations due to disruptions in user agreements or new sanctions prohibiting services in specific countries. Investments in the use of open-source software can increase analytical expertise without the cost of deploying proprietary tools. As a rule of thumb, there is no one-size fits all solution; policymakers need to be flexible in exploring multiple technological solutions.

Operational Recommended Practices

Use technology to ensure data is accurate and timely. Quality data is needed for national and local governments to respond to crises and develop strategic plans. Technology has a clear role in fortifying data processing to improve accuracy (through data quality checks) and timeliness (through shortening the length of time between data capture and data use).

Automate data validation processes. When possible, automate continuous processes, such as checking on data quality, to further increase efficiency of data processing procedures. Data quality issues are best addressed at the point of potential error; creating automated quality flags can address issues in a timely manner and improve overall data quality. Standard validation processes enhanced by technology include the incorporation of geolocation checks, outlier identification, survey duration, and internal consistency checks.

Develop processes to share analyses. Technology encourages sharing analytic processes and deployed software both within an organization and across the EiE sector. Ensure work is reproducible through deploying best practices across all analytic software. When possible, share analytic techniques and approaches through Github or other collaborative forums.

Triangulate data sources. When possible, triangulate data sources using both quantitative and qualitative methods to show results of an analysis. Data triangulation can also be used to evaluate the effectiveness of a specific tool; when determining when whether a new data collection tool can substitute an existing process, triangulate the results of the collection with both the original and new tool to validate the new tool's efficacy.

5.2.3 RECOMMENDED PRACTICES FOR DATA COMMUNICATION TECHNOLOGIES IN EIE

Strategic Recommended Practices

Approach the use of new tools as change management. Users will not adopt a new tool in a vacuum. Getting a number of tools and systems to talk to each other is a heavy lift, and these efforts can be ignored if people resort to what they know (such as spreadsheet software). When deploying new tools or ways to look at and use data, a detailed change management approach is essential.

Identify data champions. A culture of DDDM needs to be built at the top to set expectations that using data in decision making is the norm not the exception. One example is the AENN datahub model where four roles (Data Manager, IT Support Lead, Data Use Advisor, and Data Collection Manager) were identified to ensure that the deployment of new technology had local ownership and local champions.

Other respondents described the successful adoption of new systems as particularly likely when a client has an executive sponsor.

Create a culture of data sharing. Whenever you get a chance to share data, either internally or externally, share it (with the obvious caveat of following data security procedures and removing PII!). This helps build experience and data literacy. As some individuals have lower levels of data literacy, deep dive with partners on specific issues with salient features that were uncovered during data analysis to scaffold learning through data.

Operational Recommended Practices

Provide data in a timely fashion. Provide relevant data to the appropriate stakeholders. Data related to teaching and student performance can be provided to teachers to help improve their practice. Aggregated data on students and teachers and school environment data can be shared with school leaders. Technology should be used to facilitate information sharing to key stakeholders in a timely fashion.

Follow data visualization best practices and simplify reporting. Data literacy acts as a barrier to interpreting information; best practices are needed to present data in meaningful ways that allow users to quickly grasp key information. Schwabish's *An Economist's Guide to Visualizing Data* [41] and Tufte's *Envisioning Information* [42] are two great resources for improving how information is presented. Schniederman's seminal *The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations* [43] provides a strong foundation for understanding how best to design graphical interfaces, such as dynamic dashboards. Boil down information to the essential parts (such as a one-pager report or a single visualization) and use innovative ways of reporting such as including the bottom-line up front with a highly visual four-page summary.

Use row-level security to limit access to data. Dashboards are effective when they show detailed information, but this could potentially be sensitive especially in an EiE context. Control who has access to specific data through row-level security. This can be hierarchical in nature, for example a teacher can enter data and read data for his or her class while the headteacher can see information about their school.

Ensure systems are maintained. Keep visuals and systems up to date. When data is no longer refreshed, users stop using the data source. Visuals should include information on when they were last updated and should have clear mechanisms for reporting any errors found in either the system or underlying data.

