NANO PILOT

FINAL REPORT FOR THE GLOBAL INNOVATION FUND (GIF)

9 October 2019
Authors

Valentina Brailovskaya: Valentina.Brailovskaya@IDinsight.org
Ana Chaves: Ana.Chaves@IDinsight.org
Christopher Chibwana: Chris.Chibwana@IDinsight.org
Eric Dodge: Eric.Dodge@IDinsight.org
Martin Gould: Martin.Gould@IDinsight.org
Kondwani Mumba: Kondwani.Mumba@IDinsight.org
Nate Vernon: Nate.Vernon@IDinsight.org
Paul Wang: Paul.Wang@IDinsight.org

About IDinsight

IDinsight uses data and evidence to help leaders combat poverty worldwide. Our collaborations deploy a large analytical toolkit to help clients design better policies, rigorously test what works, and use evidence to implement effectively at scale. We place special emphasis on using the right tool for the right question, and tailor our rigorous methods to the real-world constraints of decision-makers.

IDinsight works with governments, foundations, NGOs, multilaterals and businesses across Africa and Asia. We work in all major sectors including health, education, agriculture, governance, digital ID, financial access, and sanitation.

We have offices in Bengaluru, Dakar, Johannesburg, Lusaka, Manila, Nairobi, New Delhi, San Francisco, and Washington, DC. Visit www.IDinsight.org and follow on Twitter @IDinsight to learn more.
Chief Mukobela with the Nano application on his tablet

Executive summary 4
1. Introduction 5
2. Design and execution 7
3. User actions based on Nano data 14
4. Evaluation of Nano pilot 23
5. Dissemination of lessons learned 30
6. Next steps for Nano 32
7. Appendix 34
EXECUTIVE SUMMARY

Information systems (IS) frequently fail to affordably produce timely, high-quality data that are linked to decisions. In partnership with the Global Innovation Fund (GIF), we have worked to create Nano, a locality-focused information system designed to be faster, cheaper, and flexible with data quality controls to provide relevant data to decision-makers in any sector as demanded.

OVERVIEW OF NANO PILOT

Nano was piloted in Mukobela chiefdom, Southern Province, Zambia. This setting was chosen because traditional leaders are key decision-makers in their communities and are underserved by other IS efforts. In the course of the pilot, IDinsight conducted household and school-level surveys to test system functionality across survey types. We additionally trialled a process for remote survey assignment, training enumerators on a new survey via a messaging platform and later assessing comprehension and execution of the task. We present here lessons learned on the viability of remote tasking as a method for reducing costs in a Nano model.

ACHIEVEMENTS

The Nano team demonstrated several technical and systemic achievements in the pilot. These include developing a series of processes for collection and presentation of information relevant to a localised key decision-maker, statistical tools to maintain data accuracy, and an Android application customised to the preferences, limitations, and capacities of Chief Mukobela.

The chief has expressed that he is “fully satisfied” with Nano and that he had expected a much slower data turnaround. We successfully met his stated information needs, and the chief has appropriately used data for community decisions in several cases. We presented two rounds of data to Chief Mukobela, recording how key indicators in each affected his beliefs on the state of his community. Chief Mukobela has since shared data with several district-level, national-level, and NGO stakeholders to advocate for improvement on critical indicators in his community. On at least one indicator, Chief Mukobela has already initiated direct action himself by coordinating construction of a health facility in the area our data identified as the most underserved in the chiefdom.

At the start, we defined success as creating a system 1) that people love, and 2) that produces high-quality data that informs impactful decisions and actions. Chief Mukobela has expressed strong support for Nano, and six other Chiefs and three district councils have expressed interest in using the system. Stakeholder feedback and action indicate that the Nano pilot was successful on both fronts.

ADVANCING NANO LEARNINGS AND IMPLEMENTATION

The Nano team will document learnings from the pilot in a ‘do-it-yourself’ manual. These learnings will be widely distributed with the intent of enabling others to replicate our approach. We are eager to continue to develop Nano innovations and work towards a closer approximation of a full-scale Nano model. We will work closely with GIF to identify the best path forward, whether that includes financial support from GIF or partnership with other organisations. The most likely path forward with continued engagement from GIF is a pivot towards data tools for frontline workers with an emphasis on building capability to collate, quality-check and augment data from pre-existing sources.
1. INTRODUCTION

OVERVIEW OF NANO MOTIVATION AND PRINCIPLES

Efforts to implement information systems (IS) to enhance state capacity are widespread. However, these efforts typically fail due to top-down, complex, slow, and expensive design and execution.¹ We envision Nano as a solution to these issues – an aggressively consultative, low-cost, and flexible information system that comprehensively adheres to predictors of IS success.

Nano’s theory of change (Figure 1) is that access to relevant, accurate, timely, and well-presented data that users demand will yield data-based actions to influence government, civil society, and population behaviour to improve lives.

Figure 1: Nano’s theory of change

We expect that users will love Nano because it quickly produces critical and decision-relevant information they demand in an intuitive and engaging format. Nano has five components (see Table 1), and we have developed each of these in this project.

Table 1: Nano components

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Consultation and stakeholder buy-in</td>
<td>Process by which we identify needs/demand and gain the support of all relevant stakeholders</td>
</tr>
<tr>
<td>2. Survey modules and data collection modes</td>
<td>Library of survey modules for streamlined data requests, tailored to the manner in which data can be collected (household interviews, service provider interviews, observations)</td>
</tr>
<tr>
<td>3. Sampling</td>
<td>Sampling strategy and sample size determination process</td>
</tr>
<tr>
<td>4. Enumerator management</td>
<td>Recruitment, training, and management (in-field and remote) including incentives/remuneration</td>
</tr>
<tr>
<td>5. Data presentation</td>
<td>User-friendly interface with data reports and guidance on how to act on data</td>
</tr>
</tbody>
</table>

DESIGN OF THE PILOT IN ZAMBIA

Nano was piloted in Mukobela chiefdom, Southern Province, Zambia, supported by the Global Innovation Fund.

Traditional leaders are important decision-makers, and policy makers in Zambia. As locally embedded and widely respected leaders, chiefs are uniquely able to mobilise collective action, bypassing common challenges faced by government counterparts. However, they lack access to basic information on their chiefdoms and so, are often unable to identify and fully address issues for their communities.

Information systems implemented for decision-makers often struggle to fill this gap due to their complexity, cost, unwieldy design, and lack of consultation and context specificity. Nano attempts to solve these problems through demand-driven design, systematically identifying and incorporating traditional leaders’ information needs at each stage of the design process. It does this through frequent consultation with the end user and design iteration.

A paucity of decision-relevant information is not exclusive to Southern Province or traditional leaders, affecting actors across domains in developing countries. The learnings from the Nano pilot in Zambia are applicable for other levels of government across Zambia and governance structures in other developing countries.

---

2 Mukobela Chiefdom is located in Namwala District, Southern Province. The district has a population of 102,866 across four chiefdoms. Mukobela Chiefdom is approximately 807 km², part of which includes the district centre and central business district. Mukobela Chiefdom has eight village clusters and approximately 130 villages.


2. DESIGN AND EXECUTION

The Nano team demonstrated several technical and systemic achievements in the pilot. These include developing a series of processes for collection and presentation of information relevant to a localised key decision-maker, a lean sampling method, and an Android application customised to the preferences and capacities of Chief Mukobela. We successfully conducted two rounds of data collection and presented results to the chief on access to services, preventative health behaviours, attitudes toward education, beliefs and behaviours on child marriage, and school resources and challenges. This chapter describes the activities undertaken, decisions taken at each point, and lessons learned from the process.

Relevant sections from the progress report are available in Section 1 of the Appendix.

DEVELOPING THE NANO APPLICATION

We produced an Android-based application to present data to Chief Mukobela. The interface is intentionally streamlined, simple, and suited to Chief Mukobela’s needs and preferences. The visual design communicates to the user the function and value of Nano: there are only two buttons on the landing screen buttons: “View Data” and “Request Data.”

The data visualisations are accurate but minimal, tailored to an audience with limited statistical background and intended to facilitate decision-making. Figure 2 presents screenshots from the Nano application, showcasing the landing screen and example visualisations.

![Screenshot of Nano application](image)

**Figure 2: Screenshots from the Nano application**

We opted for creating our own mobile application, instead of using an existing application or displaying data in a stand-alone deliverable like a slideshow, in order to preserve scalability, speed, and interactivity. Unlike a stand-alone deliverable, an application enables the user to submit feedback or requests at any time. This dynamic and interactive function is better aligned with the principles behind Nano than a stand-alone deliverable could be.

