Collecting data with smart forms on mobile phones

Yaw Anokwa, Carl Hartung
http://nafundi.com

Hi, I'm Yaw and that's Carl.

In June, we finished our PhDs in Computer Science at UW with Gaetano Borreillo. As part of our dissertations, we created Open Data Kit, a free and open source platform for collecting data with mobile phones.

In this talk, I'm going to focus on our motivations in building ODK, I'll show what some of the tools are, and I'll also give you a few examples of ODK use at scale. I'll end with a quick look at some of the work Carl and I are doing at Nafundi.

I've got two goals today. I want to convince you that ODK is fantastic for anyone who wants to collect data accurately, quickly, offline and at scale. I also want to convince you that technology can help solve important problems -- even in the most challenging environments.

I want to make this informal, so interrupt to ask questions at any point, and I'll leave some room at the end for even more questions.

This picture is a village we worked in on the edge of Lake Victoria in Eastern Uganda, and I think it's an example of the places where organizations like PATH need to collect data.

You can see in a place like this there is almost no infrastructure to enable communication with the rest of the world. There's no electricity, there's actually no wired infrastructure of any kind.

For those of you who have traveled to a developing country, we didn't see any Coca Cola for sale, which is the one thing you can usually find anywhere in the world. So that's how you know when you've really left the map.

So how do you collect data in a place like this?

You probably use paper, but that can be very limiting.

Hard to be accurate. You get plenty of handwriting and transcription mistakes.

Hard to have complex questions. You want to ask different questions based on previous answers, but that's hard to train for.

Hard to capture things like location, multimedia. Adding a GPS and camera adds more points of failure. How will you attach it to the paper record to you can use it in the analysis?
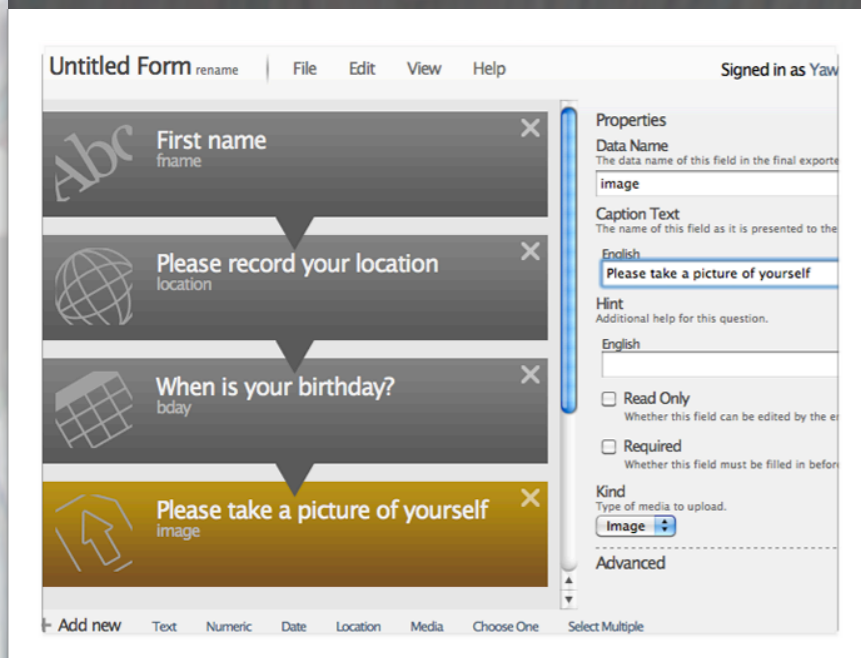
Hard to get the data in quickly. It's really important to have data as it comes in so you know when there are problems. You don't want to wait a few months after the study is done to know something went wrong.