5.3 WHAT ARE THE GREATEST OPPORTUNITIES FOR TECHNOLOGY IN EIE?

While the study team set out to gather information on what technologies are being used by practitioners in other related sectors that could potentially be leveraged for EiE, respondents were more likely to highlight *how* to best leverage technologies rather than *which* technologies. Based on interviewees' responses, the greatest opportunities include expanding use of existing platforms to collect data, empowering beneficiaries, field staff, schools, and teachers with data through data feedback loops, and embracing promising emergent technologies such as data from learning management systems, automation of data cleaning and analysis, deploying machine learning techniques, and chat-bots and enhanced information access.

Embracing promising emergent technologies. While early on in the research it became clear that most practitioners focus on adapting technologies to be appropriate for the local context within resource constraints rather than focusing on cutting-edge technology (and, in fact, that continuously updating systems to the newest technology could have negative repercussions on sustainability), the technologies in Table 5 hold the greatest potential in benefiting the EiE sector based on their current use in other sectors as well as alignment with the guidelines laid out in 5.1 General Principles for Data and Technology in EiE.

Expanding use of existing platforms to collect data. Instead of relying on deploying new technologies to collect data, using existing platforms to harvest insights pertinent to monitoring and research questions on hand is a clear opportunity for the EiE space. For example, the proliferation of social media for communication (such as WhatsApp or the use of Facebook) provides the opportunity to collect data via surveys within these platforms. UNDP's *WhatsApp Surveying Guide: Lessons Learnt from Two Qualitative WhatsApp Surveys in Lebanon* [44] demonstrates the effectiveness of this human-centered design approach while also showcasing enhanced time and cost effectiveness and discussing limitations of this approach. This could be further improved by incentivizing data collection. Data collection through digital ledger technologies can be monetized allowing small financial incentives in exchange for reporting; this approach can in turn make MEL processes more equalizing and less extractive [27] and place data collection in the hands of a trusted community member rather than an outside data enumerator.

Table 5: Promising Technologies

Process	Description	Quote	How will this impact EiE?
Collecting Data	Collecting Data through LMS: Data collected through the use of Learning Management Software as learners directly interact with learning technologies.	"I feel like the systems for teaching children…applications and devices… [are just now] getting to the stage where they are starting to work, and they might work more because of the use of Artificial Intelligence to do recommendation of the next questions that should be given or doing things like recognizing writing to coach children." Humanitarian Organization, Global (11)	Insights into how learners are engaged with content including completion rates, embedded assessments, and user feedback can help guide how technology-driven approaches to providing education can be the most effective within specific contexts.
Processing Data	Automated data cleaning and analysis: The application of machine learning techniques to accomplish data cleaning objectives (such as estimating missing values or dealing with outliers) and conduct exploratory data analysis.	"AI that will help you walk you an analysis, saying that if you're collecting this data, that data, this data [you do XYZ]. So, you need less of statistician support. It will walk you through those steps and then you can generate and then we click a button and then it will help generate some of the statistics for you." Academic (21)	Data cleaning is generally the most resource-intensive part of data analysis. Speeding up these processes can shift more attention to leveraging insights from analysis. Tamr and Inductive are two companies that are working toward this goal.

	Sector-appropriate machine learning techniques: Artificial intelligence technique using computer algorithms that improve through experience.	"The big promise [in AI] is you can have economies of scale, you can better serve kids in a larger fashion, particularly when they don't have the privilege of having a very good teacher in front of them. You can …provide extra time for learning after the class is happening [and provide a] customized learning experience." Academic (35)	Improved AI techniques can help classify unstructured data (such as information collected through social media) or help predict natural or man-made disasters. Enhancements in AI tutoring systems can develop new modalities to deliver lessons and materials targeted to individual learner needs.
Using Data	Chat bots and enhanced information access: Chat bots (textual or via voice recognition) deployed to provide information to families in refugee camps or emergency settings.	"We would like people to be able to walk up to a box in the camp where they're living and speak to it and that box will speak back. So, using a kind of chatbot function, but transcribed and translated … and the humanitarian organization [is able to provide] spoken feedback." Humanitarian Organization, Global (15)	Advances in natural language processing can be leveraged to provide curated support to beneficiaries through chatbots or an information access center. Imagine a refugee asking a chatbot (in their native tongue) how they should register their child for secondary school.