---

6 While the ‘request data’ function is not currently active, we would like to build out this function in the next Nano iteration.
We additionally considered using an existing application but found that those available were not well-suited for our audience. Creating our own application also gave us full control over how the data is displayed, instead of limiting us to any specific chart types or aesthetic defaults. This is an important component because the data is complex, particularly with respect to sampling uncertainty, and our typical user will have little to no statistical background. We intentionally produced the application as mobile-based, rather than web-based, to maximise functionality in a connectivity-limited region, and chose to build an application for Android devices because the chief uses an Android tablet, and Android devices are most common in low-resource settings.

**INTEGRATING USER VOICE IN THE NANO MODEL**

At each stage, Nano has prioritised stakeholder buy-in and user experience, aiming to adhere to a ‘demand-driven’ principle. In this pilot, Chief Mukobela is our primary user, and we have worked to develop a model of Nano aligned with his stated information needs. Simultaneously, we have maintained contact with other ministries at the district and provincial levels to generate enthusiasm and familiarity with Nano.

**Survey module selection**

In the early stages of the pilot, we consulted closely with Chief Mukobela to elicit his information needs and understand his ability to act in different topic areas. Accordingly, we generated an assessment of the viability and value of each potential module (Figure 3), prioritising his preferences alongside consideration of other factors that influence social impact, such as module actionability, interest from other stakeholders, and the feasibility of collecting high-quality data for each module.
Figure 3: Assessment of viability and value of potential modules

<table>
<thead>
<tr>
<th>Survey module</th>
<th>Interest from Chief</th>
<th>Interest from others</th>
<th>Decision relevance</th>
<th>Feasibility of quality data collection</th>
<th>Training ease</th>
<th>Involves other ministries</th>
<th>Overall assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child marriage</td>
<td>Very high</td>
<td>Very high</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Yes</td>
<td>Implement</td>
</tr>
<tr>
<td>Access to services</td>
<td>High</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>No</td>
<td>Implement</td>
</tr>
<tr>
<td>Education</td>
<td>High</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>Yes</td>
<td>Implement</td>
</tr>
<tr>
<td>Training needs</td>
<td>Medium</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>No</td>
<td>Do not implement</td>
</tr>
<tr>
<td>Preventative health</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>Yes</td>
<td>Implement</td>
</tr>
<tr>
<td>Demographics</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>No</td>
<td>Implement</td>
</tr>
<tr>
<td>Land rights</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Yes</td>
<td>Do not implement</td>
</tr>
<tr>
<td>School resources</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>Yes</td>
<td>Implement</td>
</tr>
</tbody>
</table>

Notes: In assessing the viability and value of modules, we considered several factors that we determined could affect the eventual social impact of the data. The initial list of potential modules was constructed solely based on interest from Chief Mukobela, as we viewed his engagement as the primary factor in determining social impact.

We additionally considered interest from other stakeholders, particularly ministry counterparts, because action in this context typically requires collaboration across ministries. We similarly prioritised modules in which we expected Chief Mukobela had the authority and capacity to take meaningful action. We also favoured modules for which we expected we could feasibly collect high quality data and train inexperienced data collectors to measure properly. We sought out an opportunity to collect data on a subject that involves other ministries, in order to establish a proof of concept for a data collection system that cuts across sectors.

Lessons learned on module design

We designed short and flexible data collection modules, basing questions on those from standardised household surveys, e.g. the Demographic and Healthy Survey (DHS). Modules were designed to be usable across a wide variety of regional, country, and user contexts as part of a scaled Nano, and this is something we will continue to prioritise in the design of new modules.

Results from the surveys translated easily to simple visualisation types and were informative and decision-relevant to Chief Mukobela (see 14). In the semi-structured interviews, Chief Mukobela demonstrated comfort with interpretation of results and identification of potential actions in response to the data. In particular, data on the average state of resources or behaviours is highly actionable, and future modules should prioritise these types of indicators. In the future, we may explore condensing questions to reduce the time needed surveying and further tailoring of question design to the types of actions that the indicator could prompt.
**Stakeholder management**

We maintained touchpoints with all relevant ministries and sought appropriate permissions when applicable. In addition to frequent meetings with Chief Mukobela, we have held meetings with the District Commissioner\(^7\), and both district and national-level officials from the Ministry of Chiefs and Traditional Affairs, as well as district-level representatives from the Ministry of Health and the Ministry of General Education. In these engagements, we have updated stakeholders to maximise action from the data, and have explored avenues for continuation of Nano in the future (see Chapter 6 for more information).

**MANAGING ENUMERATORS**

Enumerator management is central to producing accurate, fast, and low-cost data. There are five components for successful and low-cost enumerator management: 1) selection of high-quality embedded enumerators, 2) targeted training, 3) task automation, 4) enumerator performance assessment, and 5) provision of incentives and feedback.

To produce a nimble and cost-effective IS, we used locally embedded enumerators. We created a nomination system and tools for identification of individuals who were trainable and able to execute high quality data collection.

Throughout the process of data collection, we employed enumerator management methods focused on limiting the amount of direct oversight needed. Among these methods were:

- In-person supervision by local supervisor in the first round of data collection
- High frequency checks in the form of daily review of data submissions to identify any data irregularities
- GPS tracking of enumerators to ensure fidelity to assigned enumeration areas
- Audio audits: random recording of survey snippets with external party verification
- Bonus payments for high-quality performance
- Regular, remote feedback to enumerators on their performance metrics

Following data collection, we administered a survey to enumerators for their feedback on training, data collection, wages, and bonus payments. Overall, enumerators reported feeling well prepared for surveying – on a 5-point Likert scale\(^8\), the average response on preparedness was 4.57. However, the enumerators also felt the timeline for surveying was too short. Enumerators additionally felt that the payment was too low, suggesting instead an hourly wage of about 19 ZMW. Similarly, though they expressed the bonus payment system was well explained and effective in incentivising effort, enumerators felt that the bonus amount should be higher. While we will not necessarily lengthen the timeline and increase the pay, we will consider this feedback, alongside aforementioned factors, when setting pay and timeline targets in the future.

---

\(^7\) A District Commissioner is appointed by the President to oversee field administration in a district. There are 72 districts in Zambia, and a district covers a greater area and population than a chiefdom.

\(^8\) A Likert scale is a tool used to measure opinions ranging from one extreme to another. Respondents are asked to assign ratings to items according to a linearly increasing scale.
Efficacy of enumerator management and data quality methods

Audio audits and high-frequency checks were particularly effective in identifying low-quality data and points of misunderstanding among enumerators. For audio audits, we intentionally selected questions with multiple sections or those in which an enumerator may shirk their full responsibilities. We employed an outside assessor to listen to snippets of these questions and independently record the answer given by the respondent. From comparison of the assessor’s and enumerators’ recorded responses, we were able to identify underperforming enumerators and provide feedback accordingly. Similarly, we created a robust set of high-frequency checks to track responses to key questions and proxies for quality data collection (e.g. time spent in administering the consent process). We conducted high-frequency checks nightly on the incoming data and flagged outliers on key variables in an Excel dashboard. From this data, we produced and delivered individual feedback to enumerators on a near-daily basis.

These methods were effective in identifying issues and moderately effective in improving performance. However, each required significant input from IDinsight staff. For example, conducting high-frequency checks and delivering feedback typically required an hour and a half of a team member’s time each day. In a scaled-up model, we would reduce these costs by: a) further automating the process of running high-frequency checks, b) building capacity among supervisors to translate the high frequency checks output into personalised feedback, and c) standardising feedback messages.9

We offered enumerators bonus payments based on performance, with the intent of incentivising effort and improvement. We produced a short rubric to assign points according to signals of high effort and quality. Daily, enumerators were eligible for a 5 ZMW bonus for meeting a threshold of 12 points out of a possible 24, and the top two performers received an additional 10 ZMW. In practice, these bonus payments were less effective than expected in influencing behaviour and effort because of connectivity constraints. Several enumerators were not able to submit forms on a daily basis, encountering issues in uploading forms to the server. As a result, receipt of forms was often staggered, which complicated bonus point calculations. To some extent, this constraint is a feature of working in regional areas with low connectivity. In the future, however, we may consider paying for transportation for enumerators to travel daily to town to access better quality internet, in order to ensure daily receipt of survey forms. This is important for data quality and would simultaneously reduce noise in bonus point calculations, improving the efficacy of the tool in incentivising high performance.

