

The Science Behind Data Collection:

How to choose the best tools and approach to collect data considering the culture, context and existing partnerships.





Khulisa

Khulisa supports African people and institutions to better demonstrate results and use high quality data for decision making, to ensure accountability and build knowledge

Agenda

Time	Activity	Facilitator
14:00-14:30	Welcome and Ice breaker	Jesse Webb
14:30-14:45	<ul style="list-style-type: none"> • Introduction to the online seminar and introduction of the presenters. • The role of AI in M&E 	Margie Roper
14:45-15:00	Survey tools and approaches: JotForm and CATI survey	Jesse Webb
15:00-15:15	Learner assessments: Tangerine	Margie Roper
15:15-15:30	Mapping data: Heatmaps ArcGIS	Thembi Mahlangu
15:30-15:45	A diverse tool for various kinds of data collection: Kobo Toolbox	Tamar Boddé-Kekana
15:45-16:00	Conclusion and Q&A	Jesse Webb

Webinar Rules

- This webinar is being recorded. Your participation in the webinar acts as your consent for this recording to be posted on our website.
- While presentations are taking place, keep microphone muted and camera off to eliminate background sound and poor bandwidth connection.
- During the presentation, use the chat function and Q&A function to ask questions and make comments.
- During Q&A section, raise your hand to ask a question. When the webinar host calls on you, unmute your mic.
- Please refrain from vulgar and insensitive language. Act and speak with cordiality and kindness while engaging with the webinar presenters and attendees.
- For any follow up information after the event, please email jwebb@khulisa.com

Let's hear from you...

What expectations do you have for this session?

What question would you like addressed?

Write it in the chat.



Margie Roper

“

Margie is a technical expert with over 25 years of experience leading and executing monitoring and evaluation (M&E), social research, learning and development, and capacity building and training assignments.

She has expertise in early childhood development (ECD), early grade reading/learning, education, child development and protection, resilience, youth development, behaviour change, social justice and protection.

Margie leads the Education and Development Division at Khulisa. She is currently completing her PhD in Education Research, Evaluation and Enhancement



Data collection science?

- **Data collection** or **data gathering** is the process of gathering and measuring information on targeted variables in an established system, which then enables one to answer relevant questions and evaluate outcomes
- **Data science** is an interdisciplinary academic field that uses statistics, scientific computing, scientific methods, processes, algorithms and systems to extract or extrapolate knowledge and insights from noisy, structured, and unstructured data.
- Data science is multifaceted and can be described as a science, a research paradigm, a research method, a discipline, a workflow, and a profession.



WIKIPEDIA
The Free Encyclopedia



The big question

The rise of AI and the implications
on data collection and analysis?



Jesse Webb

“

Jesse is a Monitoring, Evaluation, Research and Learning (MERL) Associate in the Education and Development Division at Khulisa.

Jesse has experience in evaluating refugee programming, primary education, early childhood development (ECD), and trafficking in persons.

She has a Master's degree in International Relations and Diplomacy.



5 main considerations when creating data collection tools

BUDGET

RESPONDENTS AND
ACCESSIBILITY

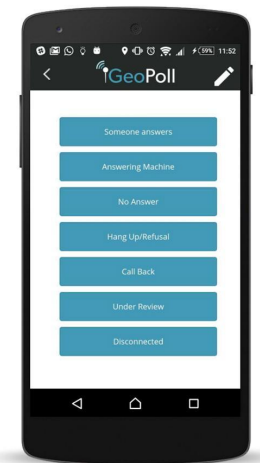
KIND OF DATA BEING
COLLECTED

HOW THE DATA WILL
BE USED

ANTICIPATED
CHALLENGES TO
DATA COLLECTION

Survey options for hard-to-reach places

JotForm and Computer Assisted Telephonic Interviews (CATI)



Accessibility



Computer Assisted Telephonic Interviews (CATI) are surveys conducted over the phone. Beneficiaries are called and asked questions by an interviewer, and responses captured into the computerized system being used.