Empowering beneficiaries, field staff, schools, and teachers with data through data feedback loops. Technology can empower beneficiaries to access data collected; this could be through individually tailored reports, dynamic visualizations accessible across school systems, or immediate feedback at the point of data collection. The aforementioned examples of AENN and World Vision using collected data to guide coaching discussions demonstrate how technology can immediately bring data back to teacher practices. UNRWA's EMIS system, with key data and reports available to staff at the regional and individual school level, ensures that this data can be used for local decision making rather than solely higher-level strategic decision making.

Clear opportunities exist for the further integration of technology to enhance EiE's capacity to collect, process, and use data, even if technology is not a panacea. The proposed guidelines and best practices outlined above can act as the framework needed to avoid the negative impacts of technologies while maximizing positive benefits. Finally, the study team notes that this work is solely the foundation for practitioners and donors to approach the use of data collection, processing, and communication technologies. Regular evaluation of the technology landscape as well as guiding strategies for a country or region's adoption of technologies in the EiE space are needed to ensure the successful integration of data technologies in EiE.

APPENDIX

APPENDIX 1. INTERVIEW PROTOCOL

FHI 360
MEERS – EiE Tech Data - Interview Guide

INTRODUCTION:

Welcome. Thank you again for agreeing to speak with us. Let's start by introducing ourselves. My name is _____________. I work for ____________________ through the Middle East Education Research, Training, and Support Initiative (MEERS). We are doing a research study for which we are interested in learning more about how EiE organizations and professionals use technology for data collection, analysis, use and visualization – specifically, we are looking at the intersection of technology, EiE, and data.

For the purposes of our conversation, when we say "Education in Emergencies" we are referring to education for those affected by disasters, conflict, and displacement. When we speak about "technology" we are interested in both hardware and software.

CONSENT:

[Read consent form. If a participant chooses not to give consent, he / she cannot participate in the interview and the interviewer must stop the interview.]

Did you get oral consent from this participant? ___ Yes ___ No

INTERVIEW:

Demographics

First, we would like to ask some information about your background and your work.

1. What is your current position within your organization? What are your major responsibilities?
2. How long have you been working in the field of education?
3. How long have you been involved with work related to education in emergencies?
4. Would you consider your work to focus on a local level, national level, or global level?

Technology in EiE

Now we would like to ask you about how your organization uses technology.

1) Does your organization use technology to collect and analyze data related to EiE?
   (This could include using technology to systematize and analyze existing data and information (e.g., data mining techniques for social media or newspapers).
   a. If yes, can you tell us more about the specific technologies and how you are using them?
   b. How does technology assist you with data collection?

2) What are the advantages of using these technologies? What prior issues or concerns is the use of technology addressing?
   a. Does the use of open-source software offer benefits in education data collection? If yes, how?

3) What requirements (such as electricity, internet connectivity, other technologies, etc.) must exist to enable the use of the current technologies that you are using?

4) What are the costs or unintended consequences of using these technologies in data collection, use, and analysis?

a. How are concerns around privacy and security accounted for in data collection processes?
b. What measures are in place to mitigate and avoid these fallouts?

5) What are barriers to the application and continued use of these and more data collection technologies in EiE contexts?
   a. Do you find any barriers in terms of interface preferences across different national, organizational, or individual cultures?

6) With attention to these barriers, what are recommended practices for using these technologies in EiE contexts?
   (This could relate to cultural and contextual appropriateness, infrastructure requirements, time and resources required, professional capacity, bias reduction, and safety and security.)
   a. Are there recommendations/requirements that are unique to the MENA region?

7) Do you know of any other promising technologies that could be used to supplement the tools you are already using? If so, which ones and how will they help?

