Replacing paper forms with smart forms on phones and tablets

Dr. Yaw Anokwa
http://nafundi.com

Hi, I'm Yaw. I'm CEO of Nafundi.

Nafundi's mission is to create technologies that improve the lives of the underserved. These days, we are best known for Open Data Kit, a project that I helped to create as part of my Ph.D. work at UW.

Open Data Kit (or ODK), is a free and open source platform for replacing paper forms with smart forms on phones and tablets.

In this talk, I'll show what some of the ODK tools are and give you a few examples of their use at scale.

I've got one goal today.

I want to convince you that ODK is fantastic for anyone who wants to collect data accurately, quickly, offline and at scale.

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.

Collecting data in low-income regions is difficult.

This a picture I took at the edge of Lake Victoria in Eastern Uganda.

I think it's a great example of the places where the Foundation's grantees 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 are no computers, and when it rains, there are no roads.

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

Paper is common, but can be very limiting.

You probably use paper. Here is an example of a nicely organized records room at a hospital in Kenya.

Even in this best case scenario, there are problems because paper is very limiting.

It's impossible to search. It'll take you a few hours just to tell me how many of these patients have malaria.

It's impossible to visualize. It'd be nice to have the GPS locations for these records so you can see a map of disease spread.
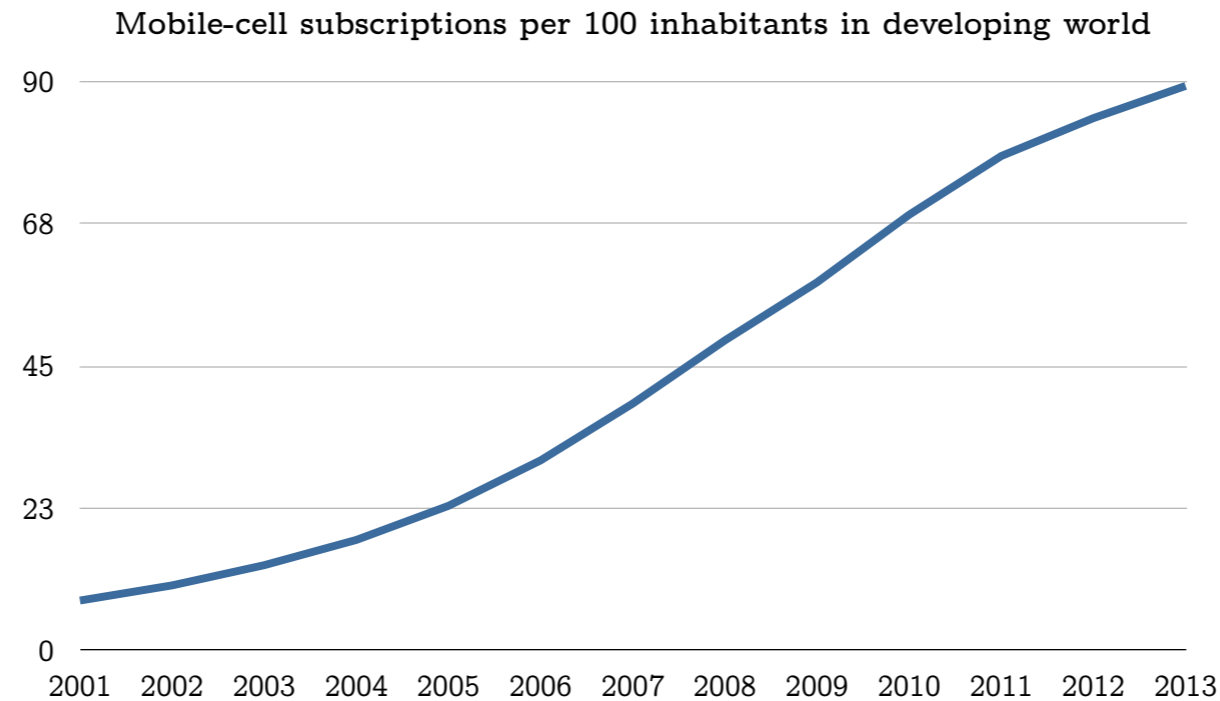
It's impossible to validate. If someone enters an age of 40 instead of 4 into the paper form, your data will be very dirty. You don't want to wait a few months after the enumerators have left the field to realize that the data was all wrong.

Our research group saw these problems a lot when we did our fieldwork, and so that was the driving motivation behind our work -- addressing these limitations of paper.

We chose to do this using smartphones and "cloud" servers.

Let me explain why we thought this was a good idea.

Growth of cellular technology enables innovation

Mobile-cell subscriptions per 100 inhabitants in developing world

90
68
45
23
0

2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013

In 2008 when we started ODK, there was growing availability of cell technology in developing countries.

