

## LEARNING BRIEF

*“Learning from what we’ve done and how we did it”*

Maximizing opportunities in data management of a large-scale development project through advances in digital technology

**Date:** May 2017

**Author/s:**

**Lynn van der Elst**  
(CEO, MIET Africa)

**Megan Cockerill** (Girls’  
Education specialist,  
MIET Africa)

**Contact:**

[megan@miet.co.za](mailto:megan@miet.co.za)

**Abstract**

Large development programmes that are heavily data-dependent pose considerable challenges, both logistical and analytical. This paper seeks to highlight the challenges associated with traditional systems (i.e. paper-based) of data collection and analysis, and demonstrates how the utilization of digitized technologies (as demonstrated in the **Young Women and Girls Programme**) can enhance efficiencies in logistics, analysis and overall project management.

MIET Africa’s learning briefs are developed to **share** new ideas, good practices and learnings arising from our implementation experiences, so as to **grow knowledge** with the **ultimate aim** of *improving the lives of children and youth*.

### INTRODUCTION

Large development programmes such as the **Young Women and Girls (YW&G) Programme** are invariably data-heavy because they target large numbers of individuals and therefore create a great many data points. The volume of data, the rate at which it is generated, and the need to interrogate it all pose considerable challenges, both logistical and analytical.

In recent years, the advent of modern digital technology has impacted virtually every aspect of daily life, but particularly for the management of information. It has created opportunities in information management and, in many cases, has revolutionized both data collection and data analysis. Certainly, it has contributed to a proliferation of data. But paradoxically, while the analytical improvements brought about by digitization are apparent, the expanded scale and volume of data poses new challenges associated with logistics of project implementation. For example, the exponential growth of data is accompanied by the need for a (re) trained staff and special infrastructure, and brings to the fore issues of privacy, intellectual

property rights and security. There is therefore an added need to include people and to develop policies and processes to ensure that these challenges are addressed. Developing countries, in particular, have characteristics that further exacerbate the challenges (Luna, DR 2014).

This learning brief describes the data management opportunities and challenges brought about by the implementation of a large-scale girls' education project in South Africa, and how the subsequent shift from a paper-based and manual system to an electronic tracking biometric system has provided solutions to challenges associated with both the management and the interrogation of large-scale data.

## **THE SETTING**

South Africa is the recipient of a three-year Global Fund grant (2016-2019) aimed at bolstering the country's national response to HIV&AIDS and TB, thereby adding value to the substantial commitments from the South African government and other funding partners. The grant supports several programmes, one of which is the YW&G Programme.

As a sub-recipient of the Global Fund grant under NACOSA<sup>1,2\*</sup> the principal recipient, MIET Africa, an education development NGO, is responsible for the implementation of two key components of the YW&G Programme: Keeping Girls in School (KGS) and Sexual Reproductive Health (SRH). This is the second phase of the programme, with the previous grant period running from 2013–16.

## **PROGRAMME OVERVIEW**

The YW&G Programme is an initiative focused on YW&G aged 10–24. It aims to reduce new HIV infections, reduce pregnancy, and retain girls in school until they complete Grade 12. The two relevant components of the YW&G Programme are described below.

### **KGS Component**

The KGS model, an initiative of the national Department of Basic Education, is a school-based intervention that identifies and supports female learners who are at risk of dropping out of school due to a range of interlinking factors—including pregnancy, poor academic performance, significant responsibilities within the home, and other health and social factors. KGS is implemented in a target study area comprising 50 Quintile<sup>2</sup> 1–32 secondary schools in each of four priority districts (Zululand, Greater Sekhukuhne, Ehlanzeni and Gert Sibande). It targets 100 girls in each school, thereby reaching approximately 20 000 of the most vulnerable girls per year.

A combination package of age-appropriate services is offered to adolescent girls in Grades 8–11, and includes: a peer education programme; Sexual Reproductive Health and Rights education; homework tutoring; career guidance; SRH referral and tracking follow-up on non-school-attending YW&G. All participating YW&G receive the health education component in

---

<sup>1</sup> The Networking HIV/AIDS Community of South Africa

<sup>2</sup> In South Africa, public (i.e. government) schools are graded into five quintiles, with the poorest schools in the lowest quintile, and least poor in the highest

each year of the grant (the sessions include new, more in-depth information each year) and participate in a career jamboree<sup>3\*</sup> once in the three-year grant period. The peer education, homework assistance and SRH referral components are provided according to the individual young woman's needs.