8) Do you know of other technologies in related fields that could be relevant to EiE data collection, analysis, visualization, and use?
   a. If so, how could they be integrated into EiE data collection processes?

9) Is there any other information that you would like to share with us?

10) Is there anyone else we should speak to?
   a. In particular, do you know anyone from the ministry of education at the level of the districts or municipalities where you work?

APPENDIX 2. EXAMPLE OF TECHNOLOGY IN USE

The following narrative showcases how technology has been incorporated in the MEL processes of a real organization active in the field of EiE:

Taalim (name has been changed to avoid disclosing information that could compromise the organization's activities) is a Syrian community-based organization that provides psychosocial support and non-formal education in three provinces in Syria. As of October 2020, they operated 125 learning centers, reaching close to 33,000 children in need.

Faster data collection, even when offline

Taalim periodically conducts assessments on children's psychological wellbeing, as well as their literacy and numeracy skills. These were traditionally conducted on paper, but recently the organization transitioned to Kobo Toolbox and Formera, two platforms which allow survey creation and rollout via Android tablets and mobile phones. As a member of Taalim's MEL team explains, technology has equipped them with a "pre-analysis advantage" since now they can have an update on the data that you are collecting instantly and see what your progress in the field is."

Taalim relies on a team of enumerators and learning center staff for data collection. While the organization typically provides staff with tablets, they also rely on staff's personal devices (Android smartphone ownership is high among their staff, allowing them to run Kobo Toolbox and Formera apps) to capture data from harder-to-reach communities. By design, these platforms allow offline data collection, with completed surveys transferred to a server once the devices are connected to the internet. This allows Taalim's staff to resolve the challenge of finding reliable internet connection inside Syria.

Keeping data and staff safe

Keeping assessment data in a mobile device comes with important data privacy and security risks in many areas in Syria. Particularly in the northwest, actors engaged in armed conflict might use survey data to map the location of learning centers and enumerators and teaching staff's activities. In fact, given the pervasive use of Kobo Toolbox among humanitarian organizations operating in Syria, members of armed groups have become familiar with the mobile app and its ability to collect GPS coordinates. To mitigate these risks, Taalim has trained their staff to raise awareness on the importance of data privacy, increase transparency about how assessment data is used, address concerns that staff might be surveilled while engaging in data collection, and develop protocols to deal with encounters with armed actors.

Dashboards as monitoring tools

Data collected by Taalim's staff inside Syria is analyzed by a team of MEL specialists and case management officers operating out of Gaziantep, Turkey, a hub for cross-border humanitarian assistance. The MEL staff has built interactive dashboards in Microsoft's Power BI software to track key output-level metrics, such as the number of beneficiaries reached, progress made on specific project indicators, and identify patterns in the data that could point to irregularities in service provision or location-specific conditions. Power BI can extract data directly from the servers used by data collection apps, ensuring data displayed on the dashboards remain up to date.

Case management software: open source vs. proprietary

Taalim's case management team relies on Salesforce—a cloud-based system—to centralize data streams from projects and activities. Through Salesforce, caseworkers can get a snapshot of children's assessments, identify needs, and track their interactions with Taalim's services. The platform also allows them to leave qualitative assessments such as comments and flags, key to the case management flow. Taalim actively uses

Salesforce's row-level security features, restricting access to personally identifiable information about beneficiaries only to staff with the appropriate permissions.

Salesforce is a proprietary software, requiring yearly licenses per authorized user. While open-source options were considered when transitioning out of paper-based case management, Taalim had reservations about whether open-source solutions could provide the level of data storage security and ownership to handle sensitive information, leading them to opt for a proprietary solution.

Is more data always better?

As many other organizations working in the EiE space, Taalim is currently thinking how technology could enhance their MEL capabilities by improving the tracking of beneficiary children on the move. Displaced children often lack personal identification which can allow them to be tracked across coverage areas. In many cases, Taalim must complete entire new assessments for children, even if they have the suspicion that the child accessed services before. In this sense, Taalim faces a tradeoff between increasing the amount of data they collect from beneficiaries, and their concerns that the same data could put them in danger.

