



Mobile Data Collection and Geomapping in Private Sector Development

An innovative approach developed by
DTI GIZ Private Sector Promotion in the Philippines
and its potentials in development cooperation



On behalf of

BMZ



Federal Ministry
for Economic Cooperation
and Development



Responsible

Dr Volker Steigerwald volker.steigerwald@giz.de

Author

Uwe Sturmman uwe.sturmann@gmail.com

Editor

Leah Divina Siton Steigerwald leah22@gmail.com

Consultant

Contributor

Arvin dela Cruz arvin.dela-cruz@giz.de

Publisher

Private Sector Promotion Program PSP SMEDSEP
smedsep.ph

PSP Program Office

7F New Solid Building, 357 Sen Gil Puyat Avenue
Makati City 1226 PHILIPPINES

Volker Steigerwald PhD

Program Manager

volker.steigerwald@giz.de

September 2011

Mobile Data Collection and Geomapping in Private Sector Development

An innovative approach developed by
DTI GIZ Private Sector Promotion in the Philippines
and its potentials in development cooperation

Contents

Figures.....	3
Acronyms.....	4
Summary	5
1 Background.....	7
Situation Analysis	7
Need for a Baseline Survey	8
Disadvantages of Paper based Surveys	9
2 Description of the new Approach.....	10
Hardware - Tablet computers as survey tools	10
Software - Data Collection via <i>ODK Collect</i>	12
Google Geo Tools.....	14
3 Mobile Data Collection - The Bohol Experience	18
Design of Questionnaires / Data Collection Forms.....	18
Survey Preparations and Training of Enumerators.....	19
Implementation of Tourism Data Surveys	20
Tourist Satisfaction Survey (TSS).....	20
Tourism Enterprise Survey	21
Data Analysis.....	22
TSS Analysis	22
TES Analysis	23
Mapping of Tourism Data with Google's Geo Tools	23
Google Spreadsheet Mapper.....	23
Mapping tourism locations on Google Earth and Google Maps.....	24
Google Map Maker	25
4 Assessment of the new Approach	26
Advantages of Mobile Data Collection using <i>ODK</i>	26
Assessment of OECD DAC Criteria	28
5 Perspectives for Mobile Data Collection and Geo Mapping	30
Potentials in Tourism Value Chain Development	30
Potentials in MSME Development.....	31
Potentials for Market Surveys and Customer Insights.....	32
Potentials in Rural / Agricultural Development	32
Potentials in Environment and Natural Resources Management.....	33
Potentials in Social Development, Gender and Health.....	33
Potentials in Decentralization and Local Governance Programs	33
Potentials in Results based Monitoring (RBM)	34
6 Conclusions and Recommendations	36

Annexes.....	37
Annex 1 Glossary:	37
Annex B Tablet Computers	44
Annex C Hardware Preparation	45
Annex D Open Data Kit.....	46
Annex E Google Map Maker	48
Annex F <i>ODK Collect</i> – Screenshots	50
Annex G Links to Tutorials of Google Geo Tools	54
Annex H: ODK Deployments	55

Figures

Figure 1 The Samsung Galaxy Tab - a Google Android OS based tablet computer	10
Figure 2 ODK Collect Tourism Satisfaction Survey Question	20
Figure 3 ODK Collect Inventory of Tourism Operators Survey Question	21
Figure 4 TSS Spreadsheet Analysis.....	22
Figure 5 TSS Analysis Chart Visualization	22
Figure 6 Original Google Spreadsheet Mapper version 3.....	23
Figure 7 Bohol Tourism Map on the Google Earth App	24
Figure 8 Bohol Tourism Map online on Google Maps.....	24

Acronyms

API	Application Programming Interface
BPMS	Business Profile Management System
BDS	Business Development Services
BPLS	Business Permit and Licensing System
BTO	Bohol Tourism Office
CFIF	Cebu Furniture Industries Foundation
csv	comma separated values
DAC	Development Assistance Committee (of the OECD)
DHS	Demographic and Household Survey
DTI	Department of Trade and Industry, Philippines
DOT	Department of Tourism, Philippines
GDG	Google Developer Group
GIS	Geographical Information System
GIZ	Deutsche Gesellschaft fuer Internationale Zusammenarbeit (formerly GTZ)
GPS	Global Positioning System
GSM	Global System for Mobile Communications
GTH	Gifts, Toys and Housewares
HNU	Holy Name University, Bohol, Philippines
ISIC	International Standard Industrial Classification
kml	Keyhole Markup Language (file format in Google Earth / Google Maps)
kmz	kml file zipped
LGU	Local Government Unit
LRED	Local and Regional Economic Development
MSME	Micro, Small and Medium Enterprises
ODK	Open Data Kit
OECD	Organization for Economic Cooperation and Development
OSM	Open Street Map
PBR	Philippine Business Registry
PDA	Personal Digital Assistant
POI	Point of Interest
PSP	Private Sector Promotion (Program)
RBM	Results based Management, Results based Monitoring
REA	Rapid Economic Assessment
RCT	randomized controlled trials
SEC	Securities and Exchange Commission, Philippines
SMEDSEP	Small and Medium Enterprise Development for Sustainable Employment Promotion
TSA	Tourism Satellite Account
TSS	Tourist Satisfaction Survey
UN WTO	United Nations World Tourism Organisation
VC	Value Chain
WLAN	Wireless Local Area Network
USD	US dollars

Summary

Most private sector development programs need to generate data on MSMEs, markets and supply chains for various purposes, such as baseline surveys, business profiles for enterprise matchmaking, consumer satisfaction surveys, quantitative analysis of value chains, business advisory services, results monitoring, as well as general statistics on economic development and growth, particularly the contribution of micro, small and medium enterprises (MSMEs).

Data collection has always been a tedious task. Besides the usual hesitation of MSME owners and managers to spend time and provide confidential enterprise information to interviewers, it has been the time consuming effort of completing lengthy forms and questionnaires, followed by the equally arduous process of encoding data and transferring them from paper into PC format, that made data collection such a dreadful exercise.

After all the efforts of collecting and transforming data, it has often been experienced that analysis results were less than meaningful due to data gaps, incorrectly completed forms, errors in encoding and data input, etc. Results were often outdated by the time the data were processed. Analytical capabilities were limited among staff. Hence the results of enterprise surveys were often not worth the efforts invested into them.

During the past two years, tablet computers have become more and more popular, mostly because of the Apple iPad, but also because of the increasing functionality at a reasonable price. Smaller tablet computers like the Android based Samsung Galaxy Tabs with their 5" to 7" screen are handy enough to become working instruments for empirical enterprise surveys in the field. Special software has become available to design forms and questionnaires for data input on interactive touch screens.

The DTI - GIZ Private Sector Promotion (PSP SMEDSEP) in the Philippines has pioneered mobile data collection using tablet computers for tourist satisfaction surveys as well as enterprise profiling. The handheld devices are used to collect alphanumerical data using the touch screen, but can also geographically locate enterprises using GPS (Global Positioning System) and take photos using the built in digital camera. These features allow for placing enterprise and product information on web based mapping sites like Google Earth and Google Maps enabling users to locate the workshop facilities and / or showroom for increased sales opportunities. All data are almost instantaneously available for analysis and further action.

Many GIZ programs and partner agencies would benefit from using tablet computers to conduct their enterprise surveys in order to update business profiles in various sectors including the suppliers of raw materials and intermediate products. The GPS based mapping of MSMEs will allow for identifying enterprise clusters for targeted assistance and business development support. Maps can serve as a spatial communication tool, which add a new dimension to many socioeconomic phenomena.

Finally, the use of tablet computers can invoke the enthusiasm of program and partner staff and make enterprise data and consumer data collection an exciting and worthwhile activity. The presence of quality data will also contribute to more and effective data driven decision making in development projects and programs as well as their partners in government, the private sector and civil society.

This knowledge product provides a brief overview of the mobile data collection approach using tablet computers and Open Data Kit (ODK) software. It further showcases the experiences gained by PSP SMEDSEP and explores the fields of its potential application in the context of development cooperation. However, this text is not an introduction into the technical hardware and software issues involved. Educational videos and step-by-step tutorials on ODK and Google's geo tools are available from the *Google Earth Outreach* website. This booklet encourages development practitioners to explore the mobile data collection approach and adapt it to their work context. The author is a former GIZ Senior Adviser and is now a freelance consultant. He is available for further information and advice: uwe.sturmann@gmail.com.

1 Background

Situation Analysis

Development cooperation intervenes in complex environments to initiate positive results for the intended beneficiaries of the technical and / or financial assistance provided. This requires a thorough understanding of the living and working conditions of both target groups and intermediaries.

Planners, implementers and evaluators of development projects and programs require a solid data basis for decision-making and monitoring changes among their stakeholders. As development practitioners they need accurate and reliable information on the situation of the intended beneficiaries, the change process and the ultimate benefits of the support activities.

Similarly, support agencies require data on economic activities, agricultural production, governance aspects, social and organizational characteristics, environmental and ecological performance, climate change adaptation and mitigation, energy usage, quality, social and sustainability standards, etc.

Most project planning documents rely on secondary data, often on government statistics of the national government agencies in the partner countries. However, these databases are often either outdated or incomplete and originate from unreliable data sources established through little known data collection methodologies and error margins.

The consequences are that projects and programs are designed and implemented on the basis of dubious data that are often not even specific to the project area, but to other administrative areas or research samples, or even the national average figures. Existing data stocks are fragmented in different formats, like in word files, excel spreadsheets, relational databases and even basic hard copy lists. It comes without surprise that usually also the variables are inconsistent among the different data stocks.

This is also the typical reality in the Philippines. The Department of Trade and Industry (DTI) - for example – has both regulatory and business support functions. This does not imply that DTI has a credible data basis. While in the future all registered enterprises shall be mandatorily listed in the new Philippine Business Registry (PBR), the existing data come from different sources, namely the Local Government Units (LGUs), the Securities Exchange Commission (SEC) and the DTI operated Business Profile Management System (BPMS). While the data quality of more than 1500 LGUs lack consistency and integrity, the SEC data are limited to regulatory aspects, and the BPMS data suffer from limited scope and depth; only 140 DTI clients were surveyed per province due to the limited financial resources available.

When the PSP SMEDSEP program was requested by DTI Bohol to assist the tourism stakeholders in the province using the Value Chain (VC) approach, the PSP SMEDSEP team found a myriad of different data stocks of unknown quality and various degrees of detail, such as the hotels and resorts listed by the Bohol Tourism Office (BTO), the enterprises registered by the DTI, the establishments accredited by the Department of Tourism (DOT), the membership lists of various tourism associations, etc. No two data sources were compatible with each other, No single database appeared to be a reliable

source for decision making. In such cases the program implementers have no other option than to generate primary data by conducting a baseline survey of all relevant stakeholders.

Need for a Baseline Survey

While the monitoring guidelines of most donors and technical assistance agencies rightfully suggest the use of existing secondary data, like the use of country systems for administrative and / or statistical purposes, the reality in many developing country situations is such that the wish is father to the thought. Useful data hardly ever exist.

In these cases the donor guidelines for results based monitoring (RBM) require the project or program staff to collect baseline data for comparisons of the target group status before and after the program interventions. However, in many programs this is hardly a reality.

Implementers often neglect baseline surveys of the initial situation of target groups and program areas. Data collection is usually not considered a priority by most programs. The (perceived) pressure to get started and not to get bogged down by cumbersome (and often unproductive) data collection processes effectively prevents the programs to lay a thorough empirical basis for their future work. But there is no shortcut to progress.

Without baseline data neither program objectives can be adequately formulated nor can indicators be quantified. Consequently, the monitoring system will not be able to provide implementers and evaluators with the required information to assess changes in the status of target groups and / or in the program environment. If ever a proper baseline survey is conducted by a program, then it is done in compliance with donor requirements for having a results-based monitoring system in place rather than for having a solid basis for program planning and steering of program activities.