### SRH Component

SRH services are offered by nurses through mobile units that regularly visit schools and the surrounding community. Two mobile units have been allocated to a district, each staffed by a team of two nurses. The mobile SRH services include HIV testing services (HTS) and CD4 tests for girls who test HIV-positive. The mobile units also offer pregnancy testing, diagnosis and treatment for sexually transmitted infection (STI) and provide contraception.

HTS is provided to individuals beyond those reached by the defined package of care, and is also expanded to include this same sector of the population who are not in the programme but are living in the community and surrounds. A smaller number (limited to approximately 15% of the total tested) of young men and boys (YM&B) aged 10–24 are included to receive HIV tests.

## PROGRAMME INDICATORS AND MEASUREMENTS

The two programme components each have one core and several non-core indicators. The target levels for each have been set high, thereby creating several data management challenges.

### CORE INDICATORS

#### Component 1: KGS core indicator

*Cumulative proportion of young people in the target study area aged 10–24 reached by life skills-based HIV education in and out of schools over a three-year period*

#### Target

- Year 1: April 2017–March 2018, 25 000 young people
- Year 2: April 2018–March 2019, 20 000 young people
- Year 3: April 2019–March 2020, 15 000 young people

A young woman or girl is counted as having reached the target when once she has attended a minimum of three Health Education topics. The assumption is that most of these YW&G will also receive the other programme components, although this is based on their individual needs. Health Education topics comprise:

- Year 1: Puberty, Teenage pregnancy and HIV counselling and testing
- Year 2: Contraception, STIs and HIV counselling and testing

---

<sup>3</sup> A career guidance event at which exhibitors from various sectors (including government, business and tertiary education) are invited to make presentations on career options and to give advice on subject selections pertinent to various careers

- Year 3: Teenage pregnancy, STIs and HIV counselling and testing

The indicator is cumulative annually, meaning that beneficiaries reached are summed over the year to calculate the annual total. The count starts again from zero at the start of the next grant year. The count is cumulative since girls are gradually reached over the year with Health Education, as it is not possible to reach all the girls at the same time from Quarter 1. The aim is that the same cohort of YW&G is reached annually. However, an assumption has been made that approximately 25% of beneficiaries will leave and be replaced in the programme over the grant period.

#### Specific challenges

- Keeping accurate track of individuals and their level of service access over three years
- Generating individual and group profiles of YW&G

### Component 2: SRH core indicator

*Annual and cumulative proportion of young people aged 10–24 who have received an HIV test during the reporting period and know their results*

#### Target

- 33 000 young people per annum

An HIV test is counted each time a person in the YW&G Programme, or a person aged 10–24 not in the programme but living in the community and requiring access to HTS, receives an HIV test and knows their result.

The programme has included mechanisms to conduct HTS within the programme and therefore most individuals are tested directly by the MIET Africa team. Where an HTS referral to another service provider is necessary, only successful referrals (defined as evidence of the beneficiary receiving the service) are counted. The number of HIV tests conducted cumulate over the grant annual reporting period. This means that HIV tests are summed over the grant year and the indicator is re-set to zero at the start of the grant each year.

#### Specific challenges

- Keeping accurate records over three years
- Generating individual and group profiles of YW&G
- Generating profiles and keeping records of participants outside of the KGS component of the programme (i.e. boys and individuals out-of-school)

### NON-CORE INDICATORS

There are nineteen non-core indicators—five for the KGS component and fourteen for the SRH Services component—as set out in Table 1.