We saw these problems a lot in the field, and so that was the driving motivation behind our work -- addressing these limitations through technology. That's why we created ODK.
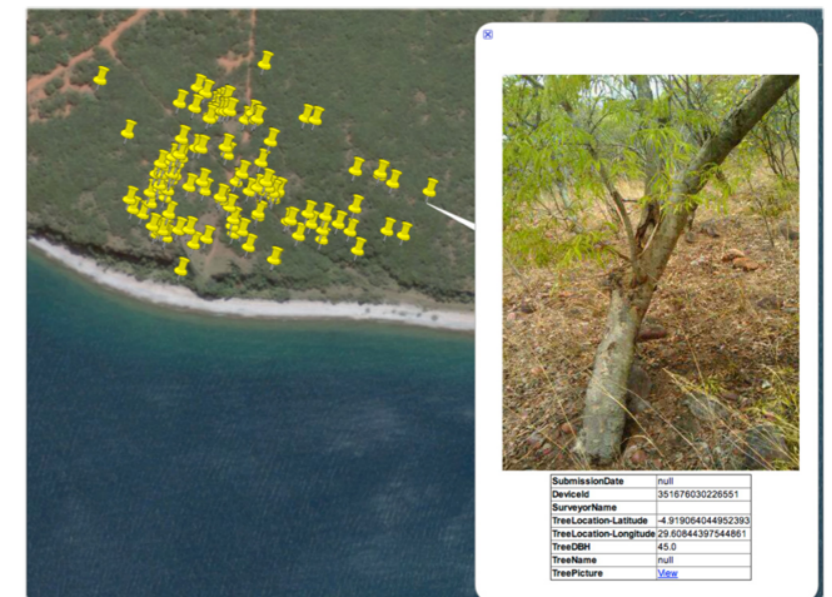
ODK uses phones and servers to digitize data collection.

1. Build form

2. Collect data

3. Aggregate results

ODK is a free and open-source set of tools which use smartphones and cloud servers to digitize data collection.