This data is from ITU. It's the cell subscriptions per 100 inhabitants in the developing world. It's close to 90% penetration. This doesn't mean everyone has a cellphone. Richer folks have multiple subscriptions and poor folks have none.

But the growth is there. And mobile broadband is growing as well. It's now at 20% in developing countries. So people are getting online on their phones. And these aren't always basic phones.

In Kenya, the price of the basic smartphone, something like an original iPhone, is about $50–$60. And it sells well at that price.

The punchline here is that when we started ODK in 2008, there was growing Internet access and smartphone usage in these developing areas. And we predicted it would keep growing exponentially.

So on that infrastructure of cell towers and that infrastructure of familiarity, we thought we could build systems that address the limitations of paper I talked about in the last slide.

That's what inspired us to replace paper forms with smart forms on phones and tablets. That's what inspired ODK.

So what is ODK?

ODK uses phones and tablets 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 phones and tablets 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 agricultural data, or even triage patients with complex decision trees and embedded videos.

ODK provides a turnkey 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. You can stop paying attention now and do so guilt-free.

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

Build: Drag and drop prompts for form creation.
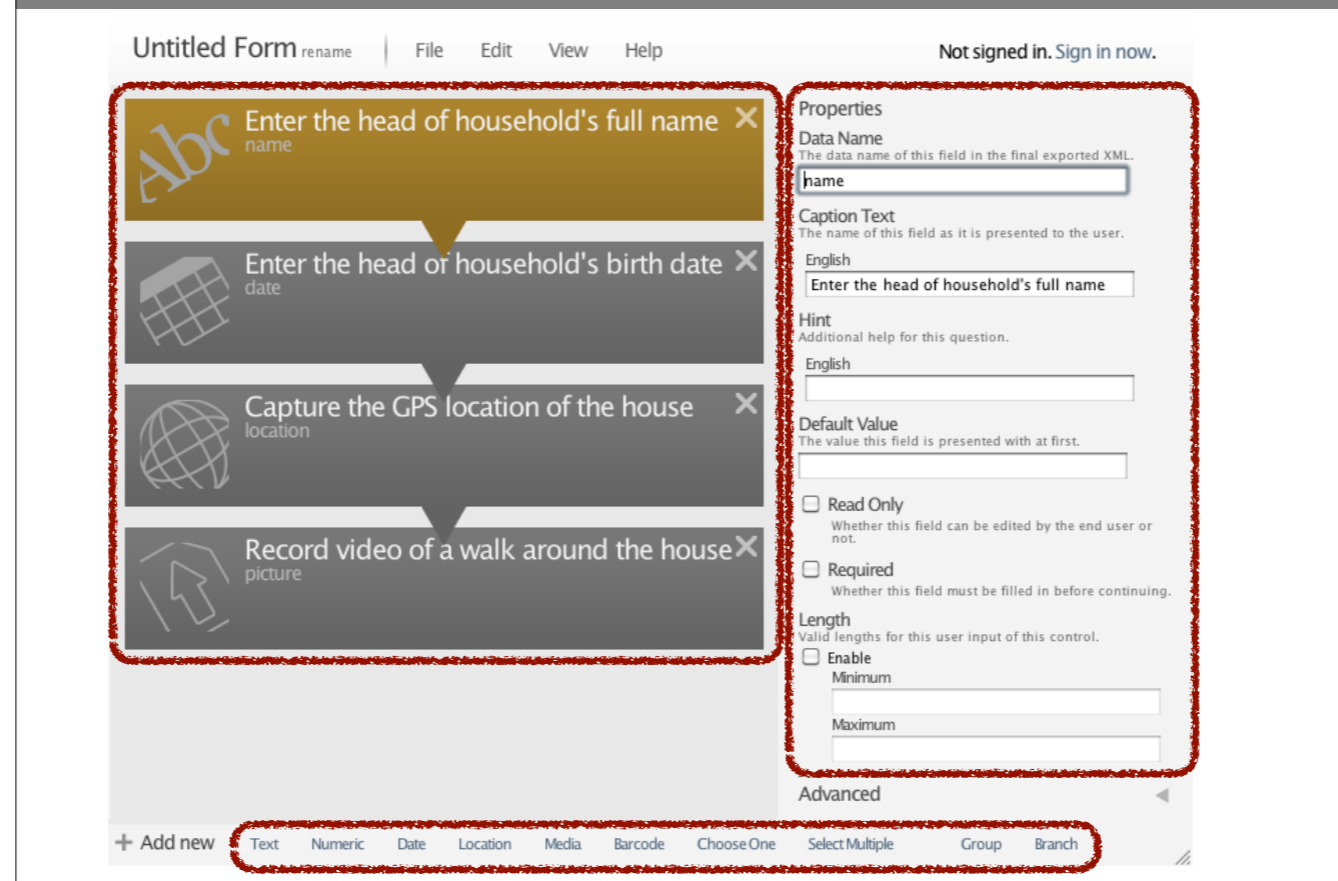
ODK Build is an web application where you drag and drop prompts to create forms.