The results orientation in development cooperation programs has unfortunately led to an amazing paradox, a focus on reporting for accountability reasons rather than a focus on results at the target group level. This trend away from target group benefits is motivated by program objectives that aim at macro and/or meso levels; for example the improved *framework conditions* and the development of new approaches and support systems. Few programs have the micro level of enterprises or households formulated in their objectives. The focus has moved to abstract changes in frameworks and systems.

However, good development work starts with a thorough understanding of the ultimate beneficiaries. Program implementers need to know their target groups, their socioeconomic characteristics, their business features, their strengths and weaknesses, their environmental situation and their vision for the future.

In order to get the tourism VC analysis in Bohol started, the PSP SMEDSEP team required a quantitative overview of the tourism related enterprises in the different functions, the tourism attractions, etc. The baseline survey design included data for the quantification of the value chain, such as products and services offered, capacity of the establishments, prices, markets and sales, etc. In order to identify the availability of resources and gaps, the PSP SMEDSEP team also required locational data of tourism enterprises and attractions to map their spatial distribution in the island destination of Bohol.

With these data available from a quantitative survey, the team was not only able to describe the current status of the tourism stakeholders with their opportunities and constraints, but

also created a baseline for future results monitoring, which observes the changes over a period of time.

The typical approach for such a survey would have been to design a questionnaire format and have enumerators / interviewers hired to collect the data into paper based questionnaires.

Disadvantages of Paper based Surveys

There are many reasons why data surveys are not popular among implementers of development projects and programs.

- Data surveys are time consuming. From the concept stage and the design of data collection instruments through actual data gathering and data compilation to data analysis and interpretation may often take several months, if not more.
- Paper-based questionnaires contain data entry errors: Subject to the quality of interviewers and enumerators paper based data collection forms usually contain numerous errors, starting from wrong choice, data format, spelling errors, to omissions and data gaps, all of which require extensive data cleaning efforts.
- Encoding is time consuming and cumbersome: The transfer from paper based forms to electronic database is manual, error prone and often cumbersome because of the poor state of raw data. Most implementers have witnessed the piles of paper forms in an office corner that wait for encoding and capturing into a database.
- Data cleanup is time-consuming and costly: While the raw data on the paper-based forms is already faulty, the encoding process is will add more mistakes into the data stocks due to flawed transcription. The necessary data cleaning process is cumbersome, and may not be able to restore the original meaning.
- The data analysis yields less than satisfactory results: Due to the many flaws usually observed in the data collection and collation process the data analysis often hits unforeseen obstacles, which cannot be overcome at the analysis stage.

2 Description of the new Approach

Hardware - Tablet computers as survey tools

During the past two to three years, tablet computers have become more and more popular, not only because of the well known iPad, but more so because of the increasing functionality of these handheld devices at a reasonable price. Smaller tablet computers like the Android based Samsung Galaxy Tabs with their 5" to 7" screen are lightweight and handy enough to become working instruments for empirical enterprise surveys in the field. Special software has become available to design forms and questionnaires for data input on interactive touch screens.



Figure 1 The Samsung Galaxy Tab - a Google Android OS based tablet computer

Box 1 Tablet Computer definition

Tablet Computer

From Wikipedia, the free encyclopedia (www.wikipedia.org)

A tablet computer, or a tablet, is a mobile computer, larger than a mobile phone or personal digital assistant (PDA), integrated into a flat touch screen and primarily operated by touching the screen rather than using a physical keyboard. (...) Besides having most PC computer capabilities, typical tablet computers purchased recently include wireless Internet browsing functions, potential cell phone functions, GPS navigation, and video camera functions, weigh less than 1 kg and typically have a battery life of three to ten hours.

PSP SMEDSEP in the Philippines has pioneered mobile data collection using Android based tablet computers for tourist satisfaction surveys as well as enterprise profiling. The electronic devices has not only collect alphanumerical data using the touch screen, but has also geographically locate enterprises using the Global Positioning System (GPS) receiver and take photos using the built in digital camera. This allows for precisely placing enterprise and product information on web based mapping sites like Google Earth and Google Maps enabling users to locate the workshop and/or showroom facilities for increased sales opportunities. All data are almost instantaneously available for analysis and further action.

While the first tablet computers, especially the Apple iPad, were very innovative with their flat design and the intuitive operation via a 10" touch screen, their size and weight of 700g was more suitable for watching multimedia and playing games in the home environment.

However, the second generation of tablets added a 7" touch screen to the available dimensions, like the Samsung Galaxy Tab based on Google's Android operating system. This unit is light weight (350g) and small enough to be held in a single hand and operated with the other without additional support. The 7" display is still big enough to see details, to touch buttons and to type on the virtual keyboard without hitches. It is perfectly usable for data entry in the field. Recent 2012 models with a 5" screen have even less weight (under 200g), and still sufficient screen space for use as a data collection instrument.

All tablet models have an integrated camera with a resolution sufficient to make basic quality photos. The advantage of an integrated camera over an external camera is that the photos can be shot from within the data collection form (see Chapter 3.2 on the *ODK Collect* software). Should high quality photos be required an external camera may be needed. However, ongoing technological advances will probably enhance the photographic capabilities of tablet computers very soon.

A very useful feature of most tablet computers is the integrated GPS receiver that can determine the precise location of the unit (within 5m) by recording latitude, longitude and altitude. The GPS capability is used for geographic mapping and navigation with software apps available for Android devices. The GPS coordinates can also be added to data collection forms using *ODK Collect* software.

Box 2 Advantages of Tablet Computers

Advantages of Tablet Computers

From Wikipedia, the free encyclopedia (www.wikipedia.org)

- Usage in environments not conducive to a keyboard and mouse such as lying in bed, standing or handling with a single hand
- Lighter weight, lower power models can function similarly to dedicated ebook readers
- Touch environment makes navigation easier than conventional use of keyboard and mouse or touch pad in certain contexts such as data collection in the field
- The ability for easier or faster entry of diagrams, mathematical notations, and symbols
- Allows, with the proper software, universal input, independent from different keyboard localizations
- Current tablets typically have longer battery life than laptops or netbooks.

There is no need for GSM cell phone capability, as the tablet computer shall be used for data collection only. Transmission of collated data is done via WLAN. The absence of GSM capability allows for very low cost. By mid 2012 the unit price for a Samsung Galaxy S WIFI 5.0" had dropped to less than 200 USD, hence making tablet computers very affordable even in developing country contexts.

Software - Data Collection via *ODK Collect*

Open Data Kit (ODK) is a suite of software tools that allows data collection using mobile devices like tablet computers and smart phones using the Android operating system. Especially users in developing countries in Africa and Asia have applied the new technology. In addition to the socioeconomic and health surveys - with GPS locations and images - ODK is being used to create decision support for clinicians and for building multimedia rich nature mapping tools.

First of all, the data collection forms have to be designed. **ODK Build** is a form designer with a drag-and-drop user interface. *ODK Build* is an HTML5 web application and works best for designing simple forms.

ODK Collect renders forms into a sequence of input prompts that apply form logic, entry constraints and repeating substructures. Enumerators work through the prompts and can save the submission at any point. Finalized submissions can be sent to (and new forms downloaded from) a server. *ODK Collect* uses the Android platform, supports a wide variety of prompts (text, number, location, multimedia, barcodes) and works well without network connectivity.

Data is submitted to an online server, even without an Internet connection or mobile phone (GSM) service at the time of data collection. There were only two instances when the Android device with the *ODK Collect* app has needed to connect to the internet:

- to download survey forms and
- to upload completed survey forms (1 to 100+ at a time).

Once data is collected in the field with *ODK Collect*, it can be uploaded and managed using *ODK Aggregate*, which is the secure cloud based storage platform that accepts the data and can send it on to external applications, if required. *ODK Aggregate* also allows for downloading datasets in aggregated formats, such as a single csv file.

Box 3 Open Data Kit

Open Data Kit (**ODK**)

Open Data Kit (*ODK*) is a free and open source set of tools, which help organizations author, field and manage mobile data collection solutions. *ODK* provides an out-of-the-box solution for users to

- build a data collection form or survey
- collect the data on a mobile device and send it to a server and
- aggregate the collected data on a server and extract it in useful formats.

ODK Collect allows for the integration of alphanumeric information, audio-visual files and locational data in a single dataset

- Text information: single choice, multiple choice and open-ended questions
- Numerical data: preformatted data fields which require numbers or dates
- Audiovisual files: photos and video with the integrated tablet camera, spoken word and sound with the built in microphone
- Locational data: recording of GPS coordinates (longitude and latitude).

ODK Collect data forms allow for multiple instances (groups) of an item (business has several products; mother has several children, etc)

If the tablet computer is equipped with a GPS chipset, users can add a location entry into their form. The question will prompt the user to record the GPS location. The user will be able to view an estimate of the GPS accuracy while the tablet's GPS receiver acquires satellite positions.

Box 4 Global Positioning System GPS

Global Positioning System (GPS)

From Wikipedia, the free encyclopedia (www.wikipedia.org)

GPS is a space-based satellite navigation system that provides location and time information in all weather, anywhere on or near the Earth, where there is an unobstructed line of sight to four or more GPS satellites. It is maintained by the United States government and is freely accessible to anyone with a GPS receiver.

More and more organizations are using mobile devices to collect data in the field. *Open Data Kit (ODK)* is a suite of tools that allows data collection using Android mobile devices and data submission to an online server, even without an Internet connection or mobile carrier service at the time of data collection. The data collection process can be streamlined with *ODK Collect* by replacing traditional paper forms with electronic forms that allow text, numeric data, GPS, photo, video, barcodes and audio uploads to an online server. Organizations can host their data online using Google's hosting platform, *AppEngine*, manage their data using *ODK Aggregate* and visualize their data as a map using *Google Fusion Tables* and *Google Earth*.

Data Analysis

Alphanumerical data can be analyzed with different software tools, depending on the kind of research. In an experimental or quasi experimental research design, statistical analysis may require sophisticated statistical software packages like SPSS or Stata, which allows for calculation of variance, confidence levels, regression analysis, etc.

In the development cooperation context such scientific analysis is usually not required. In most situations it is beyond practicability and affordability to conduct research that conforms to scientific standards. What is required are robust and credible data on target group characteristics that can be measured and crosstabulated with other variables, in order to understand frequency distributions, arithmetic means, correlations between attributes as well as changes over time.

Box 5 Geographic Coordinate System

A **geographic coordinate system** is a coordinate system that enables every location on the Earth to be specified by a set of numbers. The coordinates are often chosen such that one of the numbers represents vertical position and two or three of the numbers represent horizontal position. A common choice of coordinates is latitude, longitude and elevation.

If the focus of the survey is not on statistical analysis but on data exploration and decision making, data mining tools seem to be the better option. As much as standard statistical functions are possible, the main objective here is to understand and get a good overview of the distribution of research objects. The software tool *InfoZoom* may serve this purpose because of its sophisticated and use -friendly data visualization capabilities. Data can also be cleaned, transformed and augmented with *Google Refine* software, while *Google Fusion Tables* can be used for analyzing and visualizing structured datasets.

Google Geo Tools

Google Earth

The GPS coordinates recorded with *ODK Collect* or any other method can be exported to a KML (<http://en.wikipedia.org/wiki/KML>) file for use with virtual globes and maps. The most common one is Google Earth. Alternatives include Open Street Maps (OSM), Microsoft's Bing maps, GRASS GIS and QGIS (for Mac), among others.