Remote survey assignment and training

We trialled remote survey assignment and training with an example module to further explore how to reduce costs increase scalability. At scale, we expect that Nano will run across multiple geographies simultaneously with embedded surveyors in each region. We want to enable the user to request data as needed and deploy surveys rapidly, which will require comfort with remote survey assignment and training. A sophisticated remote tasking system will maintain data quality, while reducing time and financial costs produced by in-person travel to introduce enumerators to a new survey.

9 With respect to audio audits, we may be able to render the process more efficient by further automating the transfer of audio files from our surveying software to a repository for the external assessor.
In this trial, we communicated with enumerators over a messaging platform, sending them a short agriculture survey and related training manual. In tandem with the manual, we sent enumerators a quiz via Google Forms to test understanding of the new survey. There were no limits set on time spent on the quiz or number of attempts, but we required a grade of 85% before the enumerator could begin surveying.

Enumerators found it difficult to meet this threshold, suggesting that further work is needed to learn how to effectively teach a new survey remotely. While eight enumerators attempted the quiz, three never reached the minimum threshold, and on average, enumerators attempted the quiz 6.5 times. The threshold was set high to ensure demonstrated comprehension from enumerators before surveying. In future exploration, we do not intend to lower this standard, but we will experiment further with our teaching methods. We may also alter the quiz format to ensure that the quiz properly measures comprehension of the survey material, rather than other skills. For example, enumerators may be uncomfortable using Google Forms, and this may have erroneously contributed to a high failure rate. We were unable to thoroughly assess the data quality of surveys administered in this trial, so future work on this innovation should aim to optimise the quiz comprehension threshold more intentionally.

**Data collection methods considered and executed**

We considered five data collection methods prior to implementing the pilot:

1. Collating existing administrative data
2. In-person household surveys
3. Phone household surveys
4. Observations of facilities (e.g. schools, water sources, grazing land)
5. Interviews with service providers (e.g. teachers, headmasters, health workers)

We used the following criteria to decide which of these methods to implement:

a. **Needs alignment** – was the method going to collect the data the chief needed?

b. **Data quality** – was the method going to produce accurate data?

c. **Feasibility** – was the method feasible given our goals for Nano to be fast and cheap?

Prior to the first data collection round, we considered administering modules on public facilities, relying on enumerator observation, instead of household surveys. We additionally considered integrating phone surveys but chose not to proceed with this data collection method because most people do not have phones or reliable access to electricity in Namwala.

In our initial scoping visit, we visited several facilities in the area, including: a health clinic, schools, and the district outposts for the ministries overseeing health, agriculture, and education. At each, there is some form of data collection, but it is often aggregated very locally or more broadly than a single chiefdom, and the data collected is specific to the priorities of the ministry. Overall, data collected by each ministry does not align well with the chief’s priorities and capacities; for example, the district-level outpost of the Ministry of Health monitors medicine stock-outs at each health facility in order to initiate re-distribution of resources when necessary. Administrative data is also sensitive, as it could reflect poorly on a specific ministry’s performance, so approval typically takes a long time.
Conversely, Chief Mukobela is interested in monitoring household-level trends, e.g. community knowledge on child marriage, that are not currently tracked in the chiefdom. Given this mismatch in priorities and capacities, we decided that primary data collection would be necessary to meet the information needs of Chief Mukobela.

There is currently no formal channel for sharing data collected by other entities with Chief Mukobela. Moving forward, Nano could fulfil this function, serving as a conduit for data collation in addition to data collection.
The Nano team presented data results twice to Chief Mukobela and documented reactions and intended actions on the outcomes. This chapter describes Chief Mukobela’s reactions to the data presented, actions realised and intended by the chief, and lessons learned on the process of translating information into decisions and actions.

**REACTIONS AND OPINIONS OF THE CHIEF**

In each round of data presentation, Chief Mukobela was enthusiastic and eager to understand the results and consider possible actions in response.

In general, Chief Mukobela typically relies on *ad hoc* information sources to determine the chiefdom’s needs, such as:

- Contextual knowledge of the chiefdom. For example, the chief extrapolated an estimate of the number of deliveries occurring in health facilities based on the facilities he knows have mothers’ shelters.
- Observation. For example, the chief estimated incidence of child marriage based on his observation of young-looking pregnant girls in hospitals.

Many results contrasted with the chief’s prior beliefs. In these instances, the chief often stated that the results were surprising to him and that he would not have been able to obtain the information outside of Nano.

Chief Mukobela demonstrated flexibility in adapting his beliefs to match the data and reinforced his updated beliefs by proposing actions that were consistent with these beliefs. Table 2 and Table 3 present Chief Mukobela’s prior and posterior beliefs on select indicators in the first and second rounds of data collection, respectively.
<table>
<thead>
<tr>
<th>Indicator</th>
<th>Chief’s prior</th>
<th>Source of information for prior</th>
<th>Round 1 estimate</th>
<th>Potential actions</th>
<th>Stakeholders</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of HHs that drink treated water</td>
<td>1% to 2%</td>
<td>Observation from community</td>
<td>31% (CI: 26%, 36%)</td>
<td>Committees to sensitise on water treatment</td>
<td>Community</td>
</tr>
<tr>
<td>% of HHs with child marriage</td>
<td>40% to 60%</td>
<td>Observation of pregnant girls and school dropout rates for girls</td>
<td>14% (CI: 8%, 20%)</td>
<td>Meeting for chiefs on fighting child marriages organized by MoH</td>
<td>Community, MoH</td>
</tr>
<tr>
<td>% of women who delivered or plan to in a hospital and health clinic,</td>
<td>Majority of women deliver in hospital</td>
<td>Knowledge of facilities with mother’s shelters</td>
<td>78% in hospital (CI: 69%, 87%) 11% in health clinic (CI: 6%, 17%)</td>
<td>Work with headmen to encourage antenatal care; construction of mother’s shelter at Baambwe clinic</td>
<td>Community, village headmen / headwomen</td>
</tr>
<tr>
<td>% of women who delivered or prefer to deliver at home</td>
<td>Varies by area, 5% to 30%</td>
<td>Knowledge of facilities with mother’s shelters</td>
<td>11% (CI: 6%, 17%)</td>
<td>Construction of mother’s shelter at Baambwe clinic</td>
<td>Community</td>
</tr>
<tr>
<td>% who believe condoms effectively prevent HIV/AIDS</td>
<td>10% to 20%</td>
<td>Observation of condom sales and distribution at the hospital</td>
<td>42% (CI: 33%, 50%)</td>
<td>Create AIDS task force to sensitise communities on HIV/AIDS with MoH support at large gatherings</td>
<td>Community, MoH</td>
</tr>
<tr>
<td>% who do not believe sharing food or drink with someone who has HIV/AIDS prevents HIV/AIDS</td>
<td>95%</td>
<td>Observation of people who share utensils and drinks</td>
<td>81% (CI: 77%, 84%)</td>
<td></td>
<td>Community, MoH</td>
</tr>
<tr>
<td>% of children under 5 who had diarrhoea in the last one month</td>
<td>Chief unwilling/unable to give prior</td>
<td>Diarrhoea has reduced since they become open defecation free</td>
<td>45% (CI: 39%, 52%)</td>
<td>Task force and health workers increase knowledge of water treatment; lobby MoH for more chlorine</td>
<td>Community, MoH</td>
</tr>
<tr>
<td>% of diarrhoea cases in children under 5 that are treated</td>
<td>80% to 90%</td>
<td>Guess</td>
<td>76% (CI: 69%, 83%)</td>
<td>Mobilise community increase Chibunze school building; lobby partners to finance the construction</td>
<td>Community, MoH</td>
</tr>
<tr>
<td>Village cluster in which children travel farthest to reach schools</td>
<td>Chief unwilling/unable to give prior</td>
<td>No information provided</td>
<td>Chibunze</td>
<td></td>
<td>Community, Ministry of General Education, World Vision</td>
</tr>
</tbody>
</table>
Table 3: Round 2 priors, posteriors, and data results for Chief Mukobela with estimates and confidence intervals (CI)