It runs in the browser but can also be used offline.

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 appear or disappear based on previous answers, or make prompts loop some number of times.

When you are done designing the form, it goes to the phone.

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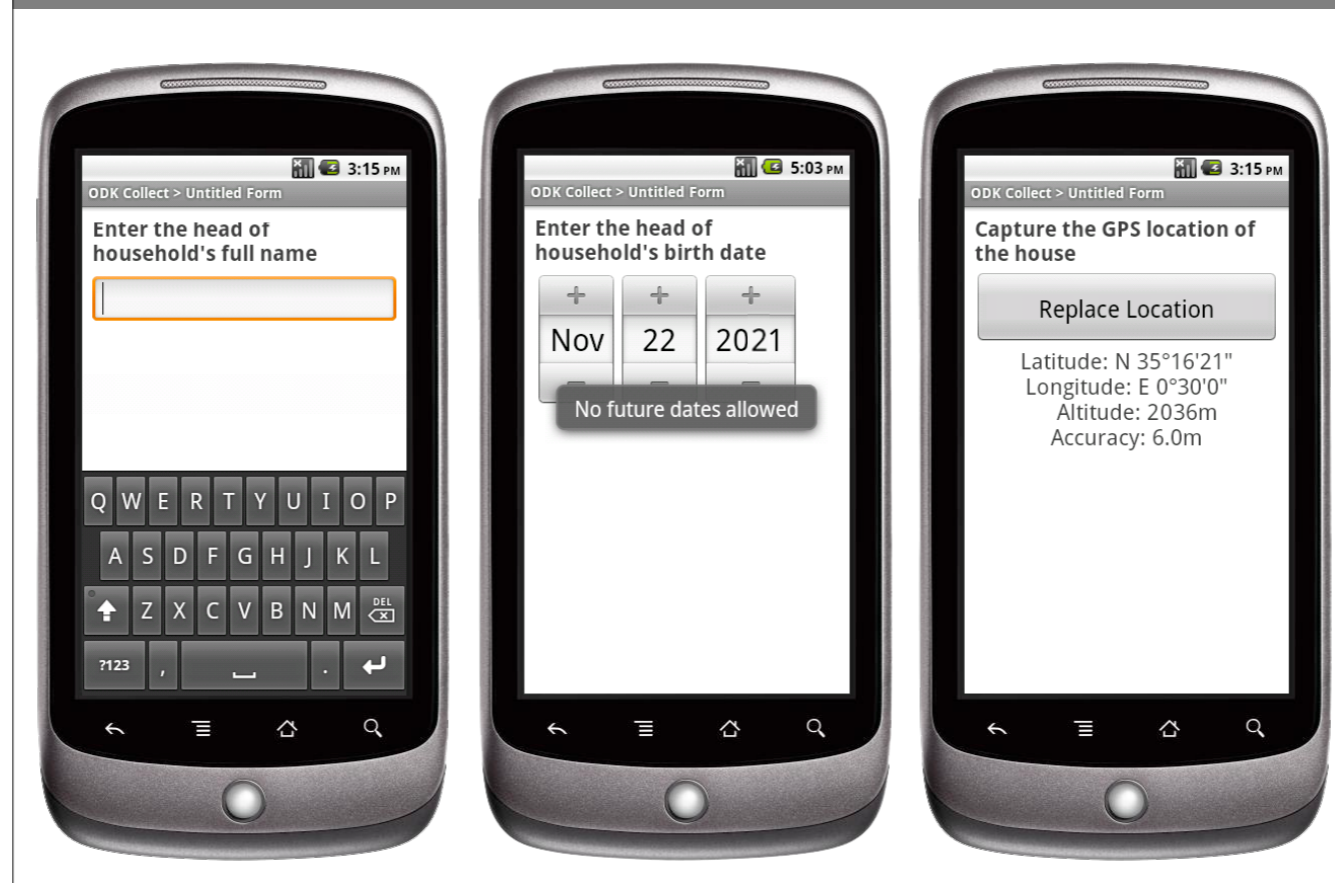
ODK Collect is an app that display prompts to collect and deliver different types of data.

Collect runs on the Android operating system so you can use it on phones, tablets and netbooks.

It runs great offline you can collect text, numbers, dates, and 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.

We even use the camera on the phone to scan barcodes.