- Budget: CATI surveys can be more affordable than going into the field
- Incentives: CATI surveys can easily give respondents data/airtime bundles as incentives
- Limitation: Mobile/telephone penetration of the area/people you are surveying

Accessibility

- JotForm is a survey tool similar to Survey Monkey that has offline capability and fillable-PDFs function.
 - Benefit: very useful for interviews requiring a lot of typing
 - Limitation: Not as many question types/skip logic features as some other survey tools

The screenshot displays the JotForm builder interface. On the left is a 'Form Elements' sidebar with categories: BASIC, PAYMENTS, and WIDGETS. The 'BASIC' category is selected, showing various form elements like Heading, Full Name, Email, Address, Phone, Date Picker, Appointment, Signature, Fill in the Blank, Product List, Short Text, Long Text, Paragraph, and Dropdown. The main workspace shows a survey form titled 'Section 1: Characteristics of the urban refugee youth population in Africa'. The form contains four questions:

1. How old are you? (If under 18, check that they have permission from their parent/guardian to conduct the interview)

(If the answer is not between 15-24 the interview ends there.)
2. Would you say you are...
 - Male
 - Female
 - Other
 - Refuse to answer
3. Which country are you originally from?
4. Who came with you to this country?
 - Alone (unaccompanied)
 - With parents
 - With relatives or people you knew
 - I don't know

At the bottom, there is a page navigation bar with 'PAGE 1', 'PAGE 2', 'PAGE 3', 'PAGE 4', 'THANK YOU PAGE', and '+ ADD NEW PAGE'.

Accessibility

- JotForm Addresses the anticipated challenge of limited internet connectivity, while still utilizing web-based forms
- Use: These tools are designed for survey questionnaires, with some qualitative responses. They do not include GIS technology or integrated analysis.



Assessing Learners to improve learner Outcomes

Tangerine TM

What is tangerine®



tangerine®

About Tangerine™

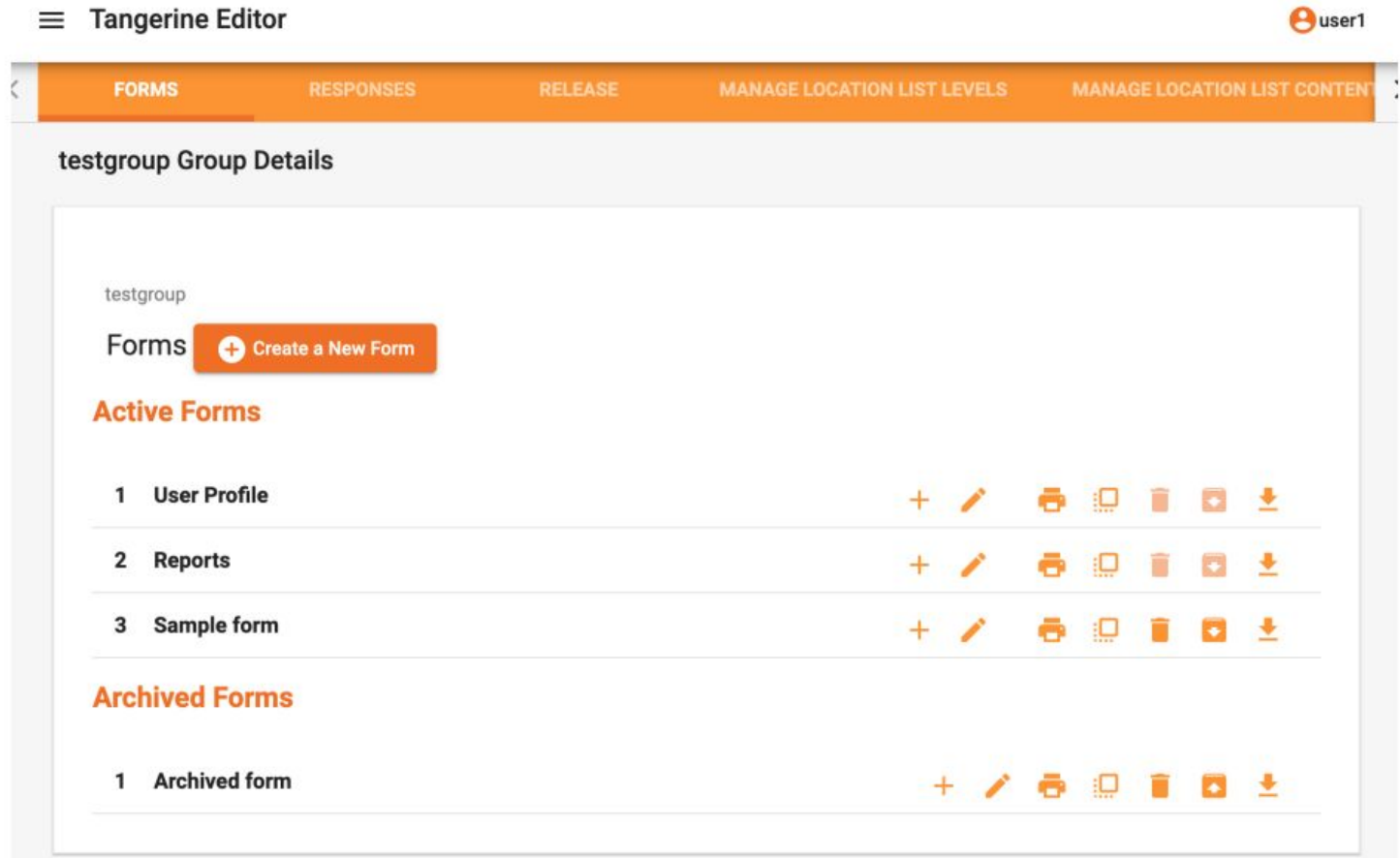
Tangerine is open source electronic data collection software designed for use on Android mobile devices. Its primary use is to enable recording of students' responses in oral early grade reading and mathematics skills assessments. Tangerine is also used to capture interview responses from students, teachers, and principals; as well as in simple surveys and other data collections.