It's great for field staff who need to collect data accurately, quickly, offline and at scale. ODK can be used to implement large socio-economic surveys, collect geo-tagged vaccine data, or even triage patients with complex decision trees and embedded videos.

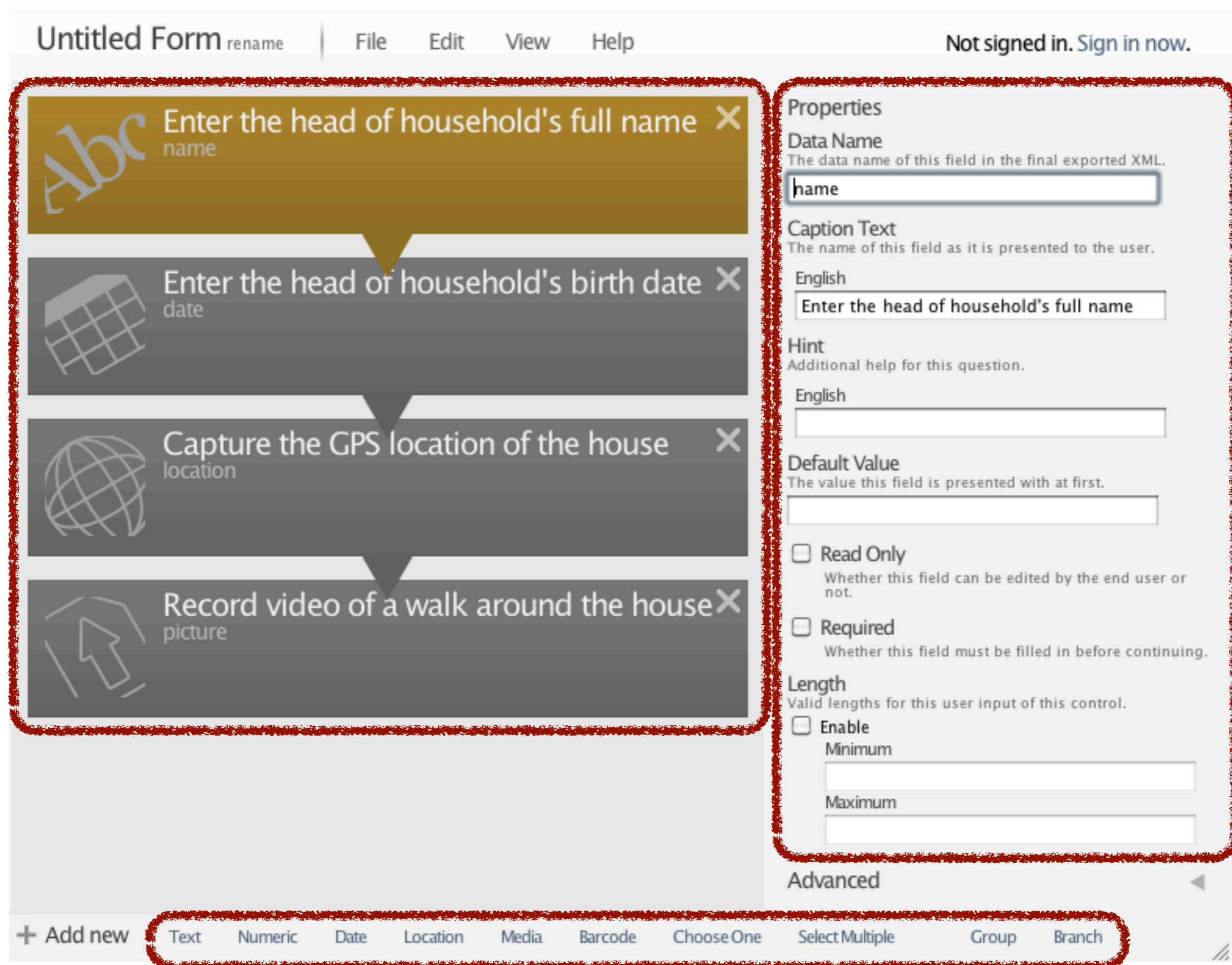ODK provides an out-of-the-box solution for users to:

1. Build a data collection form or survey;

2. Collect the data on a mobile device and send it to a server; and

3. Aggregate the collected data on a server and extract it in useful formats.

If there is one thing to take away from the talk this is it. If you were planning on taking a nap during the talk, you may now do so guilt-free.

Anyway, so what are these tools that you use to Build, Collect and Aggregate?

ODK Build is an web application where you drag and drop prompts to create forms. It runs in the browser but can also be used offline. It looks like this.
*

To design a form, you drag and drop each prompt the user will interact with from this button pane
*

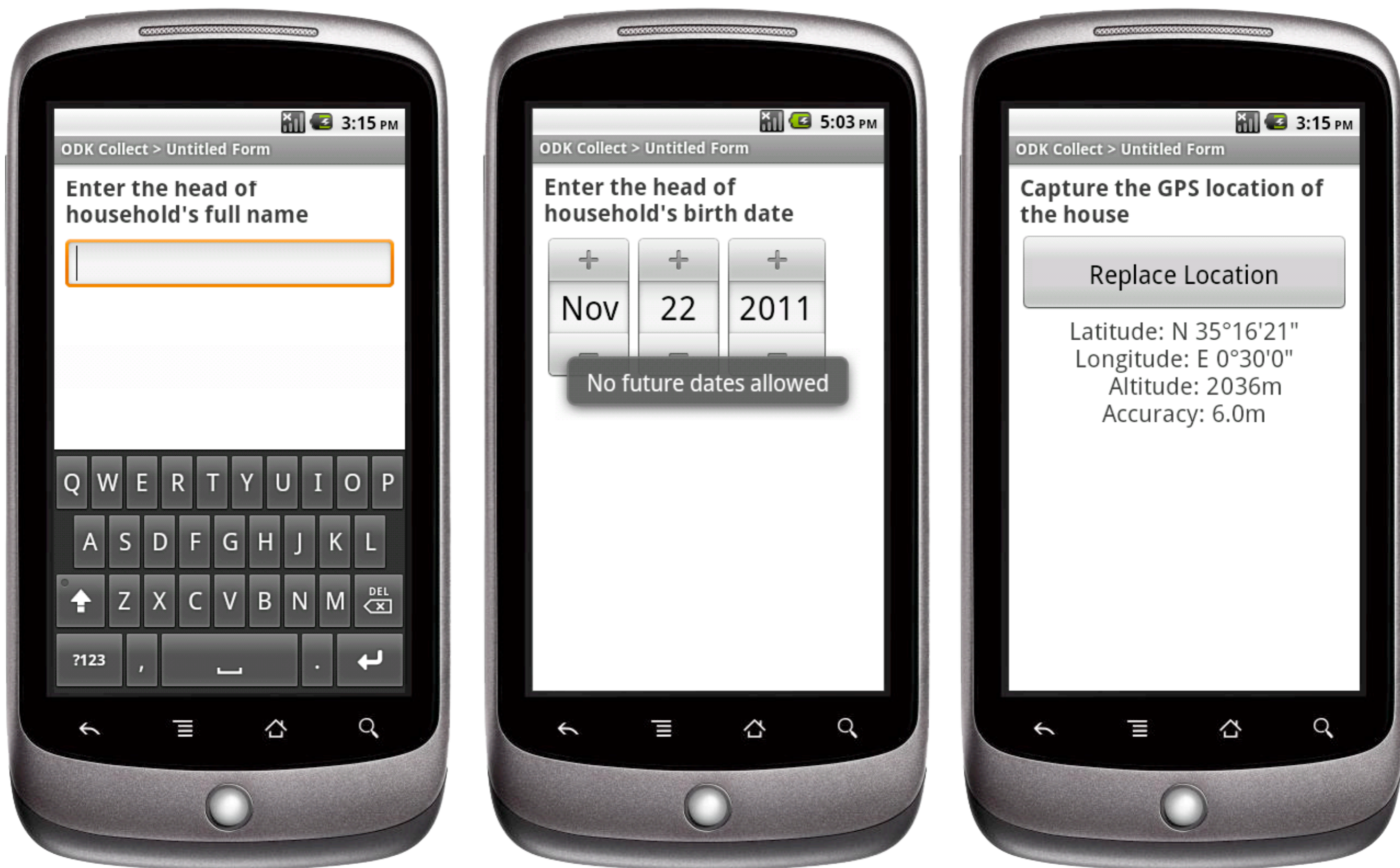to the canvas. Each prompt has a set of properties which users can edit here.
*

The prompts are pretty powerful. For example, you can have multiple languages for each prompt, make prompts relevant based on previous answers, or make prompts loop some number of times.

When you are done, thats what goes to the phone.

Collect: Display prompts for data collection and delivery.

ODK Collect is an app that display prompts to collect and deliver different types of data.