Collect: Multiple languages and multimedia ability.
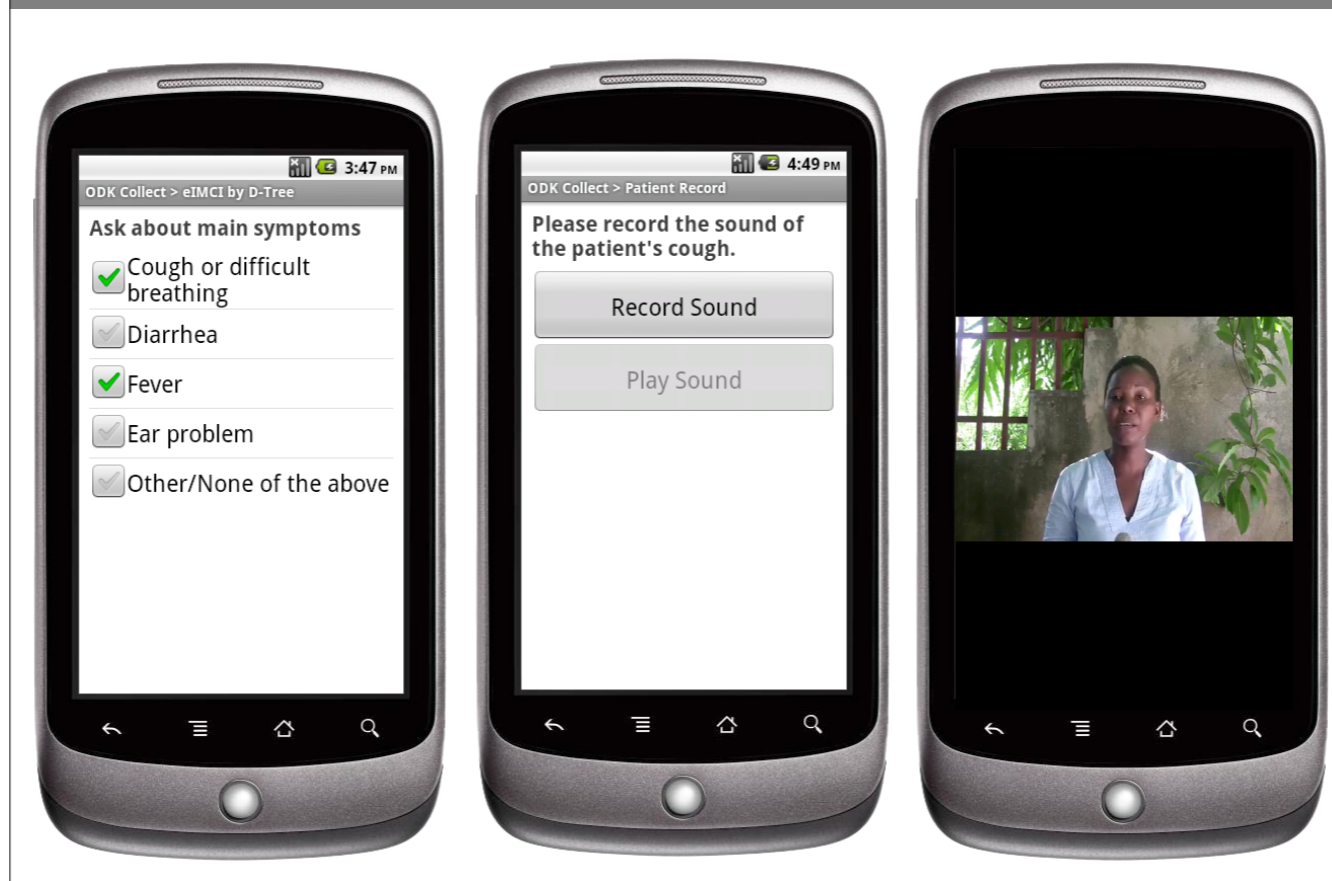
The forms themselves are really powerful. So here is an example of IMCI in Swahili.

IMCI is the integrated management of childhood illnesses. It's basically a triage protocol for sick 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?

ODK 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. You can choose to install locally and in the cloud.

This way you control where your data lives.