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Chief’s prior</th>
<th>Source of information for prior</th>
<th>Round 2 estimate</th>
<th>Potential actions</th>
<th>Stakeholders</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of HHs that would send girl over boy to school in the vignette</td>
<td>30-40%</td>
<td>Census; community concern over pregnancy in young girls</td>
<td>42.1% (CI: 37.2%, 47.1%)</td>
<td>Lobby NGOs to sensitise communities on rights of girls</td>
<td>NGOs, Chief’s Council</td>
</tr>
<tr>
<td>% of children enrolled in school who were chronically absent(^{11}) in the last month</td>
<td>Chief unwilling/ unable to give prior</td>
<td>Reports from pupils and teachers</td>
<td>34.8% for girls (CI: 29.8%, 39.9%); 54.7% for boys (CI: 49.0%, 60.5%)</td>
<td>Improve sanitation facilities in schools for female pupils</td>
<td>Latrine builders and sanitation groups in the chiefdom, NGOs</td>
</tr>
<tr>
<td>Average number of pupils per toilet</td>
<td>Chief unwilling/ unable to give prior</td>
<td>District education office</td>
<td>70.5</td>
<td>Ask toilet builders and sanitation action groups to help with latrine construction</td>
<td>Latrine builders and sanitation groups in the chiefdom, NGOs</td>
</tr>
<tr>
<td>% of schools that have access to electricity</td>
<td>2-5%</td>
<td>Personal knowledge of schools</td>
<td>27.8%</td>
<td>Lobby government to electrify schools</td>
<td>Government</td>
</tr>
<tr>
<td>% of teachers living in staff housing</td>
<td>40%</td>
<td>Observation of school infrastructure</td>
<td>23.8%</td>
<td>Encourage communities to dedicate resources for construction of staff housing</td>
<td>Communities</td>
</tr>
<tr>
<td>Average number of classroom blocks that have improved floor or weather-resistant roof</td>
<td>90%</td>
<td>Personal knowledge of schools in the chiefdom</td>
<td>84.7% improved floor 88.1% weather-resistant roof</td>
<td>Request that community members contribute materials to improve school infrastructure</td>
<td>Community members, Partners in Action (USA), World Vision, EL Foundation</td>
</tr>
</tbody>
</table>

**ACTIONS TAKEN AND INTENDED BY THE CHIEF**

In response to the data, Chief Mukobela has put significant effort towards community resource mobilisation, and stakeholder engagement for issues requiring central government support or resources exceeding those available in the community. We recently made the second round of data available to Chief Mukobela and expect to see evidence of actions based on these results at a later date.

---

\(^{10}\) Values obtained from the school survey do not have confidence intervals because we interviewed all schools, meaning that the survey sample is the full population.

\(^{11}\) Chronic absenteeism is defined as having missed 10% or more school days.
When feasible, Chief Mukobela has taken direct and immediate action on issues identified in the data. The results showed Chibunze was the area in the chiefdom with the longest travel times to health facilities. In response, Chief Mukobela organised material contributions from community members to begin construction on a clinic in the area supported by a grant from the District Council. On other issues, Chief Mukobela has appealed to government and NGO partners for support to improve critical indicators. To facilitate these conversations, Chief Mukobela requested basic paper reports showing key indicators and has shared them with officials from the District Council, Ministry of General Education, and Ministry of Health, as well as NGOs like World Vision and Oxfam.

In the table and section below, we outline responses from the Chief to specific indicators.

**Table 4: Overview of actions taken and intended by the chief based on Nano**

<table>
<thead>
<tr>
<th>Indicator/s</th>
<th>Key finding/s</th>
<th>Resulting actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diarrhoea and water</td>
<td>A high proportion of children under 5 had diarrhoea recently, and less than a third of households treat their water.</td>
<td>The chief plans to form committees to sensitise communities on water treatment. He is also brainstorming ways to reduce cost and distance barriers for chlorine usage.</td>
</tr>
<tr>
<td>Location of births</td>
<td>12% of births occur at home.</td>
<td>The chief is mobilising the community to construct a mother’s shelter at Baambwe clinic, which is far from the hospital, to provide accommodation for pregnant women. The chief is working with the Ministry of Health to provide staff to the clinic.</td>
</tr>
<tr>
<td>Access to education</td>
<td>Pupils in Chibunze village cluster spend 5.8 hours in rainy season travelling to secondary school.</td>
<td>The chief aims to build a secondary school classroom block at Chibunze Primary School. Community members will contribute labour/resources and the chief will reach out to partners (Ministry of Education and World Vision) to finance the completion of the buildings.</td>
</tr>
<tr>
<td>Child marriage</td>
<td>One fifth of households reported knowing a case of child marriage in their community in the past year, and 24% reported being unwilling to report a case of child marriage.</td>
<td>The chief intends to share the child marriage data with NGOs to spur them to act on child marriage. The chief will also use the data at a meeting for chiefs on combating child marriage.</td>
</tr>
<tr>
<td>School sanitation infrastructure</td>
<td>On average, there is one toilet for every 75 pupils, well below government guidelines. Toilet facilities are often poor quality and often lack hand washing facilities.</td>
<td>The chief aims to mobilize resources from relevant communities (including sanitation action groups) to construct toilet facilities at schools. He plans to reach out World Vision and DAPP to explore whether they can support the toilet construction.</td>
</tr>
<tr>
<td>Education beliefs and behaviours</td>
<td>57% preferred to pay for boys’ education over girls’, in a scenario we presented. This choice was driven by beliefs that girls may get pregnant at school and that boys’ education is more important.</td>
<td>The chief plans to write to Oxfam, World Vision, and the Esther Lungu Foundation asking them to implement campaigns to sensitise communities in the chiefdom on girls’ rights. He also plans to ask the Chief’s Council to develop strategies to use local leaders (village headmen/headwomen) to change attitudes on girls’ rights within their communities.</td>
</tr>
</tbody>
</table>
**Diarrhoea and water**

The data showed 45% (CI: 39%, 52%) of children under 5 had diarrhoea in the month before the survey. This surprised the chief, as he thought issues related to diarrhoea were addressed with improved sanitation and declaration of the chiefdom as open defecation free (ODF). The data further showed that only 30% of the households treat their drinking water – many households do not treat water because they do not think it is useful (10%, CI: 7%, 14%) or cannot access chlorine (57%, CI: 42%, 70%). The chief linked the high incidence of diarrhoea in children to low rates of water treatment and thus decided to consider reasons for not treating water.

The Chief plans to form committees to sensitize communities on water treatment. These committees will function similarly to sanitation action groups, which were used to eliminate open defecation within the chiefdom. He is also brainstorming ways to reduce cost and distance barriers for chlorine usage.

**Location of births**

12% (CI: 5%, 19%) of babies are delivered at home, 78% (CI: 69%, 87%) are delivered in a hospital, and 11% (CI: 6%, 17%) are delivered in health clinics. The chief is concerned about the current quality of service provision in health facilities and risks associated with home delivery, and he hopes to increase the number of hospital deliveries and the quality of health facilities.

The chief suspects that there are more non-hospital deliveries in areas farther from Namwala, where the hospital is located. He also expects that a reason for lack of delivery in hospitals is long distances to the hospital and a lack of mother’s shelters at health clinics. We were not able to provide data that supports these claims due to the small number of women who had recently given birth in our sample. We knew this ahead of time and communicated this to the chief.

Still, the chief was inspired to act based on the Nano data and is mobilising the community to construct a mother’s shelter at Baambwe clinic, which is far from the hospital (Figure 4) and does not have a mother’s shelter. The Nano data convinced the chief’s senior headmen to prioritise the shelter over the establishment of a market in Baambwe ward.

The chief has asked community members to contribute materials and labor towards this construction, and is working with the Ministry of Health to provide staff to the clinic. The new clinic will help to increase facility deliveries in Baambwe, as pregnant women will now be provided with accommodation at the clinic.
Figure 4: Average distance to health facility or hospital

Access to education

The data showed that schoolchildren in Chibunze village cluster spend the most time travelling to secondary school, especially during rainy season. While the chief had a sense that children in Chibunze had to travel long distances, and he knew that Chibunze is often cut off from the rest of the chiefdom during rainy season, he was surprised by the magnitude of travel times (1.7 hours in dry season (CI: 1.4, 2.0), 5.8 hours in rainy season (CI: 5.5, 6.1), as shown in Figure 5). This motivated the chief to think about the need to expand school facilities in Chibunze.