Collect runs on the Android operating system so you can use it on phones, tablets and netbooks. It runs great offline and looks like this.
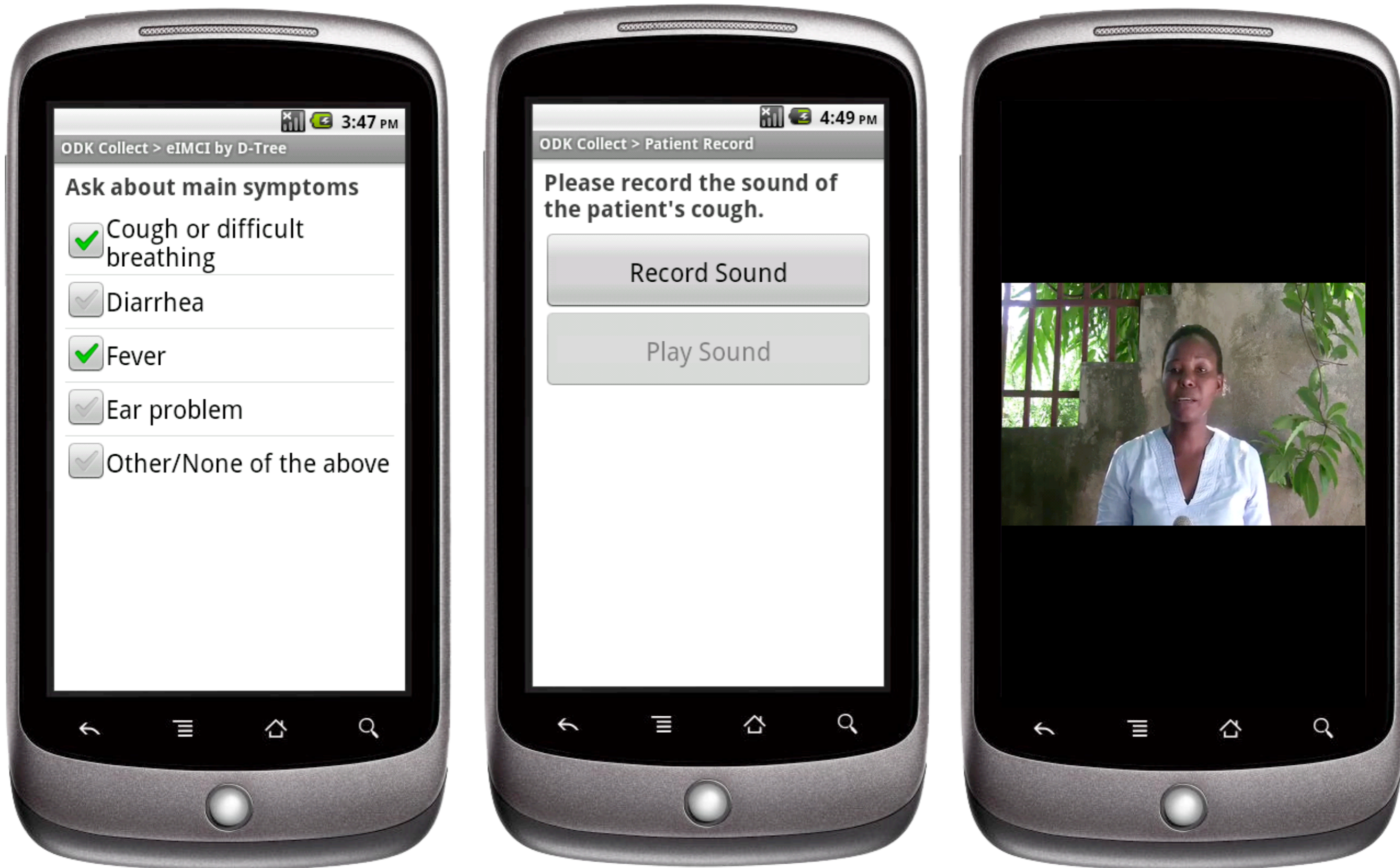*
You can collect text, numbers
*
dates
*
you can put constraints on all prompts. So for example, birth dates can't be in the future.
*
You can also capture GPS location using one click.

The forms themselves are really powerful. So here is an example of IMCI in swahili.

IMCI is the integrated management of childhood illnesses. It's basically a triage protocol for children under five.

My Swahili is weak, so let's switch to English. With ODK Collect you can do this on the fly.

And if I enter that the child is coughing and has a fever, in the next screen, I can record the sound of the cough and then I can show the mother a video about how to treat fever.

So there is this collection of data, but also a delivery of information using complex logic. Any data I gather can be stored offline and then be sent off to a server.

So that's ODK Collect. And where does the data go?

Aggregate hosts the submitted data and provides interfaces such as

spreadsheets, maps, and queries. We don't run one big server, you download an installer, and it configures one just for your organization.

Aggregate: Stores or forwards data to external systems.