APPENDIX 3. EXAMPLE OF TECHNOLOGY IN USE

Table 6: Pricing Guide for Data Tools

Process	Technology	Tool	Pricing	Pricing Link
Data Collection	Interactive voice response	Viamo	Not listed; contact for pricing	N/A
	Interactive voice response	Twilio	Varies by country and functions needed; starts at $0.0085/min to receive and $0.013/min to make a call	Link
	Interactive voice response	Magpi	Three tiers (basic, pro, and enterprise) for $0, $500, and $1500 monthly.	Link
	Interactive voice response	GeoPoll	GeoPoll On Demand pricing starts at just $3.00 per completed survey: cost varies based on the number of questions you're asking and the sample size.	Link
	Interactive voice response	RapidPro	The cost of hosting and using RapidPro will vary depending on how it is set up. For UNICEF, RapidPro is hosted centrally in the cloud, with each country office using RapidPro paying for set-up, volume-based hosting, SMS rates, aggregator services, shortcodes, customisations, and dashboards.	Link
	Offline mobile surveys	Formera	Four tiers (data collection, data management, data collaboration, and enterprise) at $25, $145, $195, and $900 monthly.	Link
	Offline mobile surveys	Survey123 for ArcGIS	Survey123 is included in the Field Worker user type at $350 annually, or it can be added for $60 annually to the Editor user type ($200 annually).	Link
	Offline mobile surveys	DeviceMagic	$25 per device per month; quotes available for high-volume contracts.	Link
	Offline mobile surveys	CommCare	Standard, Pro, and Advanced plans at $250, $500, and $1000 annually depending on needs; custom Enterprise pricing for large-scale international data collection also available.	Link
	Offline mobile surveys	Ona	Free, Standard, and Pro plans at $0, $99, and $199 monthly provide different tiers of services; custom pricing for Enterprise plans also available.	Link

Data Collection	Offline mobile surveys	Kobo	Free; unlimited storage for humanitarian organizations and up to 10000 submissions and 5gb storage monthly for others	Link
	Offline mobile surveys	ODK Aggregate	Free	Link
	Offline mobile surveys	Magpi	Basic, Pro, and Enterprise tiers at $0, $500, and $1500 monthly	Link
	Offline mobile surveys	iFormBuilder	$12.50 per user per month or custom pricing for organizations	Link
	Offline mobile surveys	ActivityInfo	Level 1, Level 2, and Level 3 tiers at 4800 euros, 9000 euros, and 15000 euros annually	Link
	Offline mobile surveys	Tangerine	Free, Member, Premium, Pro and Custom tiers at $0, $3500, $5000, $6000, and custom pricing annually. Free version allows 2000 entries.	Link
	Offline mobile surveys	REDCap	Free, but only available to nonprofit organizations that apply to join consortium.	Link
	Offline mobile surveys	Mobenzi	Community, Standard, and Essential tiers at $0, $195, and $395 monthly. Local pricing for South African organizations.	Link
	Optical mark recognition	FormScanner	Free	N/A
	Optical mark recognition	Moodle	Moodle software is free and open source	N/A
	Optical mark recognition	queXF	$0.55/form or contact for pricing at scale	N/A
	SMS	Commcare	Standard, Pro, and Advanced plans at $250, $500, and $1000 annually depending on needs; custom Enterprise pricing for large-scale international data collection also available.	Link
	SMS	Echo Mobile	Not listed	N/A