Figure 5: Average Distance to secondary school

His proposed solution is to build a secondary school classroom block at Chibunze Primary School. Community members will contribute resources and labor towards the construction of the classroom block. The Chief will then reach out to partners (MOGE and World Vision) to finance the completion of the buildings. The Chief is confident that this will work because that was the model used for construction of other schools in the chiefdom.
If Chief Mukobela’s proposed solution is successful, the school can offer grades 8 and 9, hence reducing the need for children in Chibunze to travel over 5 hours during rainy season and likely increasing school attendance levels in Chibunze.

**Child marriage**

One fifth of households reported knowing a case of child marriage in their community in the past year, and 24% reported being unwilling to report a child marriage (when asked about a hypothetical case). Only 3.5% of respondents reported that the legal age for marriage was lower than 18 years old (the legal age in Zambia). This confirmed to the chief that child marriage is an issue in the chiefdom and that knowledge of the illegality of child marriage is not a barrier to addressing the issue.

The chief is concerned that NGO partners currently working in the chiefdom are not active in following up on cases of child marriage and addressing the causes of child marriage. The chief intends to share the data with NGOs to spur them to act on child marriage. One NGO, Oxfam, has previously expressed interest in undertaking a project to combat child marriage in the chiefdom. The chief hopes that the child marriage data will provide Oxfam with the additional impetus it needs to begin work in the chiefdom on this issue.

The chief is slated to attend a meeting for chiefs on combating child marriage, organised by the Ministry of Health. He will attend this meeting with recent data on prevalence and attitudes towards child marriage in this chiefdom and will be well placed to advocate for evidence-backed approaches for combatting child marriage.

**School sanitation infrastructure**

Our field team conducted interviews with head teachers or teachers in charge at 18 government and community schools in the chiefdom. We found that on average, there is one toilet for every 75 pupils, well below government guidelines. Only 63.3% of toilet facilities in schools were improved (VIP latrine or flush toilet) and only 55% of schools have hand washing facilities.

The chief proposed to mobilize resources from relevant communities (including sanitation action groups) within the chiefdom to construct toilet facilities at schools. The propose course of action requires little intervention from external stakeholders and both materials and labour will be provided by communities. This will potentially increase the sense of ownership for public amenities by communities and lead to better maintenance – a challenge noted by government education officials. Chief Mukobela recognizes the role of NGO partners, who have in the past constructed toilets at schools – he plans to reach out World Vision and DAPP to explore whether they can support further toilet construction.

---

12 The chief has identified community attitudes and a lack of economic opportunity for young women as key causes of child marriage.

13 In Zambia, government guidelines specify a ratio of 25 male pupils per toilet and 20 female pupils per toilet – only 9% met these ratios.

14 Only 37.2% and 25.5% of hand washing facilities had water and soap respectively.

15 DAPP stands for Development and Aid from People to People.
**Education beliefs and behaviours**

To collect household’s beliefs on girls’ education, we presented respondents with a vignette. Households were asked to imagine that a respected family in their community has one daughter and one son who are the same age and attend the same school. The family is then unable to afford to send both children to school – households were asked which child the family should continue sending to school. Of those willing to make a choice, 57.1% said the family should pay for the education of the daughter over the son.\(^{16}\) This choice was driven by beliefs that girls may get pregnant at school (39.2%), boys’ education is more important (26.2%) and boys perform better in school (19.2%).

The chief was concerned that households do not prioritise girls’ education because they fear girls might drop out because of pregnancy. He plans to write to Oxfam, World Vision, and the Esther Lungu Foundation to ask them to implement campaigns to sensitise communities in the chiefdom on girls’ rights. He also plans to ask the Chief’s Council to develop strategies to use local leaders (village headmen/headwomen) to change attitudes on girls’ rights within their communities.

**LESSONS LEARNED ON TRANSLATING DATA TO ACTION**

The translation from data to action is a key step in realising social impact, and we have worked to design the system to maximise this link. In the app, at the bottom of each data module, we have listed relevant stakeholders and their contact information, nudging the chief to contact them for support on critical indicators. At the chief’s request, we have produced a simple paper report, which he uses in meetings with stakeholders. We expect that this is scalable, as the report is sufficiently basic to produce quickly, and we should be able to automate production at scale. We have provided the chief with multiple copies of this paper report to share at his discretion.

We have also aimed to build steps into the process to encourage data actionability. Following both rounds of data collection, we scheduled a full-day session with Chief Mukobela to present results, discuss possible actions, and identify associated stakeholders. Each session was guided by a thorough interview guide, which prompted Chief Mukobela to systematically consider each key indicator and brainstorm feasible actions in response. These meetings have been a valuable lens into Chief Mukobela’s interpretation of data and capacity to respond to issues, and we believe that the meetings have been highly effective in prompting action. Indeed, Chief Mukobela’s stated actions on the data results from the first round align well with the actions we brainstormed during the first-round data presentation, e.g. community mobilisation for construction of a health facility in Chibunze.

While Nano operation will become more remote at scale, it would be valuable to carry out this type of meeting once with each Nano client. We have found that this meeting is very helpful in anchoring the user on associating the indicators with clear plans for action, and early results suggest that the user does then execute on their suggested actions. In future iterations, we will continue to work towards greater facilitation of data-sharing with external stakeholders, as action in these settings typically requires collaboration across partners. Specifically, we are interested in adding a ‘share’ function to the application, allowing the user to save and email data visualisations directly from the application.

\(^{16}\) Of the full sample, 5.2% said they would not send either child to school and 2.0% refused to respond.
COMMUNITY ENGAGEMENT WITH THE DATA

Overall, households were willing and engaged participants. We surveyed some households twice, including them in the samples for both the first and second round of data collection, and encountered no substantial difficulties or reluctance to participate. In Rounds 1 and 2, on average, 9.6% of households refused to participate. We encountered a few people who refused on the basis of personal differences with the chief, but these were limited and isolated incidents, and we do not believe this affected overall Chiefdom estimates.

In future iterations, we may think more carefully about how to involve the community in the Nano process. Currently, however, sensitivities around participants’ privacy seem to outweigh the value-add of sharing results with community members. Additionally, an expectation of data results being shared with the community or excessive information on the actions that may be taken in response to the data may introduce perverse incentives to respond dishonestly to survey questions, e.g. to obtain more resources for their region by exaggerating the current state.
4. EVALUATION OF NANO PILOT

At the onset, we envisioned Nano as an alternative to standard IS that is cheap, fast, and demand-driven with high user satisfaction. In this chapter, we assess the performance of Nano on each of these dimensions. Additionally, we gauge the social impact achieved in this pilot and the potential for future impact from this system.

COST AND SPEED

To benchmark the cost and speed of Nano, we construct an estimate of collecting the data we collected in the pilot using standard data collection methods. We compare this to costs under Nano, assuming set-up in a new region.

We construct the Nano estimate based on what we expect is feasible at scale, based on evidence from the pilot. For the Nano cost calculations, we omit time spent on application development, as this will not need to be redone in new geographies, though there may be some application maintenance involved. Similarly, we do not include time spent on survey design and coding in the Nano costs, as we are working towards establishing a library of modules that can be re-used across regional, country, and user contexts.

Estimates of parameters for standard data collection are constructed based on the realised time and financial costs for another IDinsight project deployed in 2018 surveying 400 households in Southern Province, Zambia. While this is only a single counterfactual, we believe this project is the most similar to the data collection activity in Nano and serves as a reasonable comparison.

In each scenario, we assume deployment of a household survey that is approximately 20 minutes in length and administered to 700 households in a rural area similar to Namwala, Zambia. We additionally assume a team of about 20 enumerators. These parameters are informed by the set-up of the first data collection activity conducted in the pilot, in which a team of 20 enumerators counted 1,846 households in the region and administered a survey lasting, on average, 18.5 minutes to 687 of these households.