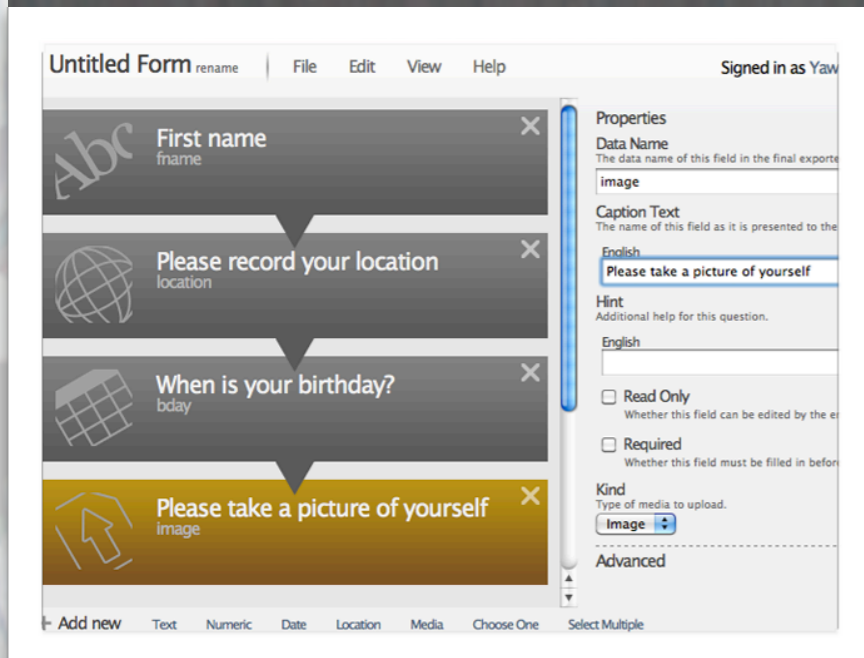| DeviceId	351676030226627
SurveyorName	Shadrack
TreeLocation-Latitude	-4.9192410707473755
TreeLocation-Longitude	29.60762321949005
TreeDBH	57.0
TreeName	Myombo
TreePicture	View

Finally, Aggregate can either store or stream your data to other systems, like Google Earth.

In this example, forestry workers with the Jane Goodall Institute in Tanzania, submitted data from Collect to Aggregate and then exported to Google Earth

Managers could then click on each yellow point and get the data that was submitted.

1. Build form

2. Collect data

3. Aggregate results

So that's what ODK is. A platform that will help you build a form, collect data, and aggregate some results.

For building forms, you can use Excel. For aggregate data, we have a desktop app for that too. You pick and choose which work best for your use case.

Now that's nice and all, but when people find out that this came out of a research project, they start asking if the software actually works.

Well, yeah! Let me give you some examples.

Kiva uses ODK to get borrower information.

Ruth

Eldoret, Kenya   Agriculture | Farming

LOAN OVERVIEW   REPAYMENT SCHEDULE   LENDER COMMENTS

PAYING BACK

This loan has be
fully funded!

A loan of $375 helped Rut
buy irrigation pipes.

16% repaid

Find a Loan

Repayment Term: 18 months (more
Repayment Schedule: Irregularly
Pre-Disbursed: Aug 30, 2012
Listed: Sep 7, 2012
Currency Exchange Loss: Possible
Default Protection: Not Covered

Carter Center uses ODK for monitoring elections.

Carbon For Water collected over 1,000,000 forms with 4,000 ODK-powered phones in 6 weeks.

Carbon For Water wanted to distributes LifeStraw Family water treatment units to 4 million people. And as part of that distribution they want to take images and location of each barcoded unit, along with some socio-economic data.

And so they used 4,000 ODK-enabled phones to collect over a million forms in six weeks. Which is incredible.

They've found that ODK enables real-time collection of image, GPS and survey data at scale.

Actually, it's kind of scary because they told us about the deployment after it happened. They just took the software and ran with it. My jaw dropped when I saw those numbers. Our largest deployment before that was a couple of hundred phones.

AMPATH's health workers have used ODK to counsel and test over 650,000 people for HIV.