Data Collection	SMS	FrontlineSMS	Pricing depends on client budget and needs; website states willingness to work with organizations with small budgets to find affordable solution.	N/A
	SMS	Magpi	Basic, Pro, and Enterprise tiers at $0, $500, and $1500 monthly	Link
	SMS	Telerivet	Starter, Pro, Custom, and Enterprise tiers at $30, $120, $480, and custom pricing monthly	Link
	SMS	TERA	Not applicable; system needs to be set up with mobile operators in a given country	Link
	SMS	Textit	$25/month for standard contract; contact for pricing at scale	Link
	SMS	Ajua	Not listed	N/A
	SMS	RapidSMS	Free	Link
	SMS	RapidPro	The cost of hosting and using RapidPro will vary depending on how it is set up. For UNICEF, RapidPro is hosted centrally in the cloud, with each country office using RapidPro paying for set-up, volume-based hosting, SMS rates, aggregator services, shortcodes, customizations, and dashboards.	Link
	Social Media	RapidPro	The cost of hosting and using RapidPro will vary depending on how it is set up. For UNICEF, RapidPro is hosted centrally in the cloud, with each country office using RapidPro paying for set-up, volume-based hosting, SMS rates, aggregator services, shortcodes, customisations, and dashboards.	Link
	Social Media	WhatsApp	Free	N/A
	Social Media	Twitter	Free	N/A
	Social Media	Facebook	Free	N/A
	Social Media	Telegram	Free	N/A
	Computer-assisted qualitative data analysis software	Dedoose	Large group (6+ users) costs 10.95 per user per month	Link

Data Processing	Computer-assisted qualitative data analysis software	NVivo	Base price for 1 non-academic organization license and copy $1249.00. Quotes available for additional copies and add-ons.	Link
	Data analysis software	Stata	Pricing varies by country; organizations can request a quote.	Link
	Data analysis software	SPSS	Standard subscription $99.00/user per month; organizations can request quotes for custom pricing based on needs.	Link
	Data analysis software	SAS	Quotes available on an individual basis	Link
	Data analysis software	SPSS	Standard subscription $99.00/user per month; organizations can request quotes for custom pricing based on needs.	Link
	General purpose programming languages	R	Free	N/A
	General purpose programming languages	Python	Free	N/A
	General purpose programming languages	JavaScript	Free	N/A
	General purpose programming languages	Julia	Free	N/A

Data Processing	Machine learning techniques	TensorFlow	Free	N/A
	Machine learning techniques	Apache MXNet	Free	N/A
	Machine learning techniques	PyTorch	Free	N/A
	Machine learning techniques	Keras	Free	N/A
	Machine learning techniques	Pandas	Free	N/A
	Spreadsheets	Google Sheets	Free with individual google accounts; included as part of Google Workspace for organizations, which has plans at $6, $12, and $18 per user per month as well as custom pricing for larger organizations.	Link
	Spreadsheets	Excel	Included as part of Microsoft 365 subscription, with business plans at $5.00, $8.25, $12.50, and $20.00 per user per month	Link
Data Communication	Dynamic data visualization	Tableau	For teams/organizations, $70/user/month billed annually for a Creator license, plus $35/user/month and $15/user/month for additional Explorer and Viewer licenses	Link
	Dynamic data visualization	Google Data Studio	Free	N/A
	Dynamic data visualization	shinydashboard	Free	N/A
	Dynamic data visualization	Salesforce	Varies widely depending on scale and functions desired; limited number of dynamic dashboards available with Enterprise, Performance, Unlimited, and Developer versions of Salesforce Classic	Link
	Dynamic data visualization	Microsoft PowerBI	$9.99/user/month	Link

	Static visualization	ArcGIS	Reduced-cost, 1-year term licenses available for nonprofit organizations at various different tiers; pricing will vary based on scale and functions needed.	Link
	Static visualization	QGIS	Free	N/A
	Static visualization	Adobe Illustrator	Included as part of Creative Cloud for teams and businesses at $79.99/month or for schools/universities at $34.99/month	Link

REFERENCES

[1] B. K. Daniel, «Big Data and data science: A critical review of issues for educational research,» *British Journal of Educational Technology,* pp. 101-113, 2019.

[2] D. Reinsel, J. Gatnz y J. Rydning, «The Digitization of the World From Edge to Core,» IDC, Framingham, MA, 2018.

[3] B. Barry y L. Newby, «Use of Technology in Emergency and Post-Crisis Situations,» Global Education Cluster Working Group and IIEP-UNESCO, 2012.

[4] M. de France y A. Matthey-Doret, «Lessons learned paper from five years of Mobile Data Collection at Terre des hommes,» Terre des hommes, 2019.