Figure 2 DTI and partners in Region 8 practice using Google Refine to clean up data

Box 6 Google Earth

Google Earth

From Wikipedia, the free encyclopedia (www.wikipedia.org)

Google Earth is a virtual globe, map and geographical information program that maps the Earth by the superimposition of images obtained from satellite imagery, aerial photography and GIS 3D globe.

As of October 2011 Google Earth has been downloaded more than a billion times. Google Earth is useful for many day-to-day purposes. The user can explore and place location bookmarks in literally every corner of the globe. One can also get directions using Google Earth, using variables such as street names, cities and establishments.

Google Earth can also function as a hub of knowledge, depending on the user's location. By

enabling certain options, one can see the location of gas stations, restaurants, museums, and other public establishments in their area. Google Earth can also dot the map with links to images, YouTube videos, and Wikipedia articles relevant to the area being viewed.

This means that points of interest (POIs) can be added to Google Earth at precisely the location where the data was collected, displaying at mouse click a textbox of selected alphanumerical information and low resolution images. Since this (business) information is accessible to the general public, no confidential data are displayed, only address and customer oriented data. Zooming in or out in Google Earth allows viewing of different degrees of detail. However, the satellite images of some of the rural areas and islands in the Philippines are of low quality, sometimes blurred and the scenery covered by clouds, making it difficult to appreciate the amount of detail which has been placed on the map. Some maps are also outdated, because the images are several years old.

Google Maps

Google Maps is the sibling of Google Earth, focusing on streets and infrastructure, which allows navigation to the point of interest chosen by the user. This is particularly helpful in developing countries where there are no street addresses available for most companies and organizations, especially informal enterprises in residential areas and unplanned settlements. Similar to Google Earth, a customized layer called “*My Places*” can be activated which shows the collected set of POIs to the user.

Google Maps can be integrated into third party websites to guide customers, visitors and tourists to better find the location. Google’s search function often displays a mini map next to the search result, provided the business location is known to Google.

Box 7 Google Maps

Google Maps

From Wikipedia, the free encyclopedia (www.wikipedia.org)

Google Maps is a web mapping service application and technology that powers many map based services, including the Google Maps website, and maps embedded on third party websites via the Google Maps API.

It offers street maps, a route planner for traveling by foot, car, bike or public transport and an urban business locator for numerous countries around the world. Google Maps provides high-resolution aerial or satellite images for most urban areas all over the world.

Google Spreadsheet Mapper

Spreadsheet Mapper allows up to 400 points of interest to be organized in a spreadsheet and mapped for export to a Google Earth and Google maps compatible KML file. Using an easy to use Google Docs spreadsheet template, the resulting output is good enough for presentations and sharing with the community.

Box 8 Google Spreadsheet Mapper

Google Spreadsheet Mapper

Spreadsheet Mapper is a tool built by Google Earth Outreach to let you dynamically publish a Google Earth layer (KML or KMZ file) directly from a Google Spreadsheet. Create different balloon templates for different parts of your dataset in Spreadsheet Mapper.

From <http://www.google.com/earth/outreach/tools/index.html>

Google Map Maker

Google Map Maker is a service launched by Google in June 2008, designed to expand the breadth of the service currently offered by Google Maps. In some countries, [like the Philippines], mapping data is unavailable, and so to fill this gap Google has decided to open up Google Maps to a collaborative community effort in certain territories. The ultimate goal of the project is to acquire sufficient high-quality mapping data to be published and used on the existing Google Maps service. Some contributions have begun to appear on Google Maps, but Google Map Maker is a separate service from Google Maps, and Maps does not necessarily adopt the changes made to Map Maker and may take a long time to adopt the ones that it does (<http://en.wikipedia.org/wiki/KML>).

In the Philippines PSP SMEDSEP has partnered with the local Google Developers Group (<http://www.gtugphilippines.org>) for training its partner DTI in using Google Mapmaker with the MSME data collected via Open Data Kit.

3 Mobile Data Collection - The Bohol Experience

These elaborations are based on PSP SMEDSEP's first time experience with mobile data collection using the Android tablet computer and *ODK* software combination in analyzing the tourism value chain in the island destination of Bohol in the Philippines.

Design of Questionnaires / Data Collection Forms

As the survey forms were initially prepared in paper format the challenge was to convert it to *ODK* format for uploading to the cloud server and downloading to the Android tablet computer.

In order to convert the survey form, the GIZ technical expert Mr Arvin de la Cruz had a choice of to either (a) learning the basic XML syntax for form design and use a HTML writing / coding application, or (b) using a free cloud based survey form designer such as *ODK Build* hosted at opendatakit.org. He opted for creating the PSP data collection forms in *ODK* formbuilder. The requirements for online form creation are a reliable connection to the internet, preferably broadband; a free *Google* account, such as a Gmail account; and a basic understanding of the *ODK* form design guidelines, especially when preparing 'filter' questions.

Box 9 One view, one question paradigm of ODK

Following the '**one page, one question**' paradigm of *ODK Collect* presented another dilemma: how to convert tabular questions to electronic format?

	Male employees	Female employees
Permanent staff	?	?
Temporary staff	?	?

- How many permanent male employees do you have?
- How many permanent female employees do you have?
- How many temporary male employees do you have?
- How many temporary female employees do you have?

After designing and saving the first draft, it was exported as an .XML file for further validation and debugging using *ODK Validate*, a Java application to ensure that it is a compliant XForm that will also work with *ODK Aggregate* and *ODK Collect*. The final step was to upload it to the *ODK Aggregate* server previously set up.

Survey Preparations and Training of Enumerators

In order to implement the research PSP SMEDSEP contracted a local university familiar with the Bohol geographic and organizational landscape, the Holy Name University (HNU), based in Tagbilaran, Bohol Province. The HNU Research Center was in charge of hiring six local university students as enumerators and arranging the research logistics like planning and supervision, quality assurance, payment of fees per interview, per diems, transport and accommodation arrangements where necessary.

HNU arranged a two day training of enumerators on sampling methodology, interview techniques and use of tablets computers with *ODK Collect* software. PSP SMEDSEP provided two trainers, one for the survey methodology and interview techniques, another for the technical expertise on hardware and software for the mobile data collection.

At the beginning of the training the attention of the student enumerators was focused on the survey methodology and interview techniques. In order to avoid distraction by the gadgets, the questionnaires were explained by using a paperbased form. However, the prospective enumerators became actively involved in understanding the meaning of questions and terminology used. Together with the HNU survey coordinators they not only familiarized themselves with the survey instrument, but also contributed to its improvement, which gave them a feeling of involvement and ownership.

The trainers explained how questions were to be asked and how to record the answer. It was stressed that the enumerator must be impartial; he or she must ask the question without passing judgment. The enumerator should also not make an interpretation of the answer.

Only after the survey introduction the tablet computers were distributed among the student enumerators. An overhead projector was used to project an Android Emulator on the screen so that the technical expert was able to demonstrate every function and button on the tablet computer. Participants were given the opportunity to familiarize themselves with the tablets by going through all the hardware buttons and basic *ODK Collect* software functions. They learned how to load the survey form, swipe through the questions, input text and numbers, select single choice and multiple choice, using the skip logic and adding multiple instances (groups) to a record. The enumerators were also taught how to deal with hardware and software errors.

Then, the enumerators were grouped in pairs of two interviewing each other and recording their data using *ODK* on tablet computers. In order to get GPS coordinates they needed to go outside the building so that for the GPS receiver can make contact with satellites, which took up to 20 minutes when the tablet had not been used in the location before. After the first successful satellite fix the location recording time was much shorter.

The number of enumerators was limited by the number of tablets that the project had available and could put in the field, on the other hand it is advisable to keep the number of enumerators small to avoid too much variance in quality of interviews.

The classroom training was followed by some practical work in the field: in the case of the tourist satisfaction surveys at airport and ferry terminals in Tagbilaran, in the case of enterprise surveys the test interviews were conducted in some tourism businesses nearby the training venue.

Implementation of Tourism Data Surveys

Tablet computers loaded with *ODK Collect* software were used by PSP SMEDSEP and HNU to initially conduct two data surveys related to the tourism sector in Bohol:

Tourist Satisfaction Survey (TSS)

In order to understand the constraints of the tourism value chain in Bohol HNU enumerators interviewed 270 tourists at two exit points in Bohol: the Tagbilaran airport departure hall and the ferry departure hall in the port of Tagbilaran. Tourist visitors were identified among the travellers and were asked about their touristic experiences in Bohol. Visitors were given the opportunity to comment on tourism products and services which they had visited, like tourist attractions, accommodation establishments, restaurants, tour operators, tour guides, tourist transport, sports or recreational facilities, shops and services, airport and / or seaport infrastructure, the tourist information office, etc. Respondents were free to choose more than one type.

Subject to the type of attraction, facility or service used, the questions differed to reflect the relevant quality criteria (skip logic). The respondents were asked to select a rating on a scale from 1 - 5 (very good to very poor). Alternatively, they could state “*don't know*” or “*not applicable*.” In case that a criterion was assessed as either “*poor*” or “*very poor*” the interviewer asked about a qualitative explanation. The responses were recorded as string into the respective text fields.

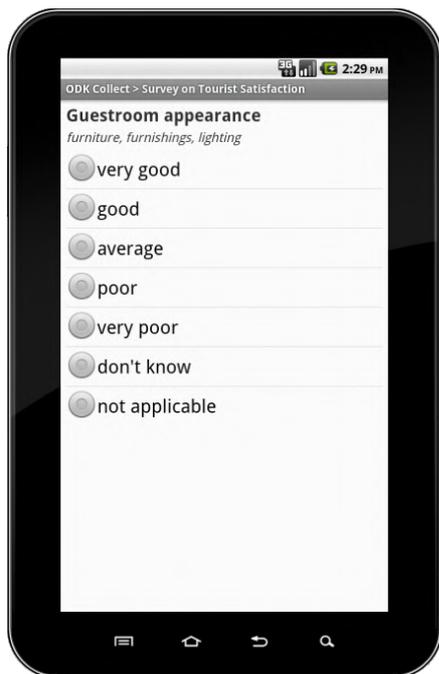


Figure 3 ODK Collect Tourism Satisfaction Survey Question

Generally, the respondents were very willing and positive towards the survey, largely attributed to the survey motive (of improving tourism quality) and the use of topnotch technology. Most interviews lasted between 15 to 30 minutes, subject to the number of attractions and services commented on. Each enumerator was able to conduct five to eight interviews per day. The whole snap survey was completed within six working days.

Tourism Enterprise Survey

In order to establish an inventory of tourism operators and destination resources in Bohol, 700 points of interests were surveyed, photographed and their GPS locations recorded. Again, using the skip logic available in *ODK Collect* the enumerators had to first determine what kind of tourism stakeholder they are about to interview. Based on that selection a different set of questions was sequenced. The survey questions were composed of text and numeric input, various groups of instances, single and multiple choice followed by conditional questions (*if ... then*), several photos to be taken with the integrated camera, and the GPS location taken from outside the premises. Each interview took about 20 minutes, subject to the availability of a competent respondent. Depending on the location and the proximity of the establishments the enumerators were able to complete an average of eight interviews per day. The survey took about three weeks to finish.