Aggregate: Stores or forwards data to external systems.
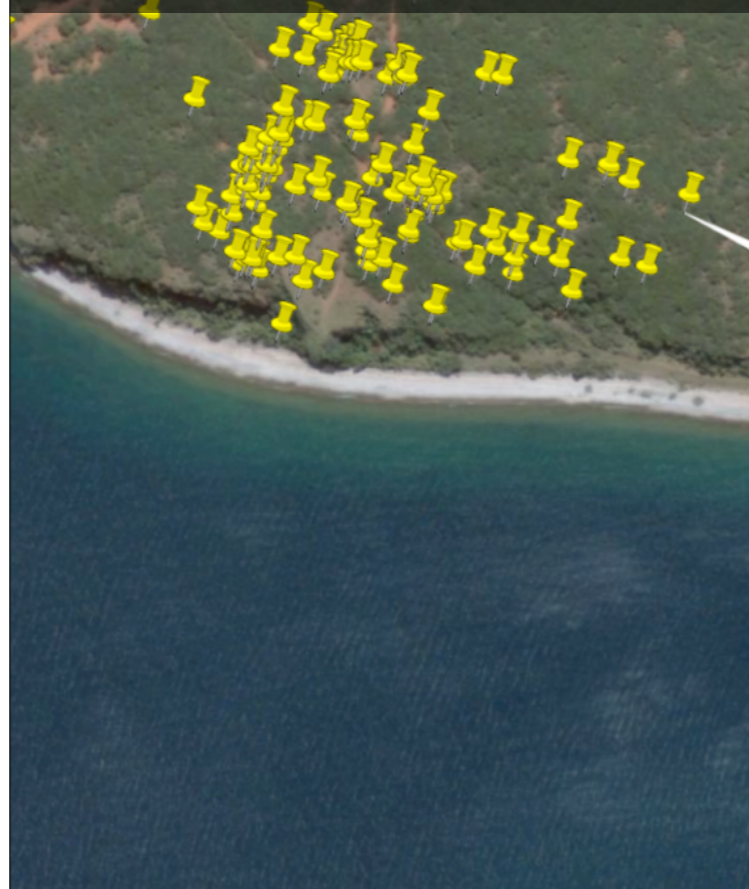
DeviceId	351676030226627
SurveyorName	Shadrack
TreeLocation-Latitude	-4.9192410707473755
TreeLocation-Longitude	29.60762321949005
TreeDBH	57.0
TreeName	Myombo
TreePicture	View

ODK Aggregate can either store or stream your data to other systems, like Google Earth or aWhere or ESRI maps.

In this example, forestry workers with the Jane Goodall Institute in Tanzania, submitted data from Collect to Aggregate and then exported to Google Earth. This particular campaign took a few hours to build forms, train enumerators, and get this data.

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.

While we provide an end to end solution, everything is designed to be modular. You can pick and choose which components you use. That's why it is called a kit.

For example, Modi Labs at the University of Columbia has some really great ODK-compatible tools that you can use instead of the ones I just showed you.

We'll start with building forms.

XLSForm: Build and share forms with Excel

type	name	label	hint	constraint
select_one occupation	a7	7. What is your occupation?	[Classify respondent on occupation. If other, list occupation stated] Researchers: please make sure not all selected respondents belong to the same occupational category (if at all possible).	
text	a7_working_class	Working class? Please specify.		selected(${a7},'8')
text	a7_other	Other? Please specify.		selected(${a7},'66')
end group	a7_group			
begin group	a8_group			
select_one educational_level	a8	8. Please could you tell me your own education level?	Researchers: please make sure not all selected respondents belong to the same educational level (if at all possible).	
integer	a8_primary	Last year of primary school?		selected(${a8},'2') and . >= 1899 and
integer	a8_secondary	Last year of secondary school?		selected(${a8},'4')and . >= 1899 and
text	a8_details	Further studies? Please specify.		selected(${a8},'6')
end group	a8_group			
select_one yes_no	a9	9. Do you have access to a radio within your home?		
begin group	a10_group			
select_one listen_radio	a10	10. How often do you listen to the radio?		
text	a10_other	Other? Please specify.		selected(${a10},'66')
end group	a10_group			

survey / choices / settings / +

So for form design. There is XLSForm, which you can use to build forms with Excel.

So it's like ODK Build, but instead of drag and drop, you just enter them into a spreadsheet.

Each row is a prompt the user sees.

It's not as easy to use as ODK Build for beginners, but for long forms and for users familiar with Excel, it's a great alternative.

So that's XLSForm.

enketo: Collect data offline in the browser.

Makubaliano iliyo toka kwenyi utafiti inayo husu Idadi Ofisini S.E.L.F. * ❷

○ Ok. Tafadhali kuendelea.

A. Utafiti

1. Namba ya Mulengwa *

2. Siku ya kujibu ku maswali ama maulizo

Utafiti kuhusu wakaaji

5. Una umri gani? * ❷

Mwisho wa utafiti

Namba ya mutafiti *

There is also enketo, which you can use to collect data in a web browser.

So it's like ODK Collect, but you can use it on your iPhone with Safari or your laptop with Internet Explorer.

It works great offline, it has all the branching and validation that ODK Collect has, and you can even change languages on the fly.

For example from English to Swahili.

It doesn't have all the multimedia features that ODK Collect has, but if you, for example, have data entry clerks who want to enter a backlog of paper forms while you use phones to collect new data, then it's perfect.

With enketo, all the data will go to the same place.