**TABLE 1: DESCRIPTION OF NON-CORE INDICATORS (A: KGS; B: SRH SERVICES)**

<b>A: KGS INDICATORS</b>	<b>DESCRIPTION</b>
Health Education	As for the Core Indicator
Peer Education	Number of girls in Grades 8–11 reached through peer education (facilitated by peer educators) on sexual behaviour
Career Jamborees	Number of girls in Grade 9 provided with opportunity to attend career guidance event
Homework Support	Number of girls in Grades 8–11 identified as requiring additional academic assistance who are then reached with homework support by peer group trainers
Home Visits	Number of girls who have either dropped out of school or have been absent for a prolonged period and who are then reached with home visits

<b>B: SRH SERVICES INDICATORS</b>	<b>DESCRIPTION</b>
# YW&G who tested HIV positive in the reporting month	A disaggregation of the HTS core indicator and a count of all the HTS with a positive result in the reporting month
# YW&G who tested HIV negative in the reporting month	A disaggregation of the HTS core indicator and a count of all the HTS with a positive result in the reporting month
# YW&G beneficiaries who disclose they are living with HIV in the reporting month	Number of YW&G (aged 10–24) and YM&B (aged 10–24) in a calendar reporting month who disclose to an SRH staff member that they are living with HIV
# YW&G & YM&B beneficiaries who decline an HIV test/do not disclose results	Number of YW&G (aged 10–24) and YM&B (aged 10–24) in a calendar reporting month who decline an HIV test or do not want to disclose their results to an SRH staff member
# newly diagnosed HIV+ YW&G & YM&B who are successfully referred for pre-ART/ART initiation	Number of newly-diagnosed (diagnosed in the reporting month) HIV positive YW&G (aged 10–24) and YM&B (aged 10–24) who are successfully referred for pre-antiretroviral treatment (ART)/ART initiation by an SRH staff member <sup>(a)</sup>
# HIV+ YW&G and YM&B successfully referred for ART adherence support	Number of HIV positive YW&G (aged 10–24) and YM&B (aged 10–24) who are successfully referred for ART adherence support by an SRH staff member <sup>(a)</sup>
# YW&G & YM&B screened for STIs	Number of YW&G (aged 10–24) and YM&B (aged 10–24) screened for STIs by an SRH staff member in a calendar reporting month

<b>B: SRH SERVICES</b>	
<b>INDICATORS</b>	<b>DESCRIPTION</b>
# YW&G & YM&B screened for TB	Number of YW&G (aged 10–24) and YM&B (aged 10–24) screened for TB by an SRH staff member in a calendar reporting month <sup>(a)</sup>
# YW&G & YM&B with a positive STI symptom successfully referred for STI treatment	Number of YW&G (aged 10–24) and YM&B (aged 10–24) with a positive STI symptom in the reporting month, who are successfully referred for STI treatment by an SRH staff member <sup>(a)</sup>
# YW&G & YM&B with a positive TB symptom successfully referred for TB investigations	Number of YW&G (aged 10–24) and YM&B (aged 10–24) with a positive TB symptom in the reporting month, who are successfully referred for TB treatment by an SRH staff member
# YW&G screened for pregnancy	Number of YW&G (aged 10–24) screened for pregnancy by an SRH staff member in the reporting month
# YW&G with a positive pregnancy test successfully referred for antenatal services	Number of YW&G (aged 10–24) with a positive pregnancy test, who are successfully referred for antenatal services by an SRH staff member <sup>(b)</sup>
# male condoms distributed	Number of male condoms distributed to YW&G (aged 10–24) and YM&B (aged 10–24) by an SRH nurse <sup>(c)</sup>
# female condoms distributed	Number of female condoms distributed to YW&G (aged 10–24) and YM&B (aged 10–24) by an SRH nurse <sup>(c)</sup>
<p><i>(a) Successful referral is defined as evidence of the beneficiary receiving the service in a calendar reporting month.</i></p> <p><i>(b) Successful referral is defined as evidence of the beneficiary receiving the service.</i></p> <p><i>(c) Individual condoms are counted and not the number of packs. For example, if there are 10 condoms in a pack, then 10 condoms are reported.</i></p>	

#### Specific challenges

- Identifying and archiving individuals who access a service only once in the programme
- Generating individual and group profiles and keeping records of individuals outside of the KGS component of the programme (i.e. boys and out-of-school individuals)

## DATA MANAGEMENT CHALLENGES AND OPPORTUNITIES

As the YW&G Programme progressed from Phase 1 to Phase 2, the operations and management for the new phase were altered to harness the new technology. Although still in the early stages of Phase 2, improvements in data management are already being experienced, with more efficient and economical operations being a valuable by-product.

