OPEN DATA KIT: Building and Refining Information Services for under-served Populations

Waylon Brunette
Department of Computer Science & Engineering

The work being discussed was done by a LARGE TEAM in UW-CSE.
In Tribute...

- Remembering Prof. Gaetano Borriello
- Open Data Kit’s visionary and leader
Organizations working in developing regions rely on field workers to collect data.
Paper Systems are difficult to Search and Transport
- Hard to search
  - “Find the patient record for ‘Hartung’ ”

- Impossible to query
  - “How many patients presented with Malaria last month? ”

- Hard to share
  - What if patient goes to a different clinic?
Ideas that shaped ODK

• Technology by itself is generally NOT a solution

• Magnify *Human* resources through Technology

• Technology requires support infrastructure
Open Data Kit (ODK)

- Project started in 2008-09
- Mobile data collection tools for Android devices
  - There was too much focus on clumsy J2ME devices
  - Not enough focus on the trajectory of technology
- Modular, open architecture
  - Too many stove-piped solutions
  - Large commercial enterprise systems
- Open source (Apache2)

GOAL: Magnify human resources through technology
Platform Shift from PCs to 
Smart Phone + Cloud
GOAL: Magnify human resources through technology

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

http://opendatakit.org
ODK 1.x: Data Collection

ODK Collect

- Renders question prompts defined by the user
ODK Collect Question Layouts

ODK Collect > Forest Plot Survey

**Plot type**

- Plot status:
  - Yes
  - No

- Logging?
- Fire?
- Grazing?
- Mining?
- Roads?
- Farming?

**DBH type**

Tree (1)
ODK 1.x: ODK Collect

Automated Survey Renderer with enhanced data types
- Pictures, Video
- GPS
- Barcode
• University College London project
• Goal to have indigenous population report poachers.
ODK 1.x

• Focus on data collection
ODK Video

- https://www.youtube.com/watch?v=CHyUS4JnajM
“Features sometimes trump cost...we needed to get more out of the phone and ODK allowed us to do that.”

D-Tree International
“It is relatively easy to train using ODK...even for people who have little or no computer experience.”

Berkeley Human Rights Center
“[Health workers felt] the system facilitated their home visits and allowed them to collect higher quality data. We actualized higher cost savings over our previous PDA-based system.”

USAID-AMPATH
USAID-AMPATH

- Began using ODK in 2009
- Provides care to over 160K HIV patients in Kenya
- Enrolls about 2K new patients a month
- In 2014, had over 1 million entries with ODK
Other Deployments

• Neglected Tropical Diseases Support Center
  – Based at the Task Force for Global Good
  – Collected 50 million data points across 200 projects in 20 countries
The Carter Center uses ODK as part of their ELMO software to monitor elections.
Training in Gombe National Park
Effort (yellow) and threats (red) collected by village forest monitors using ODK around Gombe National Park
Training with Surui

Carbon market certification and indigenous culture
Videos

• Reproductive Health Vouchers
  – http://vimeo.com/38123850
ODK’s Usage

Website visits from 225 countries/territories
(187 countries/territories last month)
Over 300,000 unique visitors from 225 countries/territories (Average of ~27,000 hits a month)
Community

- Many consulting companies (50+) “support” ODK tools

- Nafundi

- Dimagi

- SurveyCTO

- MANY MORE

SurveyCTO has been designed to make sophisticated, high-quality electronic data-collection as simple as possible – for academic research, M&E, mHealth, market research, and more.

Use Excel spreadsheets to define your survey forms, use SurveyCTO expression-builders to easily define field validation and skip logic, and use automatically-generated templates for processing in Stata or mail-merging into Word.

Synchronize, decrypt, export, and process data with two simple clicks.

Secure

SurveyCTO has been built with data-security (and Institutional Review Boards) in mind.

With 2048-bit encryption from the point of data-collection to the point of data-processing, you can be sure that your data is secure.

For your most highly-sensitive data, SurveyCTO also makes it easy to synchronize encrypted data with a cold-room computer. That way, sensitive data is never accessible from an Internet-exposed computer.
Open Data Kit

• Make tools highly modular and customizable
  – Enables organizations to compose tools that are appropriate for their deployments

• Exploit open interfaces and standards
  – Avoid “silico-ed” monolithic proprietary solutions

• Allow organizations to leverage evolving technologies
  – Avoid early obsolescence
Missing capabilities

• Updating data on the mobile device
  – Allow users to view and edit collected data
• Customizing applications to different situations without recompiling
• Collecting information from various sensing devices
• Usage of cheaper technologies (e.g., paper, SMS)
ODK Versions

CURRENT TOOLS
(Referred to ODK 1.x)

versus

NEW TOOLS
(Referred to ODK 2.x)
ODK 1.0 Architecture

**Cloud**
- Database
- Fusion Tables
- CSV
- KML

**Mobile Device**
- Collect
- File System

**FLOW**
- FILES
ODK 2.0 Architecture

CLOUD

AGGREGATE

Database

Fusion Tables

CSV

KML

MOBILE DEVICE

SCAN

SURVEY

TABLES

Database

File System

SUBMIT

SENSORS
ODK 2.0: HTML on the Device

- Highly Customizable
- Dynamic Graphs
- Watermarks
- Access Server Content (AJAX)

Only WebDev Skills

Redesign

This is a custom template that uses D3.js to plot an age and weight on a growth chart:

- male median weight for age

PATH
ODK 2.0: Data Collection

- Collect while disconnected
- Submit opportunistically

ODK Survey
(Android Cellphone / Tablet)

Bi-directional
Submit
And
Update

Local and Remote
(various tools)
Data Architecture of ODK 2.0

Local copy of **all relevant data** collected from **any** device

- **Local DB on client**
- **Sync to cloud**
- **Media files**

Redesign
ODK 2.0: Survey

• Complex workflows and technology assisted trainings
• Different approach than ODK Collect
  – Easier customizations via HTML and CSS
  – Easier branching and workflows
  – Easier to access databases that are synced with server
ODK 2.0 Example

- **Pneumonia Detection** (Ghana & India)
- Digitize complex WHO-IMCI workflows
- Guide and assist user obtaining proper patient measurements
- Display treatment based on IMCI guidelines
ODK 2.0: Tables

- Gap in Open Data Kit’s tool suite
- Needed a flexible tool for applications where previously collected data is often revisited and updated.
Vaccine Cold Chain
Cold chain end-points

4,100 doses
Polio and Measles
$635

625 doses
Rotavirus
$4687
ODK 2.0: Tables

- Visualization of underlying database
- Provide users different looks at data
- **Example:** Cold chain monitoring

**List View**
- Kandi Relais
  - Dépôt Relais
  - Alibori Kandi

**Detailed View**
- **Tchikandou**
  - Centre de Santé d’Arrondissement
  - Departement: Borgou
  - Commune: Nikki
  - Zone sanitaire: Nikki-kalalé-Pérèrè
  - Arrondissement: Ouenou
  - Total Population: 4687
  - Grid Electricity: None
  - Latitude (GPS): 9.84234
  - Longitude (GPS): 3.36392

**Map View**
- Dakar
  - Population: 1774

**Spreadsheet**
- Table: Facilities
  - Facility: Tchikandou
  - Facility Type: Health centre - MoH
  - Electricity level: Under 8 hours / day
  - Kerosene: Not available
  - Gas: Not available
  - Stock-outs: No
  - Base surplus: Over 30% surplus
  - Population: 17142
  - Fridges: MMD04, V.170.EK

**Contact Information**
- Person in charge: KARIMOU BATTA C. MAKO
- Mobile number: 97951246
- **2 Refrigerators**
ODK Tables and D3
# ODK Tables – Hope Study App

## Hope Study

**Screen Female Client**

**Follow Up with Existing Client**

**Send Data**

### Intervention Arms

- **Hope: 35%**
- **Control: 41%**
- **Ineligible: 24%**
- **HIV: 16%**
- **Not Tested: 24%**

### Table: Female Clients (List)

<table>
<thead>
<tr>
<th>Client ID</th>
<th>Age</th>
<th>Randomization</th>
</tr>
</thead>
<tbody>
<tr>
<td>14752</td>
<td>21</td>
<td>Control</td>
</tr>
<tr>
<td>88563</td>
<td>20</td>
<td>Control</td>
</tr>
<tr>
<td>98563</td>
<td>29</td>
<td>Control</td>
</tr>
<tr>
<td>25364</td>
<td>19</td>
<td>Control</td>
</tr>
<tr>
<td>23456</td>
<td>23</td>
<td>HOPE</td>
</tr>
<tr>
<td>36985</td>
<td>36</td>
<td>HOPE</td>
</tr>
</tbody>
</table>

### Graph View

- Add Client
- Search
Focus of ODK Research

- Clinic decision support
- Mixing paper and digital
- Rapid diagnostic tests
- Internal sensors - spirometry
- External sensors – cold chains, milk pasteurization
ODK SCAN
Vaccine logistics in Mozambique
SpiroSmart
ODK Diagnostics
ODK SENSORS:
Time spent gathering water in Ethiopia
ODK 2.0 Architecture

CLOUD

AGGREGATE

Database

CSV

Fusion Tables

MOBILE DEVICE

SCAN

SURVEY

TABLES

Database

File System

SUBMIT

SENSORS

Sensor Discovery

Sensor Manager

Database Manager

Driver 1

Driver 2

Sensor Drivers

USB Manager

Bluetooth Manager
ODK 2.0 Architecture

Cloud

Mezuri Platform

Mobile Device

Scan

Survey

Tables

Database

File System

Submit

Sensors
ODK 2.0 Architecture

MEZURI PLATFORM

- Track provenance of collected data
- Customizable dashboard for projects
- Data sharing capabilities with built-in security

CLOUD

MOBILE DEVICE

SCAN

SURVEY

TABLES

Database

File System

SUBMIT

SENSORS

Sensor Discovery

Sensor Manager

Database Manager

Driver 1

Driver 2

USB Manager

Bluetooth Manager
Planned activities

• Interactive voice response (IVR)
• P2P data transfer
• SMS connections
• Active scripts within Tables
• Develop/deploy several complete apps using 2.x toolset
Thank you to our partners!

• **Our NGO Partners:**
  PATH, AMPATH, VillageReach, GSID, Jane Goodall Institute, Google Earth Outreach, ....

• **Our Funding Partners:**
  NSF, Google, USAID, DARPA, NIH, Gates Foundation, Samsung, Microsoft
Questions?

Thank you for your attention.

http://www.opendatakit.org