AMPATH is one of the largest HIV treatment programs in Africa. They have about 2 million people in their catchment area and are treating 130k HIV positive people.

As part of their care, they have 300 community health workers who go house to house and do home-based counseling and testing for HIV. These guys are equipped with ODK-powered smartphones.

This image shows one of those workers using the barcode scanner on the phone to identify a patient.

We helped these guys pilot the system for two weeks, then they took over. The project started in early 2010 and as of a few months they had visited 650k individuals.

Users found the system easy to use and thought it allowed them to collect higher quality data.

Supervisors saw higher cost savings over their previous PDA-based system as well as over pen and paper systems.

Electronic collection also facilitated earlier reporting, which means more people get on treatment much faster.

I want to use RHVouchers as a last example.

A reproductive health voucher is a card that entitles a poor, pregnant woman to subsidized maternity care. These vouchers are used around the world, but do they actually help improve the quality of care?

The Population Council and Marie Stopes are evaluating voucher programs in Kenya, Uganda, Tanzania, Bangladesh and Cambodia to find out.

As one could expect, these evaluations are an intensive process of collecting, organizing, cleaning, and analyzing data.

The process happens over largely rural areas with vast distances between households and health service providers and these guys put together a video about how they use ODK.

- **Nafundi** - A company started by the guys who invented ODK. If you need professional support or require changes to ODK tools, why not try the only company the ODK team recommends?
- **Dimagi** - Provides support for core and custom development on the Open Data Kit platform. They work closely with team that founded ODK.
- **Mindflow Associates** - Have experience customizing OpenMRS, ODK data collection solutions, building Android phone based applications, utilizing JavaRosa-based technologies amongst many others.
- **Afrisis** - Has developed and deployed end to end (mobile, touch-screen kiosks, SMS, desktop, web) software solutions across all provinces of Mozambique. They are quite familiar with JavaRosa-based technology.
- **Mega Six Solutions** - Based in Kenya, and specializing in Research and IT domains, we have expertise in data management processes. Using our Android development team, Mega Six will help you customize ODK for desired data collection and aggregation.
- **Group Complete** - Will design and implement survey collection forms for users of Group Complete (and ODK-based tool) for free.
- **ULevel** - A company based in Brazil doing customizations on Collect for Android.
- **WebFirst** - Well versed in the development of data collection systems like JavaRosa and ODK in low-resource and standard environments.
- **EarlySail** - Have customized ODK Collect for a few pilot projects as well as assisted in form development. They will soon be porting ODK to BlackBerry.
- **Seeing Swans** - A company focusing on data collection and complex analytics. They have customized

ODK is more than some code and a mailing list. The project is a community and we've spent a lot of time investing and growing in that community and it's paid off.

There are companies that build entire products around ODK. There are also a lot of consulting companies around the platform –– a lot of them in developing countries.

NAFUNDI

And now that Carl and I have graduated, we've jumped into the fray with Nafundi.

Nafundi is a word that has its roots in Swahili. Loosely translated, it means "with a craftsman," and that idea describes the skilled and collaborative we when building software.

Nafundi is a software consulting company. We write code for hire.

A large part of what we do is provide services around ODK. When we where at UW, we were focused on computer science research and that's not always immediately useful to our users. Now that we are at Nafundi, we can focus on more what implementers need day to day -- everything from helping you make your survey, building extensions to ODK, helping you train and scale, etc.