**CHALLENGES ASSOCIATED WITH A PAPER-BASED DATA MANAGEMENT SYSTEM**

Under the previous Global Fund grant (i.e. during Phase 1), paper-based registers were used to track activities per school, per cohort and per learner. Fieldworkers collected information at the schools, before taking the registers to the site office to be checked by the supervisor on a weekly basis. Once checked, the registers were couriered to MIET Africa's head office in Durban where data clerks captured the information into a database. The introduction of digital and biometric systems has eliminated the associated logistical challenges inherent in this method.

Examples of the main hurdles experienced in Phase 1 that were ameliorated or resolved in Phase 2, are:

- **Inefficient and risky document transportation**

Transporting documents in a paper-based system is slow and inefficient. Project staff had to travel large distances from schools to the site office, often resulting in late submission of registers, thereby delaying monthly reporting. All sites were based in outlying areas, which meant that collection of registers was only on specific days, and delivery to Head Office was three to four working days later.
- **Backup of information**

Unavoidably, paper documents get lost, damaged, stolen or somehow displaced, which could pose serious problems for the entire programme. "Cloud storage" of the entire database eliminates this problem.
- **Supply costs**

A major drawback of paper-based data management systems are the associated costs. In addition to transport costs, several data clerks had to be recruited at MIET Africa's head office to ensure that the capturing of registers was kept up-to-date. The use of tablets by the fieldworkers allows for immediate input of data at the site, thereby reducing the number of data clerks required.
- **Verification of services**

Management was reliant on staff at site level to track and monitor the number of beneficiaries receiving services and could only verify the number at the beginning of the following month, once all the data had been captured. This delayed the identification of problems. The biometric tracking system enables immediate verification.
- **Double counting**

Changes in the spelling of names or use of nicknames meant that there was the danger of capturing the same individual twice. Much time had to be spent on data cleaning to ensure that there was no double counting. Fingerprint-recognition technology eliminates this problem.
- **Environmental credentials**

In order to provide for back-up and data security, the registers were often photocopied, thereby generating a mass of additional paper and a negative

environmental imprint. Investing in an electronic data management system reduces the amount of paper needed.

- Physical storage space

Storage of paper takes up a significant amount of space, which increases as the project is scaled up.

- Time management

Supervisors are required to check registers for accuracy and completeness. This takes time, and if there are queries to resolve, even more time is required. With large numbers of registers and tight turnaround reporting periods, meeting deadlines proved to be challenging. The elimination of outliers and having input data “filtered” has significantly reduced verification times.

## **NEW SYSTEMS INTRODUCED**

Two important digital technology systems have been introduced in the current phase of the programme.

1. *Cloud storage of the relational database*

An electronic database that is stored in the cloud is accessible to multiple users simultaneously.

2. *Biometric monitoring system*

Based on its success in two of MIET Africa’s large HIV prevention projects, a biometric monitoring system was introduced in order to strengthen the management of information and support the M&E reporting requirements of the current phase of the YW&G Programme.

The system, provided by the Broccoli Project,\* uses cloud storage and robust in-field data collection technology that accommodates all connectivity scenarios—for example, where connectivity is not available, the system will retain locally collected data until it detects a suitable Wi-Fi or GSM† link, at which point it auto-synchronizes with the central data repository. Furthermore, the biometric system provides a fully auditable view of every interaction and/or event. Every action, event or modification is linked to the administrator and (where relevant) the participant involved.

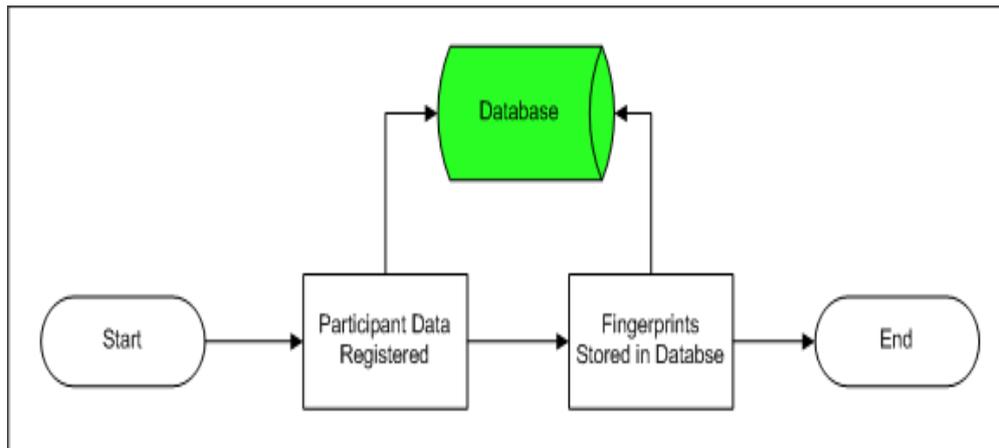
The selected biometric system involves two processes: an enrolment phase, where fingerprints are linked to a participant’s profile and saved in a database; and an identification process.