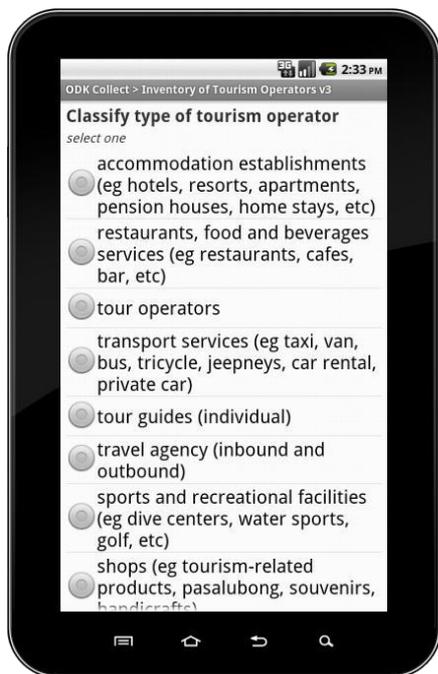


Figure 4 ODK Collect Inventory of Tourism Operators Survey Question

Except in cases where the enumerators were staying overnight in the field, they returned after work to the HNU Research Center office to connect to the internet via WLAN and upload the finalized forms to the *Google* server. In situations where no fast Internet connection was available, the large file size of high resolution photos made the upload quite time consuming and cumbersome. In case that only low resolution picture shall be used on the web, it is advisable to reduce the default file size of photos to 300dpi.

Data Analysis

TSS Analysis

After downloading the TSS data into a spreadsheet the analysis of tourist satisfaction data was performed using Microsoft Excel and / or InfoZoom. Pie or bar charts were created to communicate the frequency distribution of responses. Cross tables and bar charts were used to investigate correlations of two to three variables; such as relationships between certain opinion statements and demographic characteristics such as age, sex, nationality, etc. Qualitative explanations of negative assessments were summarized and used for the illustration of quantitative ratings.

Figure 5 TSS Spreadsheet Analysis

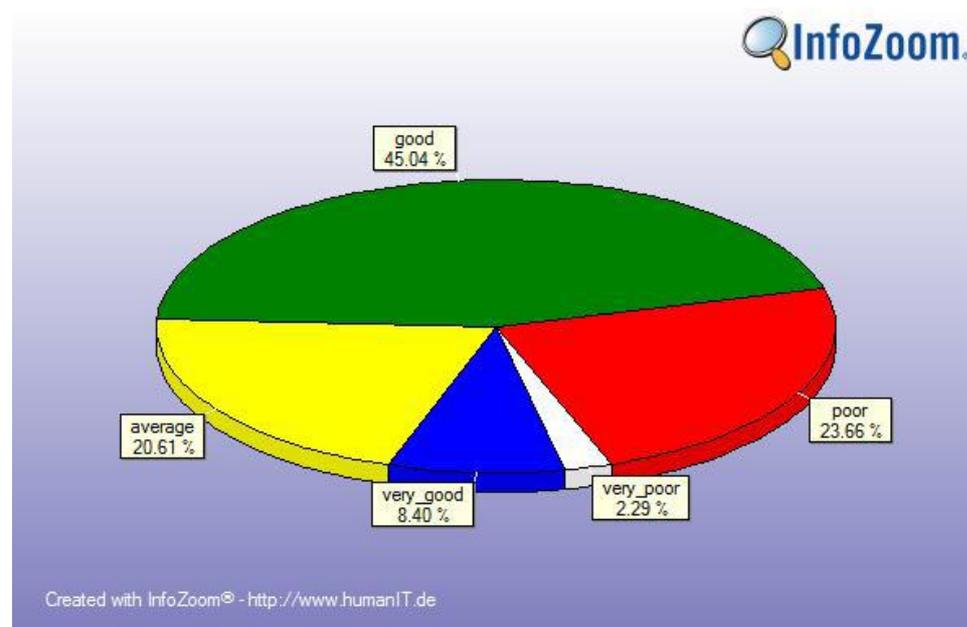


Figure 6 TSS Analysis Chart Visualization

TES Analysis

The data on tourism establishments were analyzed using Excel and / or InfoZoom to determine the aggregated figures of business indicators such as number of customers, capacities, prices, staff and availability of facilities. These figures were then used to classify the available enterprises and attractions by destination and other criteria. Key performance indicators such as sales and capacity utilization were calculated and are used for impact monitoring of value chain upgrading interventions.

Mapping of Tourism Data with Google's Geo Tools

Google Spreadsheet Mapper

The raw data collected via ODK and exported as a regular spreadsheet, was cleaned and processed into a KML file for viewing in Google Earth and Google Maps using version 3 of the Google Outreach Spreadsheet Mapper.

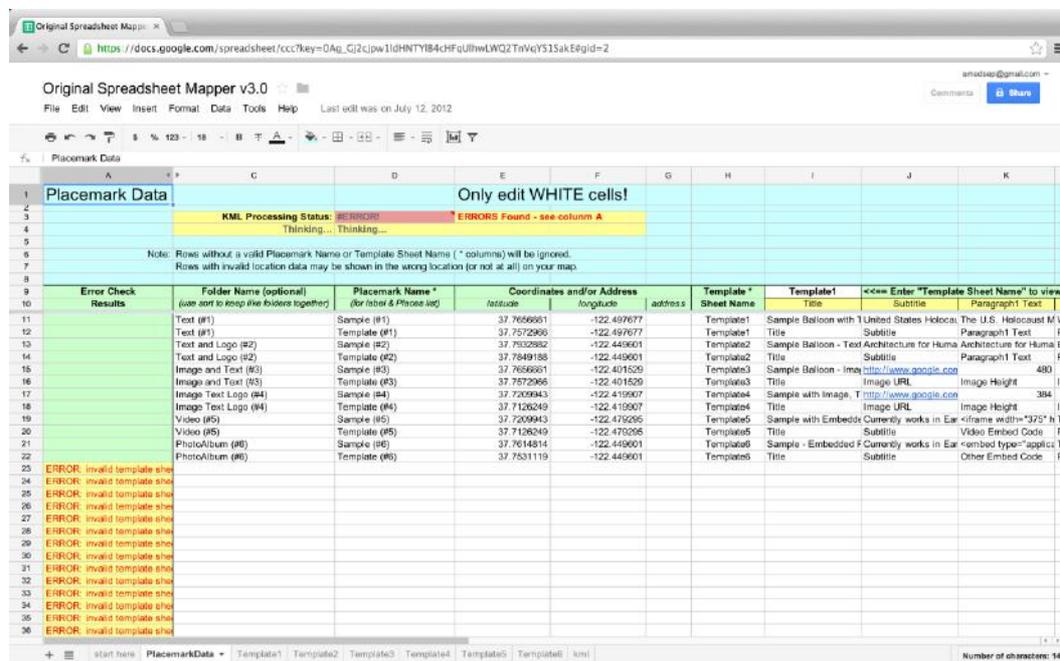


Figure 7 Original Google Spreadsheet Mapper version 3

Logging in to Google Docs, a copy of the Spreadsheet mapper template was saved and used for the exercise. On the spreadsheet template, details for the exported KML file, eg name of the organization, website and other details were inputted. Next, cleaned up or deduplicated data was copied and pasted into the mandatory fields of the spreadsheet (name, longitude, latitude). Whenever data was available, optional fields were also filled up (for example, folder names and contact information). Finally, a template for the information balloon was selected from the six predefined available, and customized for each of the points of interest. Publishing the data was completed by clicking export to KML and opening the resulting file in Google Earth. This editing cycle was repeated many times to polish and finalize the map.

See <http://www.google.com/earth/outreach/tutorials/spreadsheet3.html> for download and detailed setup instructions.

Mapping tourism locations on Google Earth and Google Maps

Even though the primary objective of the tourism enterprise survey was to build an inventory of tourism attractions and operators, such as for VC analysis and upgrading strategy, the geomapping of tourism operators on *Google Earth* was a very important side benefit. More than 700 tourist attractions and enterprises were published with selected enterprise details, photos of the establishment and located in the right place of the map.

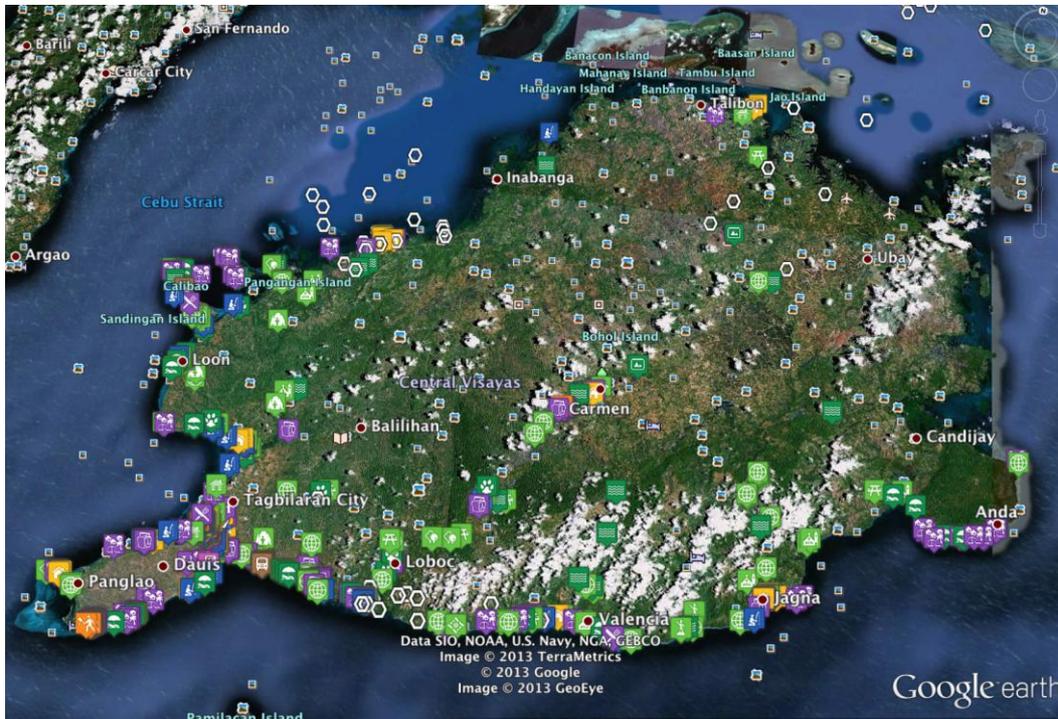


Figure 8 Bohol Tourism Map on the Google Earth App

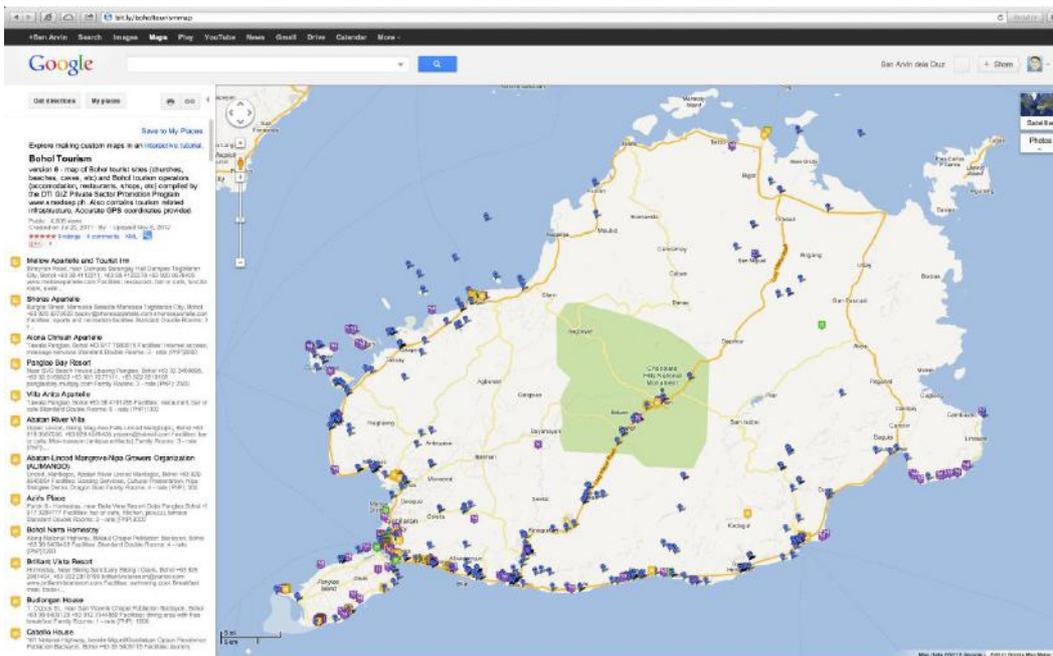


Figure 9 Bohol Tourism Map online on Google Maps

This is no uncommon marketing benefit for the somewhat larger tourism establishments in the main (sub)destinations along the prime beach areas, because they get a lot of attention from travel agencies and booking portals on the web. However, the less privileged operators in the little known pockets of tourism now have a chance to partake in the mainstream tourist activity in Bohol. More visitors to remote and less travelled attractions brings in more money into the local economy, providing income and jobs for the villagers. Local farmers can supply food and vegetables to flourishing resorts and restaurants. Local government units (LGUs) may seize the opportunity and attract additional investments into the area. As a consequence, an economic development process may be initiated with very little effort.