Tangerine was developed by RTI International with funding from RTI and Google.org.

For more information on Tangerine, visit <http://www.tangerinecentral.org>.

Accessible assessment data collection

- Easy to design assessments & surveys
- Timed activities
- Quantitative & qualitative
- Easy administration on tablets
- Must upload / sync data
- Free version or paid
- Different access permissions
- Use by fieldworkers & teachers



Key requirements

- Must train fieldworkers
- Practice, practice, practice
- Set up for data analysis
 - Partnerships & collaboration important
- Context important
- Tablet savvy



Thembi Mahlangu

“

Thembi has 13 years experience in providing technical support to project activities and deliverables, project management, training and facilitation of workshops, financial management and stakeholder management.

Her experience includes mental health programming, social services and education. She is a certified social worker.

She is currently completing her Master's in Public Health and has an Honour's Degree in Developmental Studies.



Mapping of contextual data

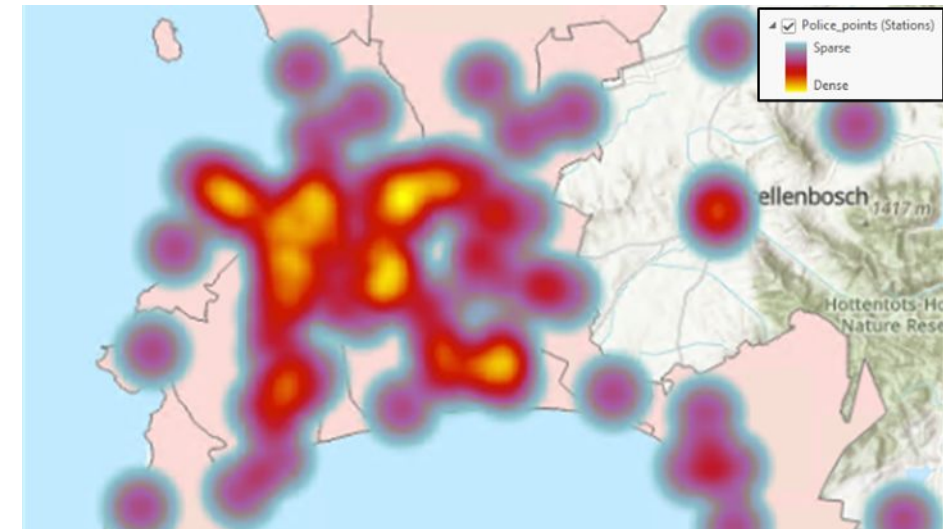
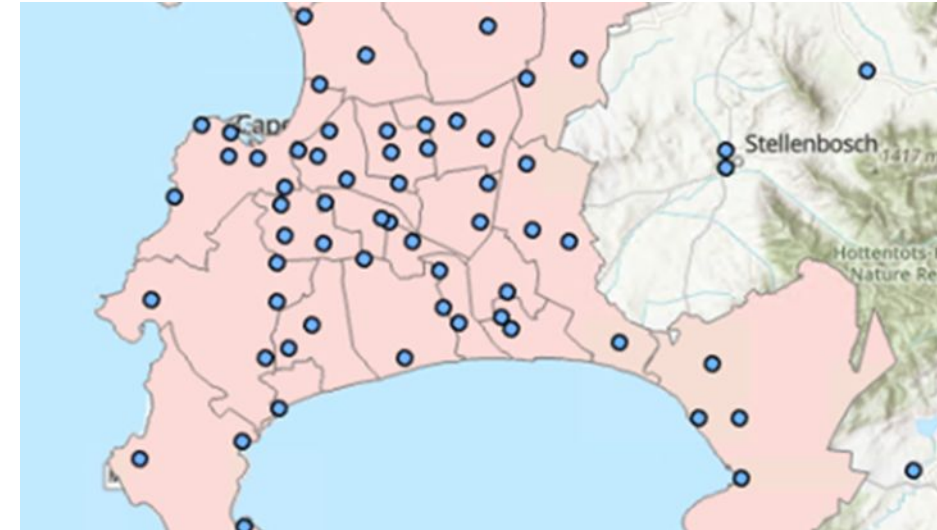
Heatmaps ArcGIS