The registration process is illustrated in Figure 1, and primarily involves setting up the profiles of the individual learners/community members. The following information is captured in this basic profile: name and surname; date of birth; gender; province; district; sub-district; name of the school and EMIS number; grade on enrolment; vulnerability criteria on enrolment into programme; enrolment date; relevant capture details.

---

\* See [www.broccoliproject.org/](http://www.broccoliproject.org/)

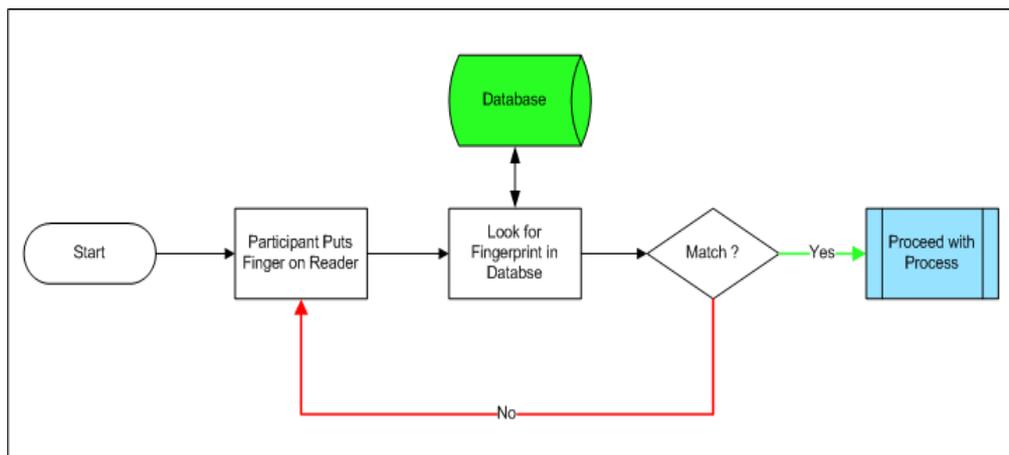
† Global System for Mobile Communications



**FIGURE 1: PARTICIPANT REGISTRATION PROCESS**

*On registration, participants' details are captured and linked to their fingerprints. This information is stored in the database.*

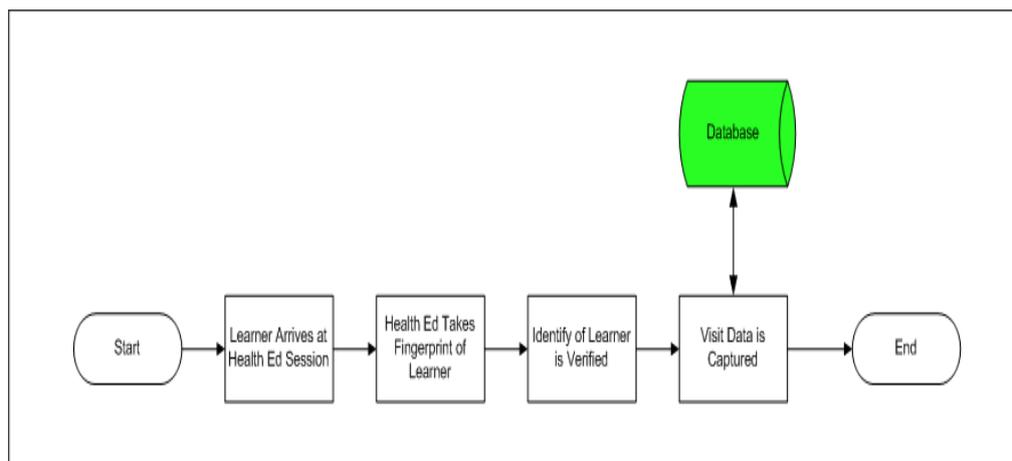
Once a profile has been created and fingerprints are linked to it, the relevant profile can be identified and located by simply using the fingerprint that was originally used in the enrolment process, as shown in Figure 2. Once a fingerprint has been matched, the process of capturing a service received by that participant can commence (see Figure 3).