We also build custom software that will work well in challenging environments -- places with unreliable connectivity, novice users, etc. It turns out our experience on ODK helps us build software that just works.
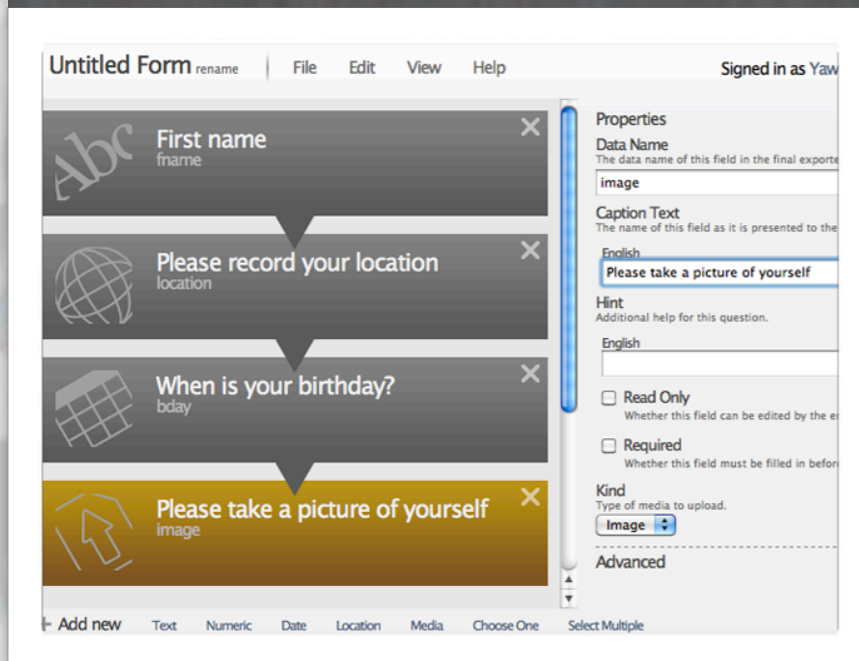
For example, we are working with IHME to build a desktop program that makes it easy for in-country researchers to automatically figure out the cause of death from verbal autopsy data.

So again, it's custom software development, but for places where normal software shops don't have any experience.

Collecting data accurately and quickly on paper is difficult. ODK uses phones and servers to digitize data collection.
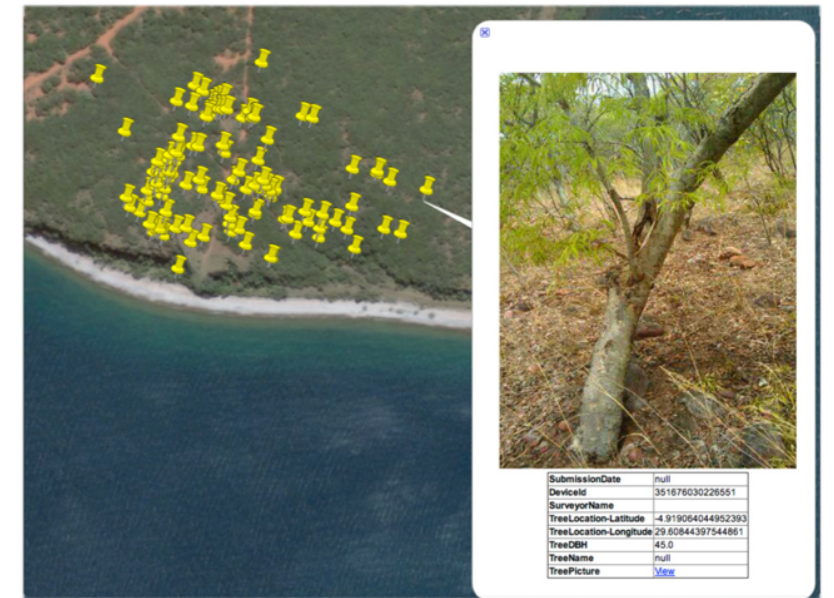

| 1. Build form | 2. Collect data | 3. Aggregate results |

http://opendatakit.org (@opendatakit)
http://nafundi.com (@nafundi)
http://anokwa.com (@yanokwa)

Using paper to collect data is difficult and inefficient. ODK is a free and open-source set of tools which use smartphones and cloud servers to digitize data collection.

ODK provides an out-of-the-box solution for users to:

1. Build a data collection form or survey;

2. Collect the data on a mobile device and send it to a server; and

3. Aggregate the collected data on a server and extract it in useful formats.

Are there are any questions I could answer?