The enterprise mapping created exposure for many micro and small tourism operators who don't have their own websites or marketing materials. In return, the effort created a lot of goodwill among stakeholders as they received some tangible benefit for providing data to and for collaborating with the PSP SMEDSEP Program. To access the publicly accessible KML file on Google Maps please click <http://bit.ly/boholtourismmap>.

The 9Mb KMZ file that can be used with Google Earth can also be provided upon request (email Arvin dela Cruz (arvin.delacruz@giz.de) or Ms Sam Raras (rhodora.raras@giz.de)).

On the down side, the satellite images of some rural areas and islands in the Philippines are of low resolution quality, sometimes blurred and the scenery covered by clouds, making it difficult to appreciate the amount of detail which has been placed on the map. Some maps are also outdated, because the images are several years old. *Google* should be requested by the Philippine Government to purchase high resolution satellite photos for better map quality.

Google Map Maker

Putting collected MSME data on the Public *Google Map* via *Google Map Maker*

Even though the KML file via *Google Maps* can already be shared as a link to other interested parties, the best way was to embed it on the public *Google Map* (<http://maps.google.com>) itself as searchable points of interest.

The PSP Team has already included the 20 Pina loom weaving MSMEs in Aklan, Philippines in the public map, as shown on the search link results for "*Pina*" "*weaving*" "*Philippines*" below:

<http://maps.google.com/maps?q=pina%20weaving%20philippines>

In Region 8 there are also around 50+ points of interest already included in the public *Google Map*, courtesy of Ms Rhodora Raras. The rest of the collected Bohol and Visayas MSME data were distributed to the respective DTI offices for cleanup and inclusion in the public Map.

4 Assessment of the new Approach

Advantages of Mobile Data Collection using ODK

Just over a year after the new data collection approach had been introduced by PSP SMEDSEP in the Philippines, the experience has been very overwhelmingly positive and promising. The optimistic assessment is due to several factors

High speed surveys: The data collection approach using tablet computers equipped with ODK software has sped up the process from first concept note to research results. The almost immediate availability of quality raw data after the interviews has stunned even the staunchest skeptics of empirical data surveys.



Figure 10 DTI Region 6 staff practice data collection using Android tablets

Higher data quality: The predefined and formatted data fields in ODK survey forms have led to less enumeration errors. The electronic form forces the enumerators to single or multiple choice. Data fields are for either text or numeric input only. Certain fields must be answered before the enumerator can move to the next question.

Due to the immediate availability of electronic data in a structured data format there are also far less transformation errors during data capture because human errors during manual encoding are avoided.

Multimedia records: The built in camera and microphone allow for various photo, audio and video recording capabilities of electronic data collection, which no paper based survey could provide without additional recording devices. While the quality multimedia capabilities of tablet computers can still be improved, the available photos of hotels and resorts in tourism promotion indicate the future possibilities.

Locational data: The easy recording of GPS coordinates of survey units adds a whole new dimension to empirical surveys. The geomapping of locational data using GPS can help in identifying patterns of geographic distribution of survey units, such as clusters of enterprises

or self-help groups and in enabling the creation of business linkages, associations and networks. The geomapping in *Google Earth* and *Google Maps* proved to be a small but tangible benefit for the tourism stakeholders in Bohol.

Low cost surveys: A comparison of cost between the electronic data collection approach and a traditional paper-based survey demonstrates that the initial capital outlay for the purchase of tablet computers is higher, the cost for data transformation from paper to PC including necessary data cleaning has been dramatically reduced. The repeated use of tablets for mobile surveys lowered the costs even further. With technological advancement the unit cost for tablet computers will decrease even further, adding to the competitive advantages of mobile data collection.

Capacity development: The decision to contract the mobile data surveys out to a local university employing local student enumerators also contributed to building local capacity in data collection. HNU Bohol has been able to expand its services using latest technology. Upon request by various stakeholders in other regions PSP SMEDSEP with the help of local *Google Developers Groups* (Cebu, Bohol) and other private sector stakeholders (*Google Inc*, *Globe Broadband*) has conducted additional trainings for the Department of Trade and Industry (DTI) in Western Visayas, Eastern Visayas, Central Visayas, Quezon province, with the Department of Tourism (DOT) and LGUs and private stakeholders joining the mapping exercises.

Higher motivation for data surveys: Project partners, enumerators and respondents in the Philippines have been enthusiastic about the possibility to work with the electronic “gadgets.” In a tech savvy environment especially young people can be motivated to conduct the traditionally dreadful research work by using the attractive tablets. A more positive mindset of government employees towards the use of data can transform the way bureaucracies work, so that databases can become the decision making basis for policy makers and program implementers.

Replication of the approach: Due to the enthusiasm mentioned above there have been several requests from PSP SMEDSEP stakeholders to replicate and adapt the approach to different environments. Besides the tourism sector in Bohol and its suppliers, the PSP SMEDSEP Program supported the geo-mapping of creative industries in Cebu, like the furniture industries (CFIF member companies) and the producers of gift, toys and housewares (GTH). Most factory workshops and showrooms are hidden in a maze of residential dwellings and difficult to locate. Potential customers can now find them with location-based navigation.

Further replications of the approach are under way in the tourism sectors of Region 8 (Eastern Visayas), Siquijor and Negros Oriental (Region 7 Central Visayas), and the profiling of micro, small and medium enterprises in Region 6 (Western Visayas). *Google GDG* Bohol is now conducting their mapping exercise in partnership with DTI Bohol, as a direct offshoot of the previous trainings and networking effort of PSP SMEDSEP.

Positive feedback: All program partners of PSP SMEDSEP involved with electronic data surveys in the field appreciate the efficiency, speed, accuracy and usability of location based data collection methods in private sector promotion, for instance for results based monitoring, enterprise development and household surveys.

Assessment of OECD DAC Criteria

Data collection is an auxiliary process to serve development cooperation projects and programs. As such, the mobile data collection approach using tablet computers with *ODK* software will assist in implementing other development interventions that directly serve the target groups, like the value chain approach in private sector promotion or tourism development. The evaluation according to the DAC criteria shall assess the suitability of the mobile data collection approach in support of development cooperation objectives.

Relevance

Implementers of development cooperation programs always require data - for project planning, implementation and results based monitoring; by way of example private sector promotion programs need to collect enterprise data for value chain analysis and promotion; rural development practitioners need data of farming households and their agricultural production; etc. They all need socioeconomic baseline surveys for impact monitoring.

Secondary data is often not available to the right degree of disaggregation and geographical demarcation. Hence the need for primary data from the respective project or program area(s). Mobile data collection using tablet computers and *ODK Collect* software takes the hassle factor out of data collection and makes data -driven program management a possible and likely reality.

This approach can be applied in all types of development interventions where data on target groups data, intermediaries and target areas is required. It helps improve the understanding of the starting situation, the quality of implementation and the monitoring of program results.

Effectiveness

In the past 24 months (January 2011 – December 2012) several electronic data surveys have been successfully completed in the Philippines. The research results are being used in ongoing projects such as value chain interventions, tourism development, MSME promotion as well as local and regional economic development (LRED).

As much as the Private Sector Promotion (PSP) Program draws to a close by end of 2012, the data collected with the tablet computers have been effectively used as baseline information in VC strategy development, intervention design and results oriented monitoring.

Efficiency

By implementing the mobile data collection approach using tablet computers with *ODK* software researchers can generate quick results at very low cost. Researchers spend less time from data collection to data analysis because the use of technology saves both staff time and financial resources.

- Enumerators complete *ODK* form swiftly and at high data quality, eliminating the need for time consuming data cleaning
- No need for data capturers or encoders transforming data sets from paper forms to PC.
- Less costs for one time investment into tablet computers and deployment of the free open source software *ODK Collect*.

The mobile electronic data collection approach is highly efficient because of the economical use of time and financial resources.

Impact

Just 24 months after the initial deployment of tablet computers for mobile data collection in the PSP Program in the Philippines, it is still too early to determine the impact of the activity. However, the prospects are encouraging. For example Tourist Satisfaction Survey (TSS): Due to the knowledge of tourism VC opportunities and constraints, the tourism enterprises and (sub) destinations can act on upgrading interventions and are likely to enhance their competitiveness.

The knowledge of performance indicators and enterprise characteristics will help service providers in assisting the tourism sector as a whole to improve quality standards and service offerings, thus contributing to higher tourist spending (sales) and employment generation in the sector and its supply chains. The geomapping will assist in targeting the poor and disadvantaged target groups.

Sustainability

Again, it is still too early to assess the long term sustainability of the mobile data collection approach. However, judging by the degree of adoption and replication in the Philippines, one can say that it is very likely to become sustainable. Several organizations have started using the approach. Some regional and provincial DTI offices have invested in their own hardware to start the enterprise data collection in their areas.

Besides the ownership and commitment by PSP SMEDSEP's partners, the mobile data collection approach is likely to become sustainable, as the tablet technology is increasingly becoming more affordable and capable. It is also environmentally and naturally resource friendly due to the absence of paper based forms.

5 Perspectives for Mobile Data Collection and Geo Mapping

Potentials in Tourism Value Chain Development

PSP SMEDSEP has piloted the mobile data collection approach in the tourism sector, and this is where the advantages are most visible to the consumer. Tourist visitors increasingly use mapping software like *Google Earth* or *Google Maps* to plan their itineraries on tablet computers while travelling. Here the potentials are to link the points of interests (POI) on the map to web based booking portals, assisting mainly micro and small enterprises without their own websites in accessing new markets.

Another potential from the perspective of the tourist is the possibility to link the public portion of the POI datasets to navigation system manufacturers like Garmin, TomTom, Magellan, Navigon etc. So far their existing POI sets cover mostly the larger, formal, main street businesses and attractions, not the many informal tourism activities. Adding those micro operators, especially in remote areas, into the navigation map would channel more tourists into areas so far neglected by mainstream tourism, spreading the socioeconomic development effects of tourism. Small restaurants and guesthouses, but also the local transport and tour guides would benefit from increased demand for their services. This, in turn, would create additional jobs or secure existing employment and also increase the demand for food and nonfood supplies from the locality.

New mapping tools like *Google Map Maker* and *Open Street Maps (OSM)* allow for participatory adding and editing of points of interests to the standard base map. Local stakeholders and tourists alike can add or modify attractions and tourist facilities on an initially very basic destination map, gradually improving its quality and accuracy. Both programs require registration of users and approval of changes by the online mapping community in order to prevent misuse.

The empirical data on tourism attractions and enterprises will assist the tourism stakeholders in destination marketing, both online and offline. The maps can also help in destination management, such as the coordination of various products and services in the tourism destination, for example business linkages between buyers and suppliers, among accommodation establishments and local tour operators, transport providers and tourist attractions. Destination planners will have a visual basis to plan for an (expanded) tourism infrastructure, transport systems and utilities like electricity, fresh water, solid waste, and wastewater disposal.