Formhub: Aggregate and visualize results.
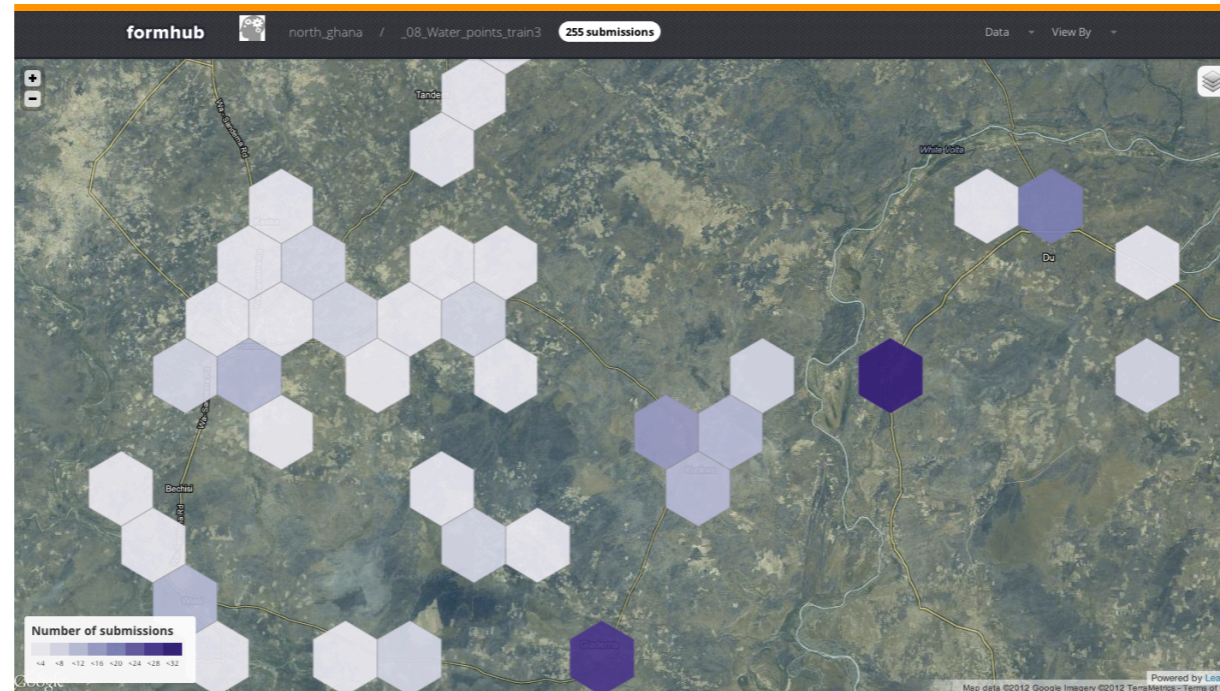
View the density of submissions with auto-generated hex-bins

The last tool from Columbia I want to show is called Formhub. It's a really nice alternative to our ODK Aggregate server.

It has a totally different look and feel. For example, it has these great visualizations like automatically showing you where your data was submitted from.

The key difference here is that there is only one FormHub website. So if for example the MOH wants you to run the server in country, or you want to control where the data goes, you might be better off using something like ODK Aggregate.

