Open Data Kit

Yaw Anokwa
University of Washington
Box 352350
Seattle, WA 98195-2350, USA
yanokwa@cs.washington.edu

Gaetano Borriello
University of Washington
Box 352350
Seattle, WA 98195-2350, USA
gaetano@cs.washington.edu

Waylon Brunette
University of Washington
Box 352350
Seattle, WA 98195-2350, USA
wrb@cs.washington.edu

Carl Hartung
University of Washington
Box 352350
Seattle, WA 98195-2350, USA
chartung@cs.washington.edu

Adam Lerer
Massachusetts Institute of Technology
77 Massachusetts Ave.
Cambridge, MA 02139 USA
alerer@mit.edu

UbiComp 2009, Sep 30 – Oct 3, 2009, Orlando, FL, USA

Abstract
Mobile phones are becoming pervasive in developing regions creating an opportunity to address data collection needs. Existing paper-based approaches are often slow and incomplete when compared to mobile phone based data collection methods. In this demonstration we present Open Data Kit, a set of mobile phone based tools used to collect, manage, and aggregate data.

Keywords
Data Collection, Mobile Phones, ICT

ACM Classification Keywords
H.5.3 Group and Organization Interfaces

Introduction
Organizations in under-served regions need to be able to collect data in any location including places with unreliable power and limited connectivity. Open Data Kit (ODK) is designed to leverage mobile devices to improve community-based data collection (e.g. citizen science, public health informatics, etc) by providing a set of tools that enable users to create data collection solutions in days rather than months or years. Mobile devices such as phones and PDAs have proven to be a faster and more complete method of data collection than traditional methods of pen and paper [1], [2].

ODK is designed to automate data collection, aggregation, and analysis in places with limited

technical infrastructure. ODK abstracts technical details from the user; thereby allowing domain experts (e.g. doctors, foresters, etc.) to rapidly create and deploy surveys with limited effort and minimal assistance from a programming expert.

Many existing systems that support data collection have little interoperability. ODK seeks to bring disparate efforts together by creating a modular architecture with standardized interfaces. ODK is working with the OpenRosa Consortium [3] and the Open Mobile Consortium [4] to create tools for mobile data collection, aggregation, analysis, and reporting by developing open source solutions conforming to the W3C Xforms specification. Other OpenRosa projects such as clinical trial software [5] and community health worker management [6] are building solutions upon a common framework with ODK.

Open Data Kit Demonstration

The demonstration will showcase the following tools:

Data Collection - ODK Collect
ODK Collect renders an XForm into as sequence of input prompts that provide navigation logic, entry constraints, and repeating sub-structures. Data entered by users can then be upload via GPRS, Wi-Fi, and/or a direct connection (SD card/cable). Currently, ODK collect uses the Android platform to take advantage of its GPS and camera capabilities. The various methods of input (e.g. video, audio, photos, location) enable the collection of rich data for later off-line analysis.

Data Storage Server- ODK Aggregate
ODK Aggregate provides a ready to deploy server repository to: manage collected data, provide standard interfaces to extract data (e.g., spreadsheets, queries, etc.), and integrate with existing systems via web requests. ODK Aggregate is currently implemented on Google's App Engine allowing users to avoid the hassles of setting up their own scalable web service.

Device Management - ODK Manage
ODK Manage maintains a centralized list of a phone's current state enabling remote device management by allowing changes to be propagated to all phones in a particular deployment. To start an update, manage sends a message to the phone with a pointer to a URL that specifies the pending transfers of forms, data, and applications (for in-field upgrades).

Example Deployments

In November 2008, a version of ODK was used in Uganda to collect over 1000 surveys evaluating SMS applications deployed by Grameen AppLab Kampala. Currently, ODK-Collect is being readied for deployment with an AMPATH/USAID project for HIV monitoring in Kenya that will use 300 G1 phones and will eventually bring HIV health services to more than 2 million people.

References

[1] K. Shimira et al., *The use of personal digital assistants for data entry at the point of collection in a large household survey in southern Tanzania*. Emerging Themes in Epidemiology, 2007.

[2] T. Parikh, et al., *Mobile Phones and Paper Documents: Evaluating a New Approach for Capturing Microfinance Data in Rural India*. CHI, 2006.

[3] OpenRosa, http://openrosa.org

[4] Open Mobile Consortium, http://open-mobile.org

[5] MEVAC, http://epihandy.org/index.php/OMEVAC.

[6] CommCare, http://commcare.cs.washington.edu