ArcGIS

Heatmaps

- Heat mapping is a visualization tool to identify and highlight areas where the density of features or a parameter of interest is elevated relative to surrounding areas. The visualization can be based on discrete areas (such as zip code polygons) or different study area demarcations.
- In both cases, the level of the parameter of interest is symbolized with a color ramp where higher values appear darker and lower values lighter
- This type of visualization uses color ramps of gradually varying intensity ranging from pale blue (lower values) through red to yellow (higher values) thus resulting in the name “Heatmap”. For example, mapped Police stations in CPT.
- Heatmaps represents the density of features of the magnitude of a parameter of interest and is highest of the locations



Measures to Counter Trafficking in Persons [MCTIP]

- Create a map representing the relative risk (i.e., probability) of encountering someone who is engaging in or is a victim of human trafficking.
- Risk factors for vulnerability include economic insecurity, housing and environment insecurity, lack of education, discrimination, and social and cultural exclusion
- Geographical Coverage – COJ & CPT
- Map the locations of identified risk factors, each in their own spatial layer and assign a quantitative value to rank the risk from that factor.
- The Back – End managed by Statisticians with mathematical formulars
- The software being used for this is ArcGIS, with the mobile application Field Maps.

Heat map Domains

Economic indicators

Child development-related indicators

Migration flow, routes or activities

Qualitative hot spots

Use of the heatmaps

The resulting heatmaps are useful visual aids for prevention and advocacy services. Such maps can be used for resource allocation by local governments for anti-trafficking prevention activities and to help inform outreach campaigns. The heatmaps methodology will be documented potentially for sharing in the sector for future planning purposes, targeting and resource mobilization

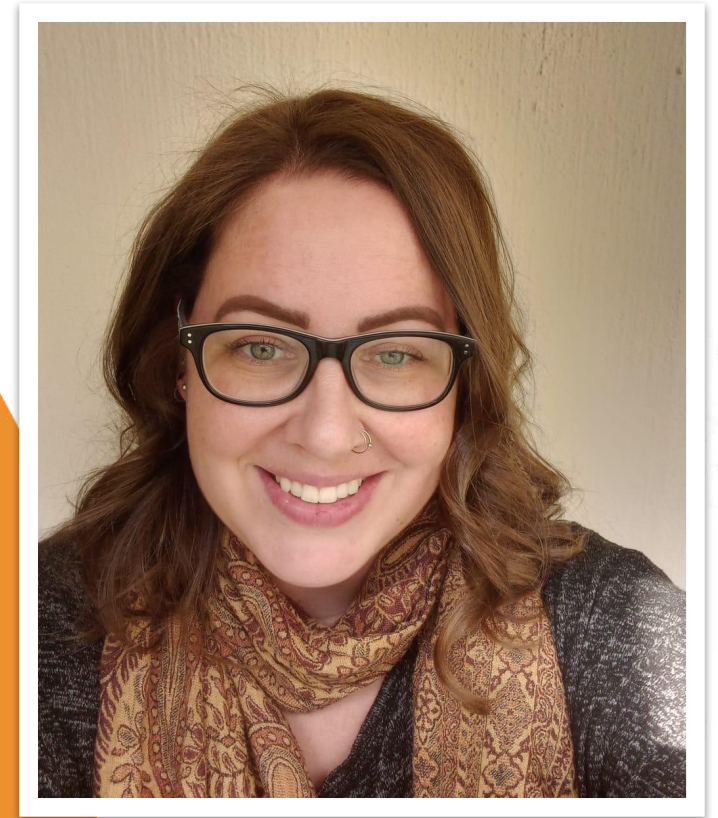


Tamar Boddé-Kekana

“

Tamar is is a Senior Monitoring, Evaluation, Research and Learning (MERL) Specialist at Khulisa Management Services with more than 10 years of experience in quantitative and qualitative research methodology, data analysis, monitoring and evaluation and quality assurance. Her sectoral experience includes education, career guidance and skills development, hybrid learning technology and development, child and youth development, peer-to-peer mentoring, community engagement, youth mobilization, social entrepreneurship, and cross-cultural learning and engagements.