[5] M. W. Unwin, M. Brugha y D. Hollow, «The Future of Learning and Technology in Deprived Contexts,» Save the Children, 2017.

[6] N. Dahya, «Education in Conflict and Crisis: How Can Technology Make a Difference? A Landscape Review,» Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, Bonn, 2016.

[7] Humanitarian Practice Network, «Humanitarian innovation,» Humanitarian Practice Network at ODI, London, 2016.

[8] B. Meiches, «Non-human humanitarians,» *Review of International Studies,* pp. 1-19, 2019.

[9] V. P. Korff, N. Balbo, M. Mills, L. Heyse y R. Wittek, «The impact of humanitarian context conditions and individual characteristics on aid worker retention,» *Disasters,* n° 39(3), pp. 522-545., 2015.

[10] D. Loquercio, M. Hammersley y B. and Emmens, «Understanding and addressing staff turnover in humanitarian agencies,» Overseas Development Institute.

[11] K. Anderson, L. Read y E. Losada, «Academic Learning Measurement and Assessment Tools in Education in Emergencies: Identifying, Analyzing, and Mapping Tools to Global Guidance Documents,» *ECC Network,* 2020.

[12] P. Montjourides y J. Liu, «Data and Evidence on Education in Emergencies: Linking Global Concerns with Local Issues,» *NORRAG Special Issue 2,* 2019.

[13] P. Kim, «Action research approach on mobile learning design for the underserved,» *Educational Technology Research and Development,* vol. 57, n° 3, pp. 415-435, 2009.

[14] UNHCR, «Data protection is part and parcel of refugee protection,» 23 May 2018. [En línea]. Available: https://www.unhcr.org/blogs/data-protection-part-parcel-refugee-protection/.

[15] T. Lüge, «Benchmarking SMS Tools,» UNHCR Public Health Section, Geneva, 2015.

[16] M. Tauson y L. Stannard, «EdTech for Learning in Emergencies and Displaced Settings: A Rigorous Review & Narrative Synthesis,» Save the Children, 2018.

[17] World Bank, «Somalia Education Programmatic Technical Assistance- Somalia: Status of Education Management Information System,» World Bank, Washington, D.C., 2018.

[18] European Parliamentary Research Service, «Technological innovation for humanitarian aid and assistance,» STOA, Brussels, 2019.

[19] USAID, «Delivering Distance Learning in Emergencies: A Review of Evidence and Best Practice,» USAID, 2020.

[20] GPE Secretariat, «How did Sierra Leone implement radio instruction during the ebola crisis?,» 29 April 2020. [En línea]. Available: https://www.globalpartnership.org/blog/how-did-sierra-leone-implement-radio-instruction-during-ebola-crisis.

[21] K. Muralidharan, A. Singh y A. Ganimian, «Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India,» American Economic Review, 2019.

[22] M. Trucano, «Using mobile phones in data collection: Some questions to consider,» 25 April 2014. [En línea]. Available: https://blogs.worldbank.org/edutech/using-mobile-phones-data-collection-some-questions-consider.

[23] M. Trucano, «Tablets in education,» 28 July 2015. [En línea]. Available: https://blogs.worldbank.org/edutech/tablets-education.

[24] K. Namit y T. T. Mai, «Digital School Census in 10 Weeks? How was it done in Sierra Leone,» 06 February 2019. [En línea]. Available: https://blogs.worldbank.org/education/digital-school-census-10-weeks-how-it-was-done-sierra-leone.

[25] Nayomi, «How might we improve educational outcomes for children and youth - particularly girls - in emergency situations?,» 5 August 2017. [En línea]. Available: https://challenges.openideo.com/challenge/education-emergencies/ideas/textit-for-education-in-emergency-in-vietnam-textit-for-eie.

[26] M. West y H. A. Chew, Reading in the Mobile Era, Paris: UNESCO, 2014.

[27] K. Bruce, V. J. Gandhi y J. Vandelanotte, «Emerging Technologies and Approaches in Monitoring, Evaluation, Research, and Learning for International Development Programs,» MERL Tech, 2020.