1. Build form **2. Collect data** **3. Aggregate results**
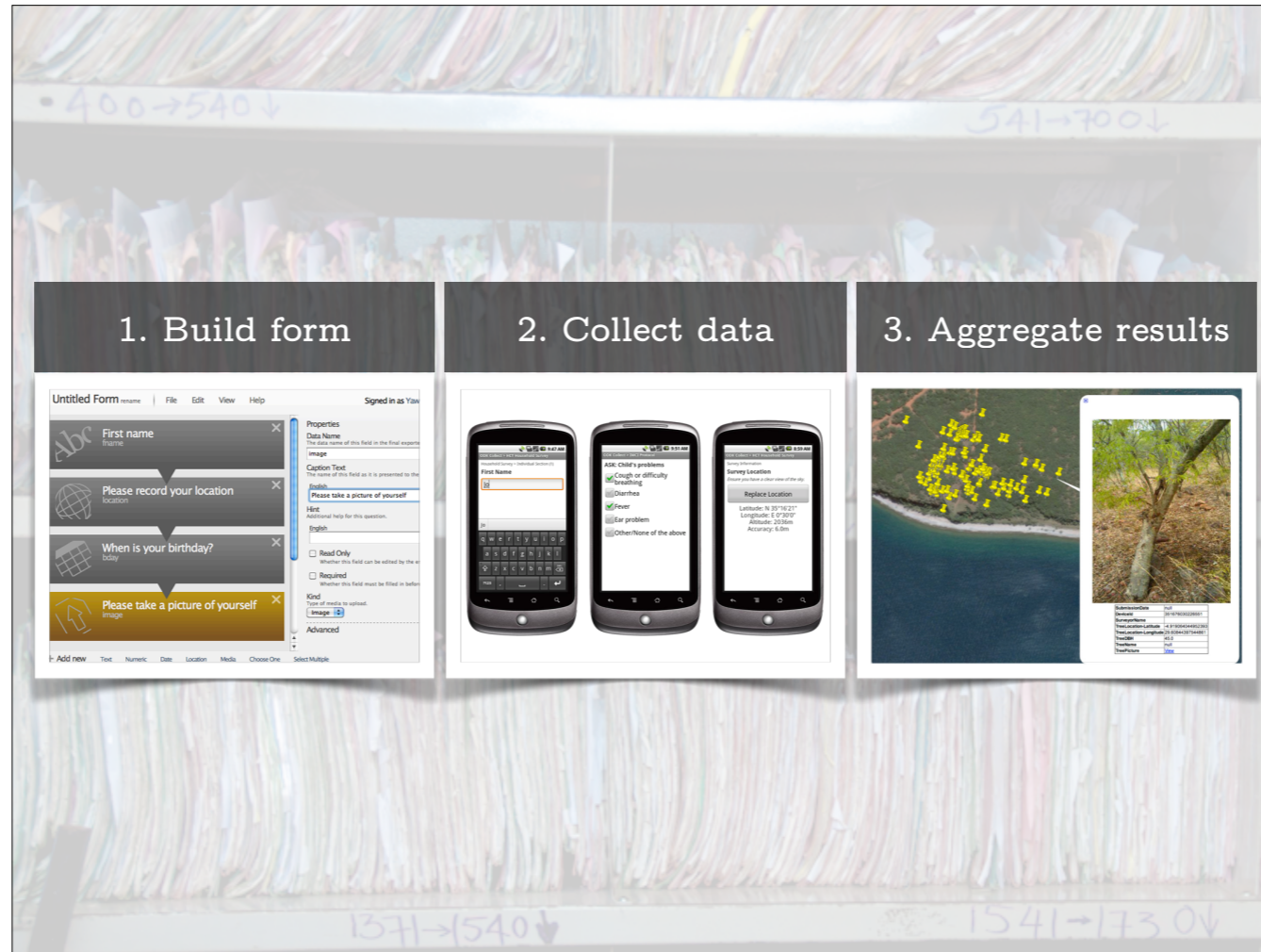
So hopefully that gives you a sense of what we mean by a kit.

If you are an Excel expert, use XLSForm.
If you need to collect data in a browser, use enketo.
If you don't like ODK Aggregate, try FormHub.

They are all free. They all work together seamlessly together. You can pick and choose individual tools and you are never locked in.

Now that's nice and all, but when people find out ODK came out of a research project, and that everything is free and open source, they start asking if the software actually works.

Well, yeah! ODK is used by tens of thousands of people who have collected millions of forms.

Let me give you six examples of some of the big projects.

Jane Goodall Institute uses ODK for forest mapping.

The Jane Goodall Institute uses ODK for forest mapping in Tanzania and Uganda.

In these communities, trees are chopped down to be sold or used as firewood. This is a problem because if there are no trees, then there are no chimps.

With ODK, JGI can empower communities themselves to document the health of their forests.

And because the process is transparent and efficient and near-real-time, the community can then show the data as the proof they need to sell carbon credits and get income.

They no longer need to cut down trees.

Kiva allows you and I to give a small loan to an entrepreneur anywhere in the world. They do this through a network of micro finance groups that administer these loans.

The key to convincing people to loan money is providing lots of information about the entrepreneur and updates about whatever it is their doing.

Some of Kiva's partners use ODK to make the process of gathering all this data easier.

With the phone, they can take pictures, track loan status and get updates to the Kiva website quickly and accurately.

Carter Center uses ODK for monitoring elections.

The Carter Center uses ODK for their election monitoring. They've used it in Egypt, Nigeria, Nepal, etc.

With ODK, observations from polling stations across a country to be transmitted to headquarters immediately, allowing a richer picture of an election to emerge in real time.

And that real time is key to being able to determine quicker if an election is credible.

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

Carbon For Water wanted to distribute 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.

To do this, 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.

Interesting side note, one of this project's members is an astronaut. And apparently they have the internet in space now, so he was using ODK on the space station. Pretty impressive.

AMPATH's health workers have used ODK to counsel and test over 775,000 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. The phone's logic helps them collect socio-economic data and also decide when to do tests for HIV or TB. All the data gets sent to their OpenMRS medical record system.

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

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

CGIAR uses ODK for baseline data collection for agriculture and climate research in Asia and Africa.

CGIAR uses ODK for agriculture and climate research in Asia and Africa.

I just found out about this a few days ago. That's what's great about ODK. Folks are empowered to use it without us getting involved.

"With paper forms, we would have spotted errors weeks or months later – far too late to resolve it. These types of issue are inevitable in a project of this scale. By using the Open Data Kit, we were able to resolve many of them during data collection. This will make analyzing and archiving the data much easier."

http://vimeo.com/38123850

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.

http://vimeo.com/38123850

Growing number of companies around the platform.