From the pilot, we find evidence of lower time and financial costs under a Nano model compared to standard data collection (see detailed discussion below). We attribute this to:

1. Sampling via satellite imagery and publicly available datasets
   a. Bypasses need for an expensive household listing exercise or an existing household registry, while still achieving a random and representative sample
2. Development of modules re-usable across contexts
   a. Enables repeat use of survey document and associated code with limited additional work
   b. Enables repeat use of subsequent data cleaning and analysis files with limited additional work

---

17 We have not explicitly included application maintenance costs in our estimates. We expect that maintenance will be minimal, and at this stage, there is insufficient information to form a cost estimate on this.
3. Use of locally embedded enumerators
   a. **Eliminates need for transportation from Lusaka** to the surveying region for enumerators
   b. **Eliminates need for per diem and lodging** for enumerators
   c. **Reduces need for transportation within the surveying region**, as enumerator selection was conducted strategically to ensure representation in residence throughout the region

4. Use of remote supervision methods
   a. **Reduces costs of IDinsight staff on-the-ground**
   b. **Eliminates need to hire in-person supervisors** to shadow enumerators

Substitution of standard data collection methods with locally embedded enumerators and remote supervision may somewhat reduce data quality, but based on the pilot, we expect that we can maintain adequate data quality through the use of multiple remote supervision methods, including audio audits, high frequency checks, and GPS tracking of enumerators. We also anticipate further innovations will be developed by IDinsight over the next two years, such as machine learning identification of data quality issues, which will further increase data quality and reduce costs.

Audio audit data presents some evidence on the data quality possible under a Nano model, suggesting that high-quality enumerators can be identified via a combination of the enumerator selection process and high frequency checks. In the first round of data collection, enumerator responses aligned with the auditor’s responses 68% of the time, on average. For the second round, we reduced the enumerator pool, keeping those who performed best on high frequency checks, with consideration given for geographical distribution. In the second round, the rate of alignment between enumerator and auditor responses rose to 75%, which we believe is reflective of an increase in average enumerator quality.\(^{18}\) In the future, the enumerator selection process should include a week of piloting, after which we can select enumerators based on assessment performance and high frequency check results.

We are interested in continuing to improve data quality with greater investment in innovations that flag markers of low data quality in incoming surveys and further exploration of the optimal design of bonus payments.

In the time and financial estimates below, we assume the parameters outlined in Table 5.

---

\(^{18}\) Alignment between enumerator and auditor response means that having listened to an audio clip from the survey, the auditor recorded the same response as the enumerator. Some disagreement between the auditor and enumerators is expected. Agreement lower than 100% does not necessarily suggest enumerator error, and may instead be produced by auditor error, low-quality audio, or ambiguous response from the respondent. These agreement rates are on par with another IDinsight study that tested use of resident enumerators and general trends for non-resident enumerators. The other IDinsight study that used resident enumerators found 77% agreement between enumerator responses and supervisor back-check responses on continuous variables.
Table 5: Assumptions for comparison of costs of standard data collection and Nano model

<table>
<thead>
<tr>
<th>Standard data collection</th>
<th>Nano model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enumerator recruitment</td>
<td>Enumerators recruited from surveying region with strategic dispersal throughout region</td>
</tr>
<tr>
<td>Enumerator payments</td>
<td>Daily wages: 85 ZMW, Per diem: 0 ZMW, Bonus payments: 0 – 15 ZMW</td>
</tr>
<tr>
<td>Enumerator lodging expenses</td>
<td>Daily expense: 0 ZMW</td>
</tr>
<tr>
<td>Transportation to/from surveying region</td>
<td>Cost per enumerator: 0 ZMW</td>
</tr>
<tr>
<td>Transportation within surveying region</td>
<td>Daily cost per enumerator: 50 ZMW</td>
</tr>
<tr>
<td>Enumerator efficiency</td>
<td>Households surveyed per day: 1, Days worked per week: 6</td>
</tr>
<tr>
<td>Survey development process</td>
<td>Survey design: 0 days, Survey coding: 0 days, Miscellaneous: 1 day</td>
</tr>
<tr>
<td>Data presentation process</td>
<td>Data cleaning: 0 days, Data analysis: 0 days, Data write-up and visualization: 0 days, Modifications: 1 day</td>
</tr>
</tbody>
</table>

A detailed description of the sampling and enumerator recruitment strategies employed in the pilot is available in the Appendix in Section 1.

Cost

Below, Table 3 outlines estimated overall costs for data collection under standard methods and the Nano model. As noted above, the estimate for the Nano model is for set-up in a new region at scale, informed by early evidence from the pilot. All numbers cited below are in USD.

---

19 The daily wage was paid for six hours of surveying, as we assumed two hours of transportation per day. Wages for resident enumerators were set at this level with the following considerations:

1. Contextual norm: we were advised by MOCTA district officials to not pay a wage higher than that of teachers or other formal employment positions.
2. Minimum wage: the minimum wage for a general worker is 5.48 ZMW per hour.
3. Sustainability: we wanted to choose a wage level that could be sustained by the government in a scale-up scenario.
Table 6: Comparison of financial costs of Nano model and standard data collection

<table>
<thead>
<tr>
<th>Phase</th>
<th>Expense</th>
<th>Nano model</th>
<th>Standard data collection</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Set-up</strong></td>
<td>Scoping visit</td>
<td>$6,461</td>
<td>$6,461</td>
</tr>
<tr>
<td></td>
<td>Sampling</td>
<td>$981</td>
<td>$14,823</td>
</tr>
<tr>
<td></td>
<td>Survey development</td>
<td>$885</td>
<td>$4,143</td>
</tr>
<tr>
<td></td>
<td>Survey CTO coding</td>
<td>$704</td>
<td>$4,666</td>
</tr>
<tr>
<td></td>
<td>Enumerator selection</td>
<td>$3,472</td>
<td>$362</td>
</tr>
<tr>
<td></td>
<td>Enumerator training</td>
<td>$4,001</td>
<td>$3,611</td>
</tr>
<tr>
<td></td>
<td>Enumerator transportation to region</td>
<td>$0</td>
<td>$910</td>
</tr>
<tr>
<td><strong>Execution</strong></td>
<td>Surveying</td>
<td>$3,341</td>
<td>$14,619</td>
</tr>
<tr>
<td></td>
<td>Enumerator transportation within region</td>
<td>$1,327</td>
<td>$13,268</td>
</tr>
<tr>
<td></td>
<td>Data quality monitoring</td>
<td>$2,172</td>
<td>$2,172</td>
</tr>
<tr>
<td></td>
<td>Data cleaning</td>
<td>$885</td>
<td>$3,902</td>
</tr>
<tr>
<td></td>
<td>Data analysis</td>
<td>$885</td>
<td>$3,862</td>
</tr>
<tr>
<td></td>
<td>Data presentation</td>
<td>$885</td>
<td>$5,306</td>
</tr>
<tr>
<td><strong>Total for basic data collection in a new region:</strong></td>
<td></td>
<td>$25,999</td>
<td>$78,103</td>
</tr>
</tbody>
</table>

As the table suggests, we expect that at scale, operation in a new region will be about a third of the cost of traditional data collection costs. These reductions in cost are primarily attributable to a lean sampling method and use of resident enumerators, which significantly shorten time spent sampling and surveying, and eliminate unnecessary additional costs like lodging for enumerators.
Speed

We compare here time estimates for a household survey conducted under the Nano model and under standard data collection methods. As noted above, the estimate for the Nano model is for set-up in a new region at scale, informed by early evidence from the pilot.

Table 7: Data collection timelines under Nano model and standard data collection

<table>
<thead>
<tr>
<th>Weeks</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nano model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scoping visit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sampling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enumerator selection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enumerator training</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surveying</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data cleaning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data transfer to app</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard data collection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scoping visit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sampling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survey development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survey coding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enumerator selection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enumerator training</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surveying</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data cleaning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data transfer to app</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DEMAND-DRIVEN DESIGN AND USER SATISFACTION

Chief Mukobela has expressed high satisfaction with the Nano system in two semi-structured interviews, commenting that the system was faster than he expected and that he would not have been able to access this information otherwise. Following the presentation of data from the first round, Chief Mukobela stated, “I never expected you to do it in the time that you said,” and “I would scratch my back to pay.”

---

20 The chief used the phrase “scratch my back” to imply that he would strive to find resources to pay for Nano if it were required.
How demand-driven is Nano?

When designing Nano, we strive to be demand-driven. The information we gather should be: appropriate for the given context, of interest and relevance to the decision-maker, and needed to fill a current information gap. We aim to provide information at the granularity, speed, and accuracy the decision-maker needs to make informed decisions for their community.

In the rubric below, we assess our current performance according to this description. In each row, the description listed under the highest value provides a benchmark for our ideal at scale.