Data on the resource efficiency (energy, water, waste) of tourism enterprises can assist in the greening of the tourism destination, which is particularly important for sustainable forms of tourism like ecotourism in ecologically sensitive environments. Geospatial data on nature conservation and biodiversity can help protect the pristine environment, which tourism requires as its most important resource.

Tablet computers with *ODK Collect* software could also be used by auditors accredited by the Department of Tourism to assess and record the compliance of tourism enterprises with the new national accommodation standards, which will be implemented from 2012 onwards. The assessor would use the tablet while walking through the establishment and checking the facilities. The complete standards catalogue could be implemented in *ODK*, possibly also allowing for an automatic calculation of the score that will determine the star rating of the

establishment. Ideally, a summary report should be generated showing the areas of improvement that an enterprise should focus on in order to achieve a higher star rating.

Finally, tourism or destination management offices as well as business associations should continue to conduct tourist satisfaction surveys on a regular, say 6-monthly, basis. These surveys will provide necessary intelligence on tourism trends and developments, allowing for better destination management as well as tourism statistics, on the basis of which the Tourism Satellite Account (TSA) can be compiled in compliance with United Nations World Tourism Organization (UN WTO) standards.

Potentials in MSME Development

While the mobile data collection approach fits particularly well in the tourism sector, the question arose soon: Why not also use the approach in other fields of micro, small and medium enterprise (MSME) development? Again, existing secondary data are scarce. Support agencies often shy away from collecting primary business data. Small business advisors usually lack the empirical basis for providing business development services on an informed manner, unless they create a baseline of key performance indicators.

Using only 17 Android tablets, six simultaneous MSME monitoring surveys were conducted by the PSP SMEDSEP team during the second quarter of 2012 covering the entire Visayas regions (Region 6, Region 7, Region 8), collecting data from more than 2000+ MSMEs, LGU stakeholders and opinion leaders.

One such initiative to establish a systematic MSME database is undertaken by the Department of Trade and Industry (DTI) in Region 6 (Western Visayas) of the Philippines, namely the provinces of Iloilo, Capiz, Aklan, Antique, Guimaras and Negros Occidental. Starting from a situation where an enterprise database does not exist in the region, DTI now deploys the mobile data collection approach using tablets with *ODK Collect* to collect business profiles of manufacturing and service enterprises.



Figure 11 DTI and other partners in Central Visayas are oriented in data collection using ODK

In order to standardize the coding of business activities, the *ODK* form implements a down-scaled version of the International Standard Industrial Classification (ISIC). The enumerator will make a single choice on the main sector, then opens up a list of subsectors in that main sector, and so on. This is applied up to the 4th level of the ISIC system, which allows a detailed description of the business subsector concerned. The MSME form collects further key business figures and parameters, such as ...products, investment, sales, employment, services used, etc. Resource auditors could record data on energy and water consumption into tablet computers, collecting information allowing for green economic development in the respective industry subsector.

These data will allow business counselors to provide more effective services to the MSMEs surveyed and monitor the results of same. Such services may include business planning, technical or production-related assistance, quality assurance, financial management, marketing, etc. Ideally, the business profiles could be updated and retrieved on the same tablet computer or handheld device that was used to collect the information in the first place. This would permit effective business counseling, but also enable **'before and after'** comparisons for monitoring of results. However, there are issues of privacy and data security that remains to be resolved.

With the GPS function of the tablet computer enabled, the *ODK Collect* form will record the location of an enterprise which can be mapped on *Google Earth / Google Maps*, etc. The respective kmz-file acts like a business layer on the map and can show enterprises by certain sectors or other categories. This allows for identifying sector clusters and may facilitate business linkages along the value chain, like buyer - supplier linkages. The map may also reveal the availability and geographical distribution of business development service (BDS) providers in relation to their MSME clients.

Potentials for Market Surveys and Customer Insights

Many companies and stakeholders in private sector development conduct market surveys for their clients, such as small and medium enterprises. In order to gauge market potential or to measure brand perception they conduct (or contract out) market surveys among existing or prospective customers. Like in the case of tourism sector mentioned above, enumerators measure consumer satisfaction, aim at getting feedback on new or existing products or try to understand the purchasing behavior of their customers. All such surveys attempt to discover the key influences and perceptions of their clientele.

The mobile data collection approach using the tablet computers is a cost efficient alternative to expensive market research. It can be implemented by the principal himself or herself in very short period of time, or to be contracted out at very low cost.

Potentials in Rural / Agricultural Development

Projects and programs supporting rural development typically work with rural farming communities and support agricultural production, micro finance and income generating activities. They assist agricultural producer groups, cooperatives, farmer's associations, savings and credit organizations and work through commercial, governmental or nonprofit service providers.

In order to design their interventions, program implementers require household data, income and expenditure records, agricultural (or fishery) production figures, association

membership, loan records etc, both for implementation and for monitoring. Since rural households are often dispersed in remote areas, field workers can use tablet computers to aid their extension work. GPS coordinates may play an important role in mapping and visualizing the rural environment. Existing GIS data on agricultural production systems can also be integrated into *Google's* geo tools.

A case for low cost use of GPS data is the mapping of bamboo stands in the Province of Capiz, Philippines, where consultants hired by PSP SMEDSEP assisted the provincial DTI office in assessing the size and accessibility of the bamboo growing area, an information which is a vital precondition for investments into engineered (laminated) bamboo production.

Potentials in Environment and Natural Resources Management

In the context of climate change and environmental degradation there are many development programs assisting vulnerable communities with climate change adaptation, integrated coastal zone management, flood protection, community forestry and rural livelihoods in geographical areas under threat. Implementers of natural resource management programs need data on the environmental and biodiversity conditions in vulnerable areas, such as on endangered species.

Field workers could combine participatory rural appraisal techniques, like transect walks with local community members and the GPS function in *ODK Collect*. The data can be mapped on dedicated layers in *Google Earth*, displaying critical environmental information on the virtual globe. In addition, *Google Earth* allows for the import of existing GIS data into its maps. Such data may also be helpful for disaster and risk management as well as microinsurance programs (like crop insurance).

Potentials in Social Development, Gender and Health

The mobile data collection approach would also be an innovation for the household surveys conducted in social development programs focusing on health, gender equality, family planning, or housing, etc. Most of these surveys are still paper based, like the USAID funded Demographic and Health Survey (DHS), which has been implemented in about 120 countries worldwide. The DHS takes about 18 - 24 months from research design to data analysis, a period, which could be significantly reduced through the deployment of the mobile data collection approach.

Now, *ODK Collect* is increasingly used in healthcare projects in Africa and India. In addition to its efficiency, the GPS mapping enables the visualization of geographical patterns of demographic, socioeconomic or health status of household respondents. The locational data also permit an adequate provision of basic infrastructure and public services - water, electricity, transport, telecommunications, etc. Finally, geo mapping will be helpful in identifying migratory movements as well as the geographical patterns (and possible causes) of peace and conflict.

Potentials in Decentralization and Local Governance Programs

The mobile data collection approach has potential in an urban context, such as in housing and infrastructure development and environmental management. Geo mapping can visualize urbanization patterns and environmental hazards. It may assist public and private stakeholders in the planning of local and regional economic development (LRED) projects.

The PSP SMEDSEP has already conducted the Rapid Economic Assessment (REA) survey in *ODK* format in 2012. LRED REA trainings are also underway.

Mobile data collection using *ODK* with subsequent geo mapping can assist local government units (LGUs) in tax collection and the implementation of regulations, for example the business permits and licensing system (BPLS) in the Philippines.

Potentials in Results based Monitoring (RBM)

Probably one of the biggest potentials for the mobile data collection approach lies in results based monitoring (RBM), which is a mandatory requirement for almost all development projects and programs. The program implementers' responsibility is not primarily the process monitoring of outputs and outcomes, but the monitoring of the indirect benefits and impact beyond the so-called attribution gap. The results chains of most projects and programs aim at impact and target group level. This means that project and program implementers must provide quantitative facts and figures on the ultimate beneficiaries.

Baseline surveys need to be conducted to permit before and after comparisons. Impact monitoring and evaluation guidelines often require experimental or at least quasi experimental evaluation design involving randomized control groups to causally attribute positive changes to project/program interventions; like to prove impact. This document shall not discuss the pros and cons of rigorous evaluation designs using randomized controlled trials (RCT). However, even simple before and after comparisons require quantitative data that need to be established in empirical baseline and follow-up surveys. Plausible attribution may require additional qualitative methods, e.g. focused group discussions with target group representatives. However, even without the efforts and costs of a scientific research design implementers can generate reliable data of the target group progress, after all: rapid survey data are better than no data at all.

A *quick and dirty* survey design using the mobile data collection approach seems to be an acceptable middle ground between complete data ignorance and a situation of scientific *analysis paralysis*, thus still providing empirical insights at low costs in a short time, when data-supported decisions are needed for project or program steering and results based management. Such data collection may be done by the project or program staff themselves. It could be outsourced to local research centers, universities or freelance consultants.

The mobile data collection approach with tablet hardware and *ODK* software makes it easy and efficient to generate snapshots of target group status, and combine them to allow time series analysis. The GPS function permits the enumerator to easily locate households, groups or enterprises for follow-up surveys. The **before-and-after** comparison can also be aided by utilizing the integrated camera for photo monitoring of household or enterprise status.

The geo mapping of households, groups or enterprises can visualize the spatial distribution of target groups and control groups, and may explain development patterns, which may address the question "*what interventions work where and why?*" Similarly, the size and extent of agricultural production areas, forests, housing, infrastructure, etc can be established and changes at any point in time could be recorded on the map through the historical layer function in *Google Earth*.

The statistical analysis of alphanumerical data does not require a trained statistician and complex statistics software. RBM is not about proving or refuting of scientific research hypotheses, but about intuitively exploring relationships among different variables and presenting the results in quick reports and graphical charts. For this purpose, data mining tools like *Google Refine* or *InfoZoom* are better suited than statistical software like SPSS or Stata, because even a lay person can conduct an ad hoc analysis through intuitive data visualization technology.

Project or program managers require a management information system that provides standardized reports on the achievement of all major indicators at various levels of the results chain, from project or program outputs to their utilization by intermediaries, to outcomes or direct benefits for target groups and beyond the attribution gap to indirect results and highly aggregated impact. For this purpose a graphical front-end presentation tool can be developed, like a customized dashboard or management cockpit.

Specialized software developers have integrated the *ODK Collect* software with the '*force.com*' platform to develop monitoring and evaluation systems for non profit organizations. Besides a customized management dashboard the system allows for analysis and reporting of data.

6 Conclusions and Recommendations

The mobile data collection approach using tablet computers with *ODK Collect* has proven to be a very useful tool for different types of surveys mainly in private sector development and tourism value chain promotion. However, the approach seems to have significant application potential in many other sectors of development cooperation, and especially in intervention planning, implementation and results-based monitoring. If good and timely data at little cost is the goal, then tablets loaded with *ODK* software are suitable tools.

Furthermore, the approach using stylish gadgets for empirical surveys was appealing to many program and partner staff who otherwise would find data collection a boring and tedious undertaking. On the other hand, the efficiency and cost effectiveness of the approach will be appealing also for managers who are concerned about the typical time consuming efforts to build data “graveyards.”

The location mapping in *Google’s* geo tools (*Google Earth* and *Google Maps*) adds a new dimension of opportunities for stakeholder cooperation, business linkages, business service provision, investment promotion, direct sales and marketing, destination management, good governance, environmental and biodiversity protection, and ultimately, pro poor services and poverty eradication.