- **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, a website, 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 now companies that build entire products around ODK and that's great for spreading the word about the tools.

There are also a lot of consulting companies around the platform -- many of them in developing countries.

A few years ago, I jumped into the fray with Nafundi, a company I started with Carl Hartung, one of my co-founders on ODK.

Nafundi's mission is to create technologies that improve the lives of the underserved.

Our current focus is on designing, building, deploying, and supporting mobile data collection software.

Our expertise is working in really challenging environments. We have software running everywhere from the mountains of Afghanistan to the forests of the Amazon.

We even have a few users in Tacoma.

Let me give you two examples of the mobile data collection work we've done for the Foundation.

Nafundi helped GPEI use ODK on over 8,000 phones in Nigeria for documenting polio vaccinations.

Nafundi helped Global Polio Eradication Initiative use ODK on over 8,000 phones in Nigeria for documenting polio vaccinations.

GPEI s a project with WHO, CDC, and UNICEF. eHealth Africa led much of the mobile data collection effort with support from the Foundation.

They've deployed 8,000 phones are hoping to scale to 12,000. In their last campaign, they collected 5 million GPS tracks in 5 days.

This was a massive effort and Nafundi built customizations (like admin mode, bulk config, and large form optimizations) for eHealth Africa to make their lives easier.

These customizations were made open source and contributed to the ODK community.

Nafundi helped WFP's Purchase for Progress collect data in Rwanda, Ethiopia, Mozambique, Kenya, and El Salvador.

Nafundi helped WFP's Purchase for Progress collect data in Rwanda, Ethiopia, Mozambique, Kenya, and El Salvador.

P4P has always collected data on paper, but there've been problems with accuracy and timeliness of the data.

We were brought on by the Foundation a few weeks before their paper deployment to move to mobile data collection in Rwanda.

We designed forms, setup servers, acquired tablets, did the training, and they soon had the data flowing in. It took days from when we arrived in country to start seeing data from the field.

That's how we work. We are very agile. We deploy rapidly, then work with grantees to build capacity.

We were hands on in Rwanda. In Ethiopia and Mozambique, we were there just to monitor. Kenya and El Salvador have been run by WFP. We try to just lay the ground work and hand off quickly.

So that's the sort of work we do at Nafundi.

To wrap up,

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

1. Build form   2. Collect data   3. Aggregate results
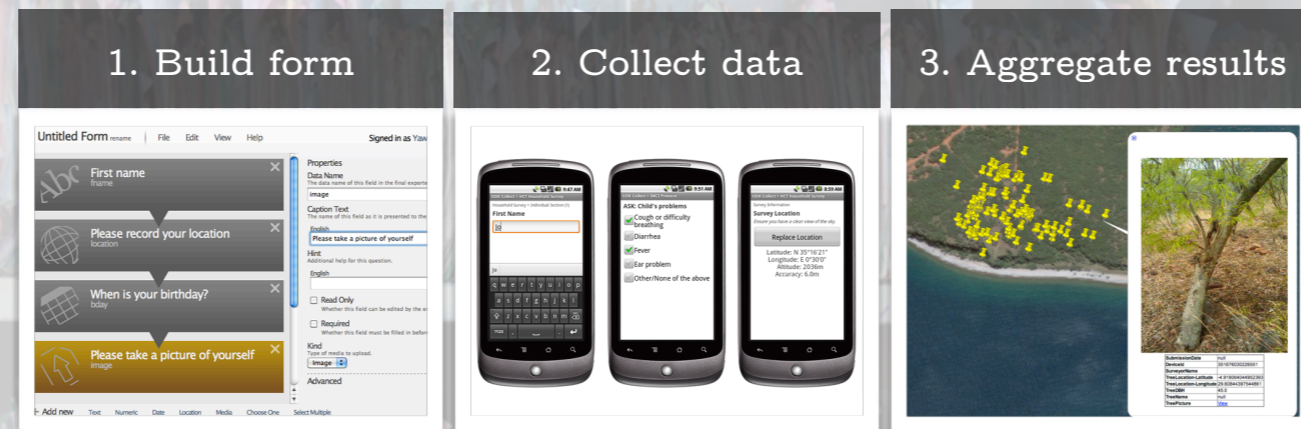
http://nafundi.com (@nafundi)
http://opendatakit.org (@opendatakit)
http://anokwa.com (@yanokwa)

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.

ODK is pretty easy to use and lots of people deploy it by themselves. But. If you need help, Nafundi is there to help.

We have years of experience – in building well-known data collection tools like Open Data Kit and in running data collection campaigns all over the world.

So with that, are there are any questions I could answer? (e.g., how is this free, what about security, what other tools are there?)