*Figure 6: Assessment of Nano performance as a demand-driven product*

<table>
<thead>
<tr>
<th>Dimension</th>
<th>1 Adequate</th>
<th>2 Good</th>
<th>3 Ideal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content: How well does the information gathered reflect user interest?</td>
<td>Specific modules are developed according to suggestion by the primary user. Module questions are drawn from standardised household surveys with no alteration.</td>
<td>Specific modules are developed according to suggestion by the primary user. Module questions are drawn from standardised household surveys and customised to maximise actionability.</td>
<td>A wide range of module topics are developed according to suggestion by the primary user. Modules can be selected and deployed by the user at will. Module questions are drawn from standardised household surveys and customised to maximise actionability, and modules are customisable: questions can be removed by the user.</td>
</tr>
<tr>
<td>Information specificity: How well does the information generated adhere to the granularity, speed, and accuracy needed?</td>
<td>Information is gathered and presented primarily at the chiefdom level. Information turnaround is at least 10% faster than expected under standard data collection methods. There are checks in place to reasonably ensure data accuracy.</td>
<td>Information is gathered and presented primarily at a local level, accompanied by overall Chiefdom estimates. Information turnaround is at least 25% faster than expected under standard data collection methods. There are checks in place to reasonably ensure data accuracy.</td>
<td>Information is gathered and presented primarily at a local level, accompanied by overall Chiefdom estimates. Information turnaround is at least 50% faster than expected under standard data collection methods. There are checks in place to reasonably ensure data accuracy.</td>
</tr>
<tr>
<td>User and context accommodation: How well does the Nano process and interface accommodate user and contextual limitations and preferences?</td>
<td>Use of the Nano interface requires strong internet connectivity. Information is presented with high technical accuracy. Visualisations are statistically complex and appropriate for an audience with statistical background.</td>
<td>Some functionality of the Nano interface is reduced, but the interface remains somewhat usable without strong internet connectivity. Information is presented with high technical accuracy. Visualisations are somewhat simplified, but require explanation for users without statistical background.</td>
<td>The Nano interface is customised for regions with low internet connectivity and minimises internet dependence. Information is presented with high technical accuracy. Visualisations are intuitive and simple, even for users without statistical background.</td>
</tr>
</tbody>
</table>
User engagement:

How well does the Nano process provide opportunities for user engagement?

The user is consulted at the start and end of the process on preferences and feedback. Preferences expressed at the start of the engagement are considered throughout, but are not the top priority.

The user is consulted multiple times throughout the process. User preferences decide the information collected.

The user is consulted multiple times throughout the process. User preferences decide the information collected, process decisions taken, and features of the interface. There is an open channel for communication between the user and the Nano team throughout the whole process.

SOCIAL IMPACT

Nano will achieve social impact when a decision-maker uses information gathered by Nano to make a decision that benefits their community, for which they would have otherwise not made a decision or relied on lower quality information.

As described in Chapter 3, there is early evidence that Chief Mukobela is incorporating information gathered by Nano into his decisions and changing course accordingly. The most promising example of this is the construction of a health facility in Chibunze, identified in Nano data as the community with poorest access to existing facilities. Chief Mukobela stated that he would not have known this information otherwise. Chief Mukobela began efforts to mobilise resources after having received this information, and we believe that he would not have taken this action at this time without Nano.
5. DISSEMINATION OF LESSONS LEARNED

We are eager to disseminate learnings from the Nano pilot to maximise the social impact of our innovations and facilitate similar projects by others. Innovations like the development of a lean sampling method are applicable to most surveying work; we aim to incorporate this technique in several IDinsight projects to enable faster and cheaper data collection. Additionally, we hope to facilitate replication of the Nano process in other areas to continue to work towards greater information access for decision-makers.

LEVERAGING INNOVATIONS

We were able to achieve faster and cheaper data collection in this pilot by using locally embedded enumerators and sampling via satellite imagery.

Locally embedded enumerators

IDinsight has experimented with use of locally embedded enumerators in other projects, including a nutrition project in Ethiopia and the ongoing Data on Demand work in India. These teams have documented best practices and learnings with use of locally embedded enumerators over standard non-local enumerators, and the Nano team has learned from these experiences. The success of the Nano team in using local enumerators provides additional evidence of the feasibility of this method, and we are eager to document this process to encourage its use across IDinsight projects. To facilitate this, we will share documentation on the selection and management of local enumerators throughout the IDinsight team.

Sampling via satellite imagery

Creation of a representative sample is often difficult in rural areas, as it requires an up-to-date household register, which is typically unavailable in these regions, or an expensive censusing activity. The Nano team addressed this issue with a geographic sampling approach, using roof data from two sources superimposed on an electronic map of the region. We then divided the region into 500x500 metre regions, referred to as Enumeration Areas (EAs), and randomly sampled EAs for surveying.

We have published a step-by-step description of this process in an internal blogpost (IDinsight staff only) and an external blogpost (public). We included code and data needed to undertake the process, such that others aiming to replicate this process have full access to the resources used. Teams within IDinsight have already contacted the Nano team to learn more about this process and consider if a similar strategy is suitable for their workstreams.

DOCUMENTING THE NANO PROCESS

We have achieved several successes in the Nano pilot that provide early evidence that this model can facilitate greater information access for decision-makers. As such, we hope to enable others outside of IDinsight to undertake similar projects.

We will produce a full ‘DIY manual’ outlining the process of Nano set-up and operation for others to replicate. In this documentation, we will include an overall timeline, detailed step-by-step process descriptions, and commentary on best practices and learnings. With approval from GIF, we will include some text directly from the progress and final reports. We will link to any tools or code that are generalisable, omitting those of our technical outputs that are either data-sensitive or context-specific. We expect to be ready to publish this document in the next 1-2 months.

This will likely take the form of a Word document – depending on capacity, we would also be interested in publishing this document on a WordPress or Google site. We will work closely with our strategic communications team to reach a broad audience with our publication and look forward to sharing our documentation with GIF.
6. **NEXT STEPS FOR NANO**

We are eager to continue our work with Nano and to work with GIF to identify the best path forward. The Nano pilot has produced several points of success and learnings, and we hope to leverage this foundation in further work.

Our interest is scaling Nano through an approach that focuses on: 1) advancing technical innovations, 2) leveraging existing systems, and 3) working with new users. As we scale up, the guiding principles remain the same – we aim for Nano to be a faster, cheaper, and flexible IS to provide relevant data to decision-makers in any sector as demanded. In the next iteration of Nano, this may include both data aggregation and collection. We would additionally like to continue work with Chief Mukobela, in order to leverage our established presence in the chiefdom and facilitate continued action on Nano data.

We outline potential channels for scaling Nano below.

**CHANNELS FOR SCALING NANO**

1. **Recommended: Leveraging data for frontline workers**

In this iteration, Nano would: centre on use by frontline workers, leverage existing data, develop a system for data validation, and work towards a more flexible technological architecture.

With frontline supervisors as the primary client, this model of Nano would introduce collation and validation of existing data. Use of existing data increases the need for data quality verification – accordingly, a key component of this model would be exploration of potential methods to reliably ensure the quality of aggregated data is sufficient for the decisions of frontline workers. The subject of rapid and cheap data validation methods is relevant to a host of IDinsight projects and would lend itself well to cross-project collaboration.

To support data collection by frontline workers, we would focus on development and deployment of flexible, customisable, and streamlined modules. This work would build on that of the pilot, continuing to refine our current data collection strategies on dimensions of cost, speed, and accuracy. We envision that the primary users here would require light-touch modules that can be modified easily to meet acute information needs. As such, we would revisit the information technology (IT) architecture produced in the pilot to standardise module formats, data cleaning and analysis files, and presentation formats. Nano would then host a library of modules designed to be deployed quickly, in any order or combination.

Overall, Nano would continue to strive for flexibility across agencies and subjects, aiming to be suitable for both mainstream and niche development domains and decision-makers. This is an ongoing topic of discussion with GIF and subject to modification later.

2. **Engaging more deeply with the Ministry of Chiefs and Traditional Affairs (MOCTA)**

In this scenario, Nano would leverage support from Chief Mukobela and MOCTA to test simultaneous deployment of Nano across several chiefdoms. Chief Mukobela has expressed strong support for Nano and satisfaction with its outputs. As described in Chapter 3, Chief Mukobela has presented data
generated by Nano to various stakeholders and has taken action on insights delivered by Nano. Following discussions with Chief Mukobela, six other chiefs have expressed interest in Nano.\footnote{Chiefs Chikanta, Chipepo, Cooma, Simwatachela, Mweenda, and Nalubamba have expressed interest in deploying Nano in their chiefdoms.}

IDinsight has presented Nano pilot data to the central MOCTA and has initiated conversations on partnership with MOCTA to expand Nano to multiple chiefdoms. Under this approach, Nano would replicate the general model deployed in Mukobela Chiefdom to several chiefdoms, continuing to serve chiefs as the primary client. As such, data collection would remain a central component of Nano, complemented by some aggregation of existing data.