The innovations of the mobile data collection approach will probably contribute to more and meaningful data driven decision making in government agencies, the private sector and in NGOs. The approach needs to be further integrated into a comprehensive management information system that provides decision makers with a condensed overview over developments in the field.

It is recommended that GIZ and / or other technical cooperation agencies provide funding to support further technical development and customization for different project environments.

Contact: Uwe Sturmman

uwe.sturmann@gmail.com

Annexes

Annex 1 Glossary:

Attribute, variable In science and research, attribute is a characteristic of an object (person, thing, etc.). While an attribute is often intuitive, the variable is the operationalized way in which the attribute is represented for further data processing. In data processing data are often represented by a combination of items (objects organized in rows), and multiple variables (organized in columns).

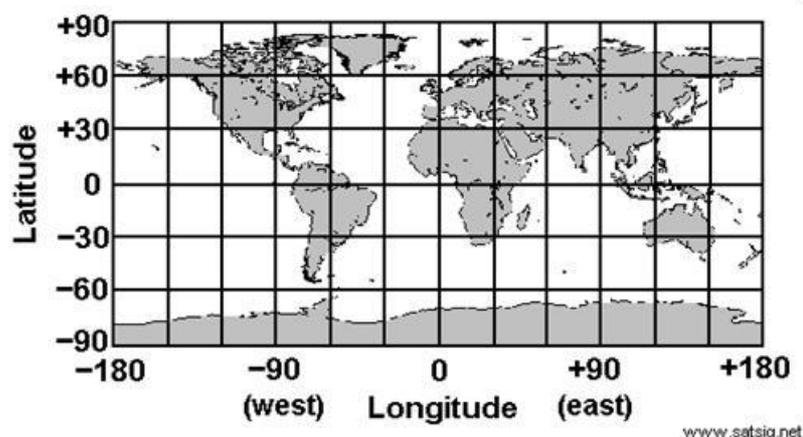
From [http://en.wikipedia.org/wiki/Variable_and_attribute_\(research\)](http://en.wikipedia.org/wiki/Variable_and_attribute_(research))

Data fields containing values A data field is a place where you can store data. Commonly used to refer to a column in a database or a field in a data entry form or web form. The field may contain data to be entered as well as data to be displayed

From http://en.wikipedia.org/wiki/Data_field

Geo tagging **Geo tagging** is the process of adding geographical identification metadata to various media such as geo-tagged photos, videos, datasets, websites, etc. These data usually consist of latitude and longitude coordinates, though they can also include altitude, bearing, distance, accuracy data, and place names. Geo tagging can help users find a wide variety of location specific information.

geographic coordinate system A **geographic coordinate system** is a coordinate system that enables every location on the Earth to be specified by a set of numbers. The coordinates are often chosen such that one of the numbers represents the vertical position, and two or three of the numbers represent horizontal position. A common choice of coordinates is latitude, longitude and elevation.



GIS

GIS - A **geographic information system** is a system designed to capture, store, manipulate, analyze, manage, and present all types of geographical data. In the simplest terms, GIS is the merging of cartography, statistical analysis and database technology. Usually, a GIS is custom designed for an organization, but describes any information system that integrates, stores, edits, analyzes, shares and displays geographic information for informing decision making.

GIS, or Geographic al Information System,

http://en.wikipedia.org/wiki/Geographic_information_system

Geographic information system (GIS) is a system designed to capture, store, manipulate, analyze, manage, and present all types of geographical data. The acronym GIS is sometimes used for geographical information science or geospatial information studies to refer to the academic discipline or career of working with geographic information systems.[1] In the simplest terms, GIS is the merging of cartography, statistical analysis, and database technology.

A GIS can be thought of as a system—it digitally creates and "manipulates" spatial areas that may be jurisdictional, purpose, or application-oriented. Generally, a GIS is custom-designed for an organization. Hence, a GIS developed for an application, jurisdiction, enterprise, or purpose may not be necessarily interoperable or compatible with a GIS that has been developed for some other application, jurisdiction, enterprise, or purpose. What goes beyond a GIS is a spatial data infrastructure, a concept that has no such restrictive boundaries.

In a general sense, the term describes any information system that integrates, stores, edits, analyzes, shares, and displays geographic information for informing decision making. GIS applications are tools that allow users to create interactive queries (user-created searches), analyze spatial information, edit data in maps, and present the results of all these operations.[2] Geographic information science is the science underlying geographic concepts, applications, and systems.[3]

Google Earth

Google Earth is a virtual globe, map and geographical information program that maps the Earth by the superimposition of images obtained from satellite imagery, aerial photography and GIS 3D globe. As of October 2011 *Google Earth* has been downloaded more than a billion times. *Google Earth* is useful for many day-to-day purposes. The user can explore and place location bookmarks in literally every corner of the globe. One can also get directions using *Google Earth*, using variables such as street names, cities, and establishments.

Google Earth can also function as a hub of knowledge, depending on the user's location. By enabling certain options, one can see the location of gas stations, restaurants, museums, and other public establishments in their area. *Google Earth* can also dot the map with links to images, YouTube videos, and other media relevant to the area being viewed.

Google Earth Engine	Google Earth Engine - A platform for environmental data and analysis, <i>Earth Engine</i> brings together the world's satellite imagery — trillions of scientific measurements dating back more than 25 years — and makes it available online with tools for scientists, independent researchers, and nations to mine this massive warehouse of data to detect changes, map trends and quantify differences on the earth's surface. e.g. deforestation and land use trends.
Google Earth Pro	<p>Google Earth Pro (US\$400) includes the same easy-to-use features and imagery as the free version of Google Earth, but with additional professional tools designed specifically for business users.</p> <ul style="list-style-type: none"> • Utilize data layers to locate your target demographic • Compute distances and areas using measurement tools • Use Movie Maker to produce media collateral • Print high-resolution images for presentations and reports • Import large vector image files to quickly map GIS data • Map addresses with the Spreadsheet Importer
Google Map Maker	Google Map Maker is a service launched by Google in June 2008,[1] designed to expand the breadth of the service currently offered by Google Maps. In some countries, [like the Philippines], mapping data is unavailable, and so to fill this void Google has decided to open up Google Maps to a collaborative community effort in certain territories. The ultimate goal of the project is to acquire sufficient high-quality mapping data to be published and used on the existing Google Maps service.
Google Maps	Google Maps is a scalable, powerful and secure platform for building and publishing your maps using the power of the Google cloud. Realize the full potential of your geospatial data by securely sharing your maps with anyone over a variety of devices.

Google Refine

Google Refine is a standalone desktop application provided by Google for data cleanup and transformation to other formats. It is similar to spreadsheet applications (and can work with spreadsheet file formats), however acts more like database. It operates on *rows* of data which have cells under *columns*, which is very similar to relational database tables. Users can filter rows to display using *facets* that define filtering criteria. Unlike spreadsheets, most operations in Refine are done on all visible rows: transformation of all cells in all rows under one column, creation of new column based on existing column data, etc. All actions that were done on dataset are stored in project and can be replayed on another dataset. Unlike spreadsheets, no formulas are stored in cells, but formulas are used to transform data, and transformation is done only once. Possible uses of *Google Refine* software

- *Cleaning messy data*: a text file with some semi-structured data can be edited it using transformations, facets and clustering to make data cleanly structured.
- *Transformation of data*: converting values to other formats.
- Parsing data from web sites: Google Refine contains URL fetch feature.
- *Adding data to dataset by fetching it from web services*; e.g. for geocoding addresses to geographic coordinates.
- Augmentation of datasets with data from Freebase.

Google Spreadsheet Mapper

Google Spreadsheet Mapper is a tool built by Google Earth Outreach to let you dynamically publish a Google Earth layer (KML or KMZ file) directly from a Google Spreadsheet. Create different balloon templates for different parts of your dataset in Spreadsheet Mapper.

GPS navigation device

A **GPS navigation device** is any device that receives Global Positioning System (GPS) signals for the purpose of determining the device's current location on Earth. GPS devices provide latitude, longitude and altitude information. They may also have additional capabilities such as:

- containing all types of maps, like streets maps, which may be displayed in human readable format via text or in a graphical format
- providing suggested directions to a human in charge of a vehicle or vessel via text or speech
- providing information on traffic conditions (either via historical or real time data) and suggesting alternative directions;
- providing information on nearby amenities such as restaurants, shops, petrol stations.

**GPS, or
Global
Positioning
System**

GPS, or Global Positioning System, is a "constellation" of 24 satellites that orbit the Earth and make it possible for people with ground receivers to pinpoint their geographic location. 21 GPS satellites and three spare satellites are in orbit at 10,600 miles above the Earth. The satellites are spaced so that from any point on Earth, four satellites will be above the horizon. On the ground, any GPS receiver contains a computer that "triangulates" its own position by getting bearings from three of the four satellites. The result is provided in the form of a geographic position - longitude and latitude – to within 100 to 10 meters accuracy.

If the receiver is also equipped with a display screen that shows a map, the position can be shown on the map. If a fourth satellite can be received, the receiver/computer can figure out the altitude as well as the geographic position. If a vehicle is moving, the GPS receiver may also be able to calculate its speed and direction of travel and provide the estimated times of arrival to specified destinations. GPS receivers are becoming consumer products. In addition to their outdoor use (hiking, cross-country skiing, ballooning, flying, and sailing), receivers can be used in cars to relate the driver's location with traffic and weather information.

Latitude

Latitude is a geographic coordinate that specifies the north-south position of a point on the Earth's surface. Lines of constant latitude, or parallels, run east–west as circles parallel to the equator. Latitude is an angle (defined below) which ranges from 0° at the Equator to 90° (North or South) at the poles.

Longitude

Longitude is a geographic coordinate that specifies the east-west position of a point on the Earth's surface. Points with the same longitude lie in lines running from the North Pole to the South Pole. By convention, one of these, the Prime Meridian, which passes through the Royal Observatory, Greenwich, England, establishes the position of zero degrees longitude. The longitude of other places is measured as an angle east or west from the Prime Meridian, ranging from 0° at the Prime Meridian to +180° eastward and –180° westward.

Latitude and longitude together are used to specify the precise location of features on the surface of the Earth.

**ODK
Aggregate**

ODK Aggregate provides a ready-to-deploy server and data repository to:

- provide blank forms to *ODK Collect* (or other OpenRosa clients),
- accept finalized forms (submissions) from *ODK Collect* and manage collected data,
- visualize the collected data using maps and simple graphs,
- export data (e.g., as CSV files for spreadsheets, or as KML files for *Google Earth*), and
- publish data to external systems (e.g., Google Spreadsheets or *Google Fusion Tables*).

ODK Aggregate can be deployed on Google's App Engine, enabling users to quickly get running without facing the complexities of setting up their own scalable web service.

ODK Build

ODK Build is a form designer with a drag-and-drop user interface. Build is an HTML5 web application and works best for designing simple forms.

ODK Collect

ODK Collect renders forms into a sequence of input prompts that apply form logic, entry constraints, and repeating sub-structures. Users work through the prompts and can save the submission at any point. Finalized submissions can be sent to (and new forms downloaded from) a server. *ODK Collect* uses the Android platform on mobile devices, supports a wide variety of prompts (text, number, location, multimedia, barcodes), and works well without network connectivity.

**Open Data
Kit (ODK)**

Open Data Kit (**ODK**) is a suite of tools that allows data collection using mobile devices and data submission to an online server, even without an Internet connection or mobile carrier service at the time of data collection. You can collect data remotely without an Internet connection or cell carrier access. Gather text, numeric data, media and more with a mobile device. Then, host your data online using Google's powerful hosting platform, *Google AppEngine*, and visualize your data as a map using *Google Fusion Tables* and *Google Earth*.