[28] UNHCR Innovation, «Using biometrics to bring assistance to refugees in Jordan,» 30 August 2016. [En línea]. Available: https://www.unhcr.org/innovation/using-biometrics-bring-assistance-refugees-jordan/.

[29] V. Corlazzoli, «ICTs for Monitoring & Evaluation of Peacebuilding Programs,» Search for Common Ground, 2014.

[30] R. Dette, J. Steets y E. Sagmeister, «Technologies for monitoring in insecure environments,» Secure Access in Volatile Environments, 2016.

[31] Oxfam, «Biometrics in the Humanitarian Sector,» 2018. [En línea]. Available: https://www.theengineroom.org/wp-content/uploads/2018/03/Engine-Room-Oxfam-Biometrics-Review.pdf.

[32] N. Joshi, «How AI Can and Will Predict Disasters,» 15 March 2019. [En línea]. Available: https://www.forbes.com/sites/cognitiveworld/2019/03/15/how-ai-can-and-will-predict-disasters/?sh=689a954e5be2.

[33] G. Akçapınar, A. Altun y P. Aşkar, «Using learning analytics to develop early-warning system for at-risk students,» International Journal of Educational Technology in Higher Education, 2019.

[34] Thinking Machines, «Mapping New Informal Settlements for Humanitarian Aid through Machine Learning,» 23 July 2020. [En línea]. Available: https://stories.thinkingmachin.es/mapping-new-informal-settlements/.

[35] USAID Middle East Education Research, Training, and Support Program (MEERS), «MEERS Data Mapping and Stakeholder Consultation Report,» 2018.

[36] A. Smiley, Y. Cao, W. Moussa, B. Dooley y J. Sullivan, «Examining "best practices" for literacy coaching and monitoring: Evidence from Northern Nigeria and Ghana,» *Social Sciences & Humanities Open,* 2020.

[37] OCHA, «Humanitarian Data Exchange,» 2018. [En línea]. Available: https://data.humdata.org/dataset/2048a947-5714-4220-905b-e662cbcd14c8/resource/c7053042-fd68-44c7-ae24-a57890a48235/download/ocha-dr-guidelines-working-draft-032019.pdf.

[38] DSEG, «A Framework for the Ethical Use of Advanced Data Science Methods in the Humanitarian Sector,» April 2020. [En línea]. Available:

https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/dseg_ethical_framework_april_2020.pdf.

[39] B. Ramalingam, «Innovations in the Nepal earthquake response: ten lessons from the DEC Response Review,» Humanitarian Practice Network, 2016.

[40] Busara Center for Behavioral Economics, «Getting the most out of your SMS Survey: Experimental Results,» Busara Center Lab, Nairobi, Kenya, 2018.

[41] J. A. Schwabish, «An Economist's Guide to Visualizing Data,» *Journal of Economic Perspectives,* pp. 209-234, 2014.

[42] E. Tufte, Envisioning Information, Graphics Press, 1990.

[43] B. Shneiderman, «The eyes have it: a task by data type taxonomy for information visualizations,» *IEEE Computer Society,* vol. 96, 1996.

[44] L. Ullrich y H. Khoudary, «WhatsApp Surveying Guide: Lessons Learnt from Two Qualitative WhatsApp Surveys in Lebanon,» UNDP, 2018.

[45] D. Waller, «10 Steps to Creating a Data-Driven Culture,» 6 February 2020. [En línea]. Available: https://hbr.org/2020/02/10-steps-to-creating-a-data-driven-culture.

[46] V. Venkatesh y H. Bala, «Technology Acceptance Model 3 and a Research Agenda on Interventions,» *Decision Sciences,* pp. 273-315, 2008.

[47] CartONG, «Benchmarking of Mobile Data Collection Solutions,» CartONG, 2017.

[48] L. G. Morales y T. Orrell, «Data Interoperability: A Practitioner's Guide to Joining Up Data in the Development Sector,» Global Partnership for Effective Development Co-operation, Dubai, 2018.