**FIGURE 2: IDENTIFICATION OF A PARTICIPANT PRIOR TO THE CAPTURE OF SERVICE RECEIVED**

*Before a service provided to a participant is captured, the participant puts their finger on the reader, which then searches in the database for that individual. Once the match is located, the process can continue.*

Both the core indicators and most of non-core indicators are captured progressively, using the biometric system, as illustrated in Figure 3.



**FIGURE 3: CAPTURE OF PARTICIPANT WHO HAS RECEIVED A SPECIFIC SERVICE**

*A participant attends a Health Education session and is fingerprinted by the health educator. The system automatically verifies the identity of the participant (see Figure 2), and their participation is captured.*

### POSITIVE IMPLICATIONS

The improvements to the overall programme management can be classified into two categories—i.e. Analytical and General.

#### Analytical

- *Data validation at time of capture*

One serious difficulty with large data management programmes lies in the original data capture being flawed through erroneous entries. This is especially true for paper-based systems. However, a well-constructed database interrogates entries according to a set of rules before capture, thereby effectively eliminating most incorrect data capture entries. It is also possible to capture the data independently twice, so that merging the two data sets will identify errors.

- *Outliers identified*

The database generates sample extracts of the captured information where appropriate, some in graphic format. This allows for rapid identification of outliers, thereby either eliminating a specific questionable entry, or confirming its validity.

- *Analysis by variables*

Historically, the number of parameters that could be analysed from a particular data set were severely restricted and was dependent largely on the format in which the data was collected. Furthermore, it was very labour-intensive and time-consuming. The new relational database facilitates analyses by various parameters. For example, the profiles of young people can now be grouped into cohorts and compared between different locations, schools, etc.

- *Calculation of means, averages, standard deviations and variances*

Statistical analysis and interpretation of the data is a prerequisite for effective monitoring and evaluation. The database facilitates extraction of selectable data so

that it can be statistically processed or inserted into other analytical packages (such as Statistical Package for Social Sciences).

## General

The use of a biometric tracking system includes the following additional benefits:

- All the relevant participant details are stored and linked to a unique identity. This prevents co-enrolment across multiple sites.
- It allows tracking of individual user-participation across multiple sites.
- It is a stable and reliable means of identifying participants; participants are not required to provide any documentation (such as ID book or participant card) when they are being identified.
- All details regarding participant interactions with the programme can be recorded and easily accessed through web-based reporting.
- It allows for anonymity, as names, ID numbers, etc. are not mandatory for tracking.

## CONCLUSION

It is clear that Phase 2 of the YG&W programme is benefiting from the introduction of selected digital technology, through greatly improved logistics, analysis and overall management.

The various challenges associated with manual-based systems that have been identified in this brief can serve as a useful first step in considering an M&E approach to a large development programme. These challenges require careful consideration from the outset, and should be included within the planning stage of project implementation. As has been discussed, the use of digital technology can often reduce these challenges significantly. However, consideration of any data digitization process should be preceded by identifying clearly the logistic and analytical hurdles anticipated. In most cases, money can be saved and efficiency boosted by investing in the technology that is right for the project.

## ABBREVIATIONS

ART	Antiretroviral Treatment
HTS	HIV Testing Services
KGS	Keeping Girls in School
SRH	Sexual Reproductive Health
STI	Sexually Transmitted Infection
YM&B	Young Men and Boys
YW&G	Young Women and Girls

## REFERENCES

Luna DR, Mayan JC, García MJ, Almerares AA & Househ M (2014) Challenges and Potential Solutions for Big Data Implementations in Developing Countries. *IMIA Yearbook of Medical Informatics*: 36–41.