Should we decide to pursue this channel for Nano scale-up, the next step will be further engagement of MOCTA to determine common objectives and develop a plan for implementation.

3. Pivoting towards District Councils (DCs)

Alternatively, Nano may pivot towards DCs as the primary client. DCs, like chiefs, are positioned for multi-sectoral influence and benefit from close engagement with various ministries. DCs may provide significant scope for social impact, given they have a broader domain of influence and present a good second context in which to test Nano, now that we have established an early proof-of-concept. As would be the case with expansion to more chiefdoms, this model of Nano would suggest simultaneous deployment to multiple DCs as the next step. We expect this expansion would be feasible, as three district councils (Choma, Kalomo, and Kazungula) have expressed interest in having Nano in their districts. Otherwise, the structure and objectives of this scenario would resemble those of the previous scenario.

To move forward with this channel, we would initiate conversations with the district councils that have previously expressed interest to gauge the feasibility of scale-up with these users.

**CONCLUSION**

We are eager to continue to develop Nano innovations and work towards a closer approximation of a full-scale Nano model. We will work with GIF to identify the best path forward, whether that includes continued engagement with GIF or partnership with other organisations. As noted above, the most likely path forward with GIF is a pivot towards data tools for frontline supervisors with an emphasis on building capability to collate, quality-check, and augment data from pre-existing sources.
7. APPENDIX

1. Detailed review of innovations discussed in progress report

CONSULTATION AND STAKEHOLDER BUY-IN

Consultation is the foundation of Nano. To build a system that users love, we need to know the user’s needs, capacity, and preferences, and users must feel heard and valued. Further, consultation facilitates stakeholder buy-in – involved stakeholders are more likely to endorse and promote the product, which is necessary to achieve ideal system functionality, use, and scale. As such, consultation was built into each step of developing Nano.

A first and crucial step was to elicit MOCTA’s feedback on and support of Nano, as Nano could not succeed without MOCTA’s endorsement. In order to achieve this, IDinsight held two meetings with the Permanent Secretary and Directors at MOCTA. MOCTA was initially reluctant to provide support but decided to do so following further consultative discussions. MOCTA has since written a letter of support to IDinsight granting us permission to work with Zambia’s traditional leaders.

Next, we developed a short-list of chiefs, based on feedback from MOCTA and NGOs. Additional consideration of each chief’s previous experience with information systems narrowed the list to two: Mukobela chiefdom and Cooma chiefdom. IDinsight met with both chiefs and ultimately selected Chief Mukobela due to his experience working with IS and interest in using data to improve lives in his chiefdom.

Since the start of our engagement with Chief Mukobela and Round 1 data collection, in a period of eight months, we have met with the chief six times to discuss his data needs, pathways to actions from the data, and the logistics of implementing Nano. We have developed a strong relationship with Chief Mukobela and earned his full endorsement of Nano.

Our early meetings focused on discussing the chief’s data needs and potential actions he could take with this data. This was supplemented by in-depth scoping of the chiefdom, in which we held consultations with multiple stakeholders and visited key facilities in the chiefdom. This allowed us to a) verify some of the information from the chief, b) understand how the chief’s data needs aligned with those of different stakeholders, and c) understand the geography of the chiefdom and potential challenges we would face (e.g. phone service coverage and accessibility).

The chief’s support for Nano has been significant, including instructing the chiefdom’s senior headmen and headwomen to sensitise the community to enumerators – this contributed to only 4.5% of households refusing to be surveyed in Round 1. At the end of Round 1, the chief spent more than a day with us to discuss the data collected and provide feedback on the data presentation.

SAMPLING STRATEGY

Our consultations with the chief indicated that data at the village cluster level was the most decision-relevant. Clusters are groupings of villages around common infrastructures (schools or health
facilities) or other resources (e.g. rivers) and contain sufficient residents to efficiently target programs.\textsuperscript{23} There are eight village clusters in the chiefdom.

To ensure the information collected is representative of the chiefdom and each village cluster, ideally, a list of households or residents would be used to randomly select households/residents to participate in the survey. Given that such a list does not exist, the following approach was used to obtain a random sample of households and, therefore, a representative sample for the chiefdom and village clusters.

Publicly available electronic maps were used to split the chiefdom into roughly 80 non-overlapping smaller areas (henceforth, enumeration areas – EA).\textsuperscript{24} The EAs were constructed around groupings of roofs obtained from a publicly available dataset.\textsuperscript{25} Figure 7 displays the map of the chiefdom (green area); each blue dot represents the location of a roof.

\textit{Figure 7: Mukobela chiefdom boundary and rooftop locations}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure7.png}
\end{figure}

Then, approximately half of the EAs were randomly chosen to be included in the survey. Figure 8 shows an illustrative example of a sampling frame generated in our code. The blue cells are those included in the sample.

\textsuperscript{23} We also discussed collecting data on individual villages, but this would have been less decision relevant, required a large sample size and could have led to confidentiality concerns due to the small size of many villages.

\textsuperscript{24} The borders of the chiefdom were obtained from digitized chiefdom boundaries map.

\textsuperscript{25} GPS locations of roofs in the chiefdom were used to determine placing of EA borders. The GPS location of roofs is a publicly available dataset obtained: HOTOSM Zambia Buildings (OpenStreetMap Export). From The Humanitarian Data Exchange: https://data.humdata.org/dataset/hotosm_zmb_buildings (1 November, 2018).
Enumerators were assigned to specific EAs and required to walk by each household within the EA. The survey tool randomly selected: a) 40% of households for all modules, b) 10% of households for only the demographics module, and c) the remaining 50% of households for GPS tagging only.\textsuperscript{26, 27}

**ENUMERATOR MANAGEMENT**

**Enumerator recruitment process**

Nano requires competent and local data collectors to efficiently and frequently collect high-quality data.\textsuperscript{28} To find such enumerators, we assessed 85 candidates, selecting 20 for Round 1. These potential enumerators were discovered through recommendations from chiefdom residents\textsuperscript{29} and then assessed by the IDinsight team. The assessment included a written test to screen for comprehension, literacy and basic arithmetic, and a face-to-face interview to test for English fluency and interview skills. We initially screened for access to means of transport (bicycle, car or motorbike), but dropped this because ownership was low (14%).

The enumerators, many of whom had never collected data before,\textsuperscript{30} proved to be generally competent. Because of this and the lack of advanced skills needed to enumerate, we do not believe that finding competent enumerators is a risk to Nano at scale.

\textsuperscript{26} There are two exceptions. The full survey was conducted on all households in a small village cluster (Nakalongwe) and all households that had a pregnant woman. The demographics module asked whether there were any pregnant women residing in the household.

\textsuperscript{27} We are currently working on refining the sampling strategy to reflect what would be more feasible at scale.

\textsuperscript{28} Local enumerators reduce data collection cost and timing, as they do not require lodging reimbursement, and there is little lag between data requests and data collection. This significantly diverges from standard data collection, which generally uses full-time enumerators that must be transported to the data collection areas and lodged.

\textsuperscript{29} We specifically received recommendations from headmen, senior headmen, clinic staff, agriculture camp officers, and teachers.

\textsuperscript{30} Approximately 10 out of 85 enumerators interviewed reported previous data collection experience. Of those selected, only 3 out of 20 reported previous experience.
Training

Selected candidates were invited to attend a one-day training held at a central location in the chiefdom. Enumerators were trained on survey protocols, mobile data collection, and use of maps to locate survey areas.

The enumerators who performed best during training and on previous assessments were selected as supervisors. Data collectors were divided into teams of four (three enumerators and one supervisor). The one-day training was followed by half day of training and piloting with each team in their operating area.

Our training approach provided important lessons for scale-up. First, 1.5 days of training are enough to prepare the most capable enumerators, while the others likely need an extra day of piloting with supervision. Second, most enumerators initially struggled with operating smartphones but learned quickly, meaning that we do not see smartphone experience as a predictor of enumerator quality. Third, training in a central location is logistically difficult, requiring enumerators to travel up to five hours, but likely efficient since it reduces total training time and overall field costs.