Tamar has an MPhil in Sociology, with a focus on monitoring and evaluation in the education sector

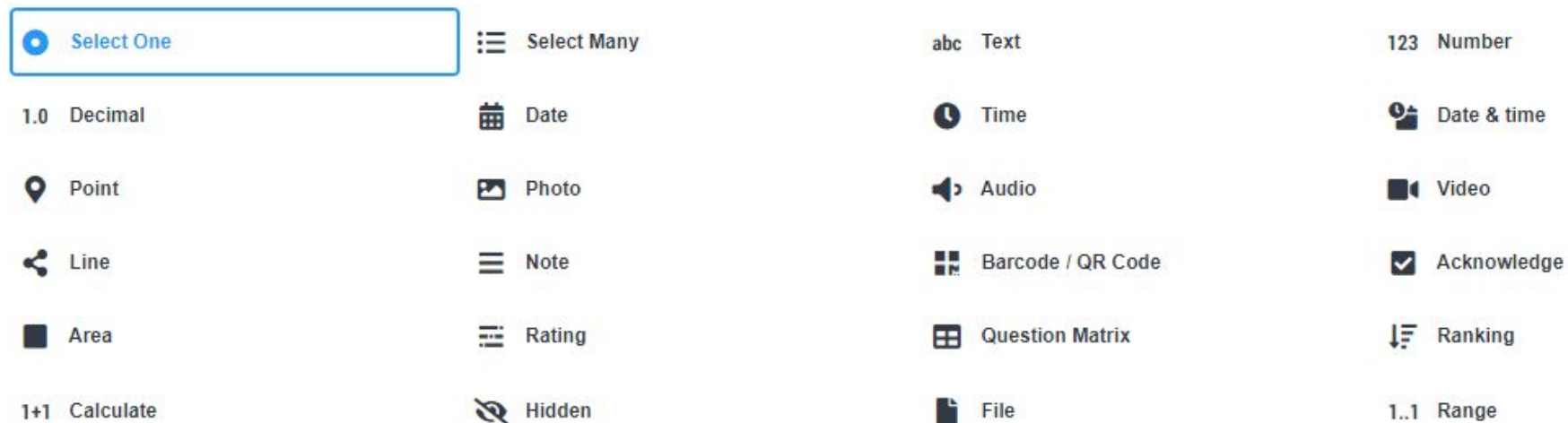


A collaborative tool for Data Collection and Data Quality checks

Kobo Toolbox

KoBoToolbox as an Inclusive Tool

- Open Source | Accessible to all | Translations in any language possible
- Fieldworkers | Remote data collection | Offline options | Data collection via app | Tablet/Phone/PC | Can be set up to do on-the-spot calculations and data checks
- Many types of questions/notes/forms of data to be collected:



KoBoToolbox Community



Any question you have, any code you struggle with, and any option that KoBoToolbox offers is discussed on this forum!

It is an amazing community of people using or looking to use KoBo Toolbox

Create your own KoBoToolbox account and play around with all the options:
<https://kf.kobotoolbox.org/>

Find the KoBoToolbox Community here: <https://community.kobotoolbox.org/>



Mentimeter:
Go to www.menti.com
Code: 7648 6783



Recent AI tools ...

Finding and synthesising literature:

- Semantic Scholar – access to 200 million research papers, links between topics, recommendations based on recent searchers
- Scholarcy – article summariser
- Paper Digest – 3 minute summaries of research papers by extracting key ideas and sentences
- Content Mine – extract knowledge from academic articles
- Elink.io – saves content from web articles, videos, cloud files, social media ...
- Elicit – find relevant papers
- Scite – where articles have been cited and finds supporting and contrasting evidence
- SciSpace Copilot – multi-lingual AI tool including comprehending maths/tables, easy summaries

Data analysis & writing:

- Excel Formula Bot
- AILYZE- interviews

- Academic writing (they won't write it for you!):
 - Trinka
 - Grammarly
 - Lex
 - Scrivener

Where to from here?

The Future of M&E: Culture, Context and Collaboration

- Reminder: 5 considerations for Data Collection design
 - Budget
 - Respondents and Accessibility
 - Kind of data being collected
 - How the data will be used
 - Anticipated challenges to data collection
- Culture
 - Resistance to/embracing technology
 - Fear/awe of AI
 - Diversity
 - Development in transition



Where to from here?

The Future of M&E: Culture, Context and Collaboration

- Context
 - Unequal access to technology within a society
 - Just transition
 - Confidentiality and data privacy
 - Progress & behaviour change
- Collaboration
 - Partnerships beyond technology
 - Utilization of data