**Open Street
Map (OSM)**

Open Street Map (OSM) is a collaborative project to create a free editable map of the world. Two major driving forces behind the establishment and growth of OSM have been (a) restrictions on use or availability of map information across much of the world and (b) the advent of inexpensive portable satellite navigation devices. The maps are created using data from portable GPS devices, aerial photography, other free sources or simply from local knowledge. The OpenStreetMap approach to mapping was inspired by sites such as Wikipedia; the map display features a prominent "Edit" link and a full revision history is maintained. Registered users can upload GPS track logs and edit the vector data. Various mobile applications also allow contribution of GPX tracks to the OSM project.

Question types**Question types**

- Multiple Choice
- Single Choice
- Rating Scale
- Text Boxes
- Demographics
- Date

Record

In the context of a relational database, a row—also called a record or tuple—represents a single, implicitly structured data item in a table. In simple terms, a database table can be thought of as consisting of rows and columns or fields. Each row in a table represents a set of related data, and every row in the table has the same structure.

From [http://en.wikipedia.org/wiki/Row_\(database\)](http://en.wikipedia.org/wiki/Row_(database))

Sample survey vs census

A census measures absolutely everyone in the whole country. A representative sample measures a small number of people who fit a particular category of people: surveying 200 white male smokers between 20 and 40 who have had surgery for throat cancer (out of a total number of say 35,000 men who fit that profile).

From

http://wiki.answers.com/Q/What_are_the_differences_between_census_and_sample_survey

Skip logic

Skip logic, or conditional branching, allows to change the course that respondents take through a survey based on answers they give to certain questions. This is implemented by creating skip rules. For example, if you were creating a survey, and you were to ask your respondents if they are male or female. You could create some skip rules based on this, that would then direct them to a certain page based on their response. You could have a page of questions directed at your male respondents and a page of questions directed at your female respondents, and could then direct your respondent to one of these pages and hide the other based on their response to the initial male/female question. There are a number of reasons why you should use skip logic. If you have questions that are only related to some of your respondents, then you would not want your other respondents to be presented with these questions, as they will get frustrated when asked to answer questions that do not relate to them.

Annex B Tablet Computers

http://en.wikipedia.org/wiki/Tablet_computer

A tablet computer, or simply tablet, is a mobile computer, larger than a mobile phone or personal digital assistant, integrated into a flat touch screen and primarily operated by touching the screen and instead of using a physical keyboard it often uses an onscreen virtual keyboard.[1][2][3]

A tablet computer may be connected to a keyboard with a wireless link or a USB port. Convertible notebook computers have an integrated keyboard that can be hidden by a swivel joint or slide joint, exposing only the screen for touch operation. Hybrids have a detachable keyboard so that the touch screen can be used as a stand-alone tablet. Booklets include dual-touchscreens, and can be used as a notebook by displaying a virtual keyboard in one of them.

An early information tablet concept was Alan Kay's Dynabook in 1968. The first commercial portable electronic tablets appeared at the end of the 20th century. In 2010, Apple Inc. released the iPad which became the first mobile computer tablet to achieve worldwide commercial success. The iPad used technology similar to Apple's iPhone. Other manufacturers have produced tablets of their own including Samsung, HTC, Motorola, RIM, Sony, Amazon, HP, Microsoft, and Archos. Tablets use a variety of operating systems such as iOS (Apple), Android (Google), Windows (Microsoft), and QNX (RIM).

As of March 2012, 31% of U.S. Internet users were reported to have a tablet, which was used mainly for viewing published content such as video and news.[4] Among tablets available in the market in 2012, the top-selling device is Apple's iPad with 100 million sold by mid October 2012 since it was released in April 3, 2010,[5] followed by Amazon's Kindle Fire with 7 million, and Barnes & Noble's Nook with 5 million.[6]

The Evolution of the Tablet Computer Infographic

<http://visual.ly/evolution-tablet-computer>

Annex C Hardware Preparation

How to get the Android phones ready to be used as data collection instruments by field enumerators. These are the things I do to the Androids to prepare them for use solely as data collection devices.

- Uninstall old versions of ODK
- Update Android operating system and ODK Collect to their latest versions
- Place the icon on the desktop
- Remove all non-relevant icons and widgets.
- Set the GPS to ON, WiFi off, data sync off, and destroy the widget. That way, no one can turn off GPS easily.
- Set screen time-out to be long for training, short for the field
- Airplane mode! saves battery.
- Set Language to the native language of the enumerator
- Turn off automatic orientation. Rotating the screen is no benefit in ODK.
- Set the keyboard to be Touch Input and change the settings to remove all suggestions and spell checking. Turn off vibrate, turn on sound.
- Calibrate the keyboard
- Copy the current survey(s) to /sdcard/odk/forms using the Android Debug Bridge.
- Load the form the first time to make sure it's loaded, it's valid, and save time later. It loads much faster the second time.
- Get a Good GPS fix, the almanac can take a while to ready itself. The first time you use GPS in a strange place, it can take 20 minutes to load. Don't make the enumerators do that. Getting a GPS fix on all the Androids updates the almanac data and the next fix in the same city will be way faster.



Annex D Open Data Kit

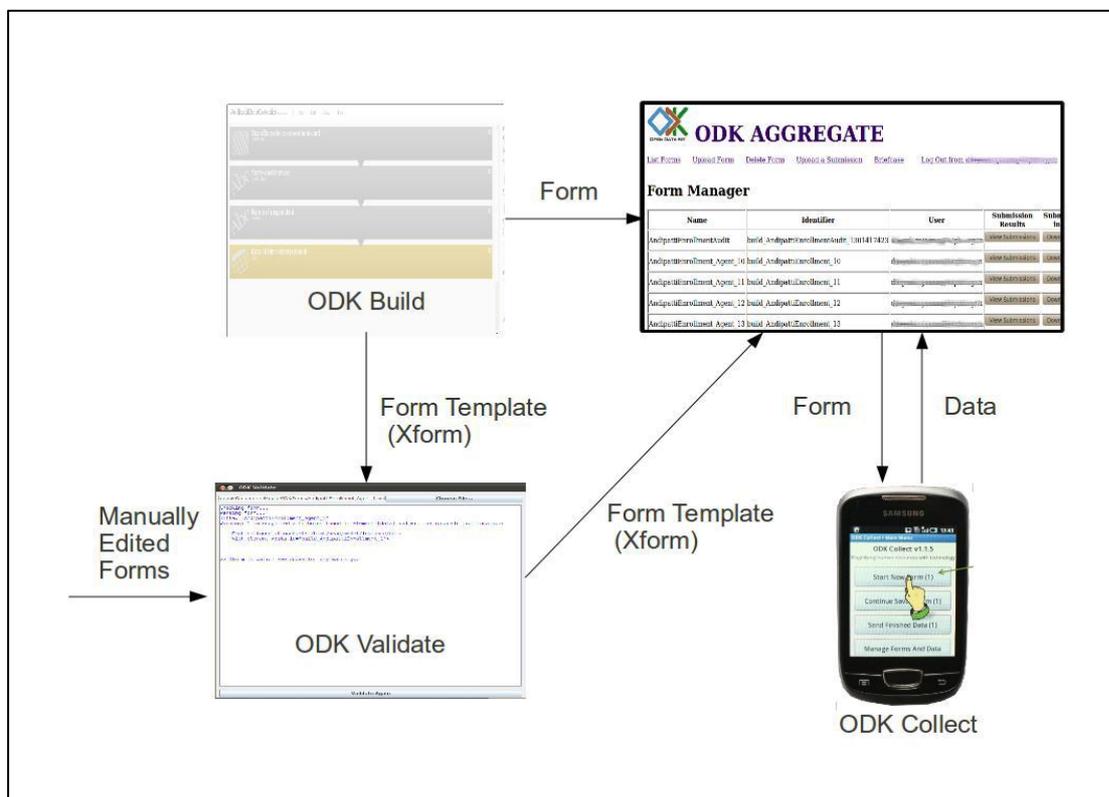
Open Data Kit (ODK) is an open-source suite of tools that helps organizations author, field, and manage mobile data collection solutions. Our goals are to make open-source and standards-based tools which are easy to try, easy to use, easy to modify and easy to scale.

Tools

ODK is funded by a Google Focused Research Award and supported by a growing community of developers, implementers and users.

Below are the released and supported ODK projects.

- **Build** - *ODK Build* enables users to generate forms using a drag-and-drop form designer. Build is implemented as an HTML5 web-based application and targets the common use case of a simple form.
- **Collect** - *ODK Collect* is powerful phone-based replacement for your paper forms. Collect is built on the Android platform and can collect a variety of form data types: text, location, photos, video, audio, and barcodes.
- **Aggregate** - *ODK Aggregate* provides a ready to deploy online repository to store, view and export collected data. Aggregate can run on Google's reliable and free infrastructure as well as on local servers backed by MySQL and PostgreSQL.
- **Briefcase** - *ODK Briefcase* is the best way to transfer data from Collect and Aggregate.
- **Validate** - *ODK Validate* ensures that you have a OpenRosa compliant form -- one that will also work with all the ODK tools.
- **XLS2XForm** - *ODK XLS2XForm* allow XForms to be designed with Excel.



Annex E Google Map Maker

From <http://support.google.com/mapmaker/bin/answer.py?hl=en&answer=155420> below are the steps to use the publicly accessible KML file on *Google Maps* as an overlay for Google Map Maker:

0. Make sure you have a *Google Account* and are logged on to <http://maps.google.com> service
1. **Enable Overlays**

First, make sure that Overlays are enabled on Map Maker. To do this, access the "Labs" by clicking the green labs icon on the top right-hand corner of the screen. Then make sure that "Overlays" is switched to "Enabled" and click "Save". The "Overlays" button should now appear before the "Map" and "Satellite" buttons on the viewport.
2. **Find or Make Content**

To access KML maps or Images in *Google Map Maker* with the Overlay feature, this content will need to be online and you will need its web address. You can use <http://bit.ly/boholtourismmap> as the overlay for the Bohol MSME survey data
3. **Add Content to Map Maker**

Once you have a web address for the content (eg <http://bit.ly/boholtourismmap>) you'd like to see in *Google Map Maker*, you can add this web address to your Overlays. To do this, first go to the *Google Map Maker* homepage: <http://mapmaker.google.com> and sign in using your *Google Account*. Next, click on the "Overlays" tab on the upper-right hand corner of the map view. (see image below)



Click the "Manage overlays" button. This will open a window where you can add KML and Image web addresses. For this tutorial, copy the following Panoramio KML url text: www.panoramio.com/kml enter it under the section that says "KML:" in the dialog box, and click "Add". (see image below) When you are finished managing your overlays, simply click "OK".

Add an overlay ✕

KML:

Add

Enter a KML URL

Image:

Add

Enter a image URL

OK

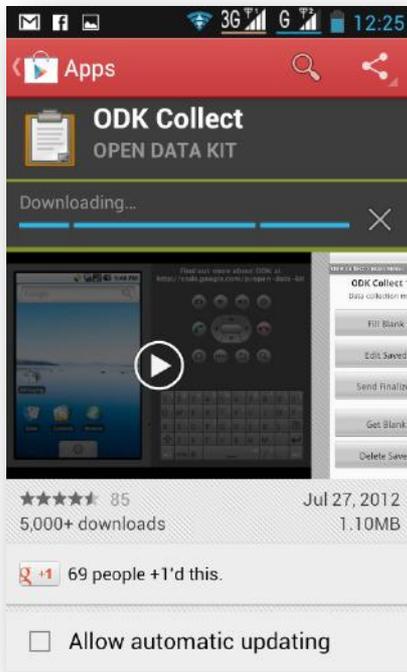
Mapping with new Data

Overlays will only be kept while you have *Google Map Maker* open your web browser as they are not stored with your user account. If you close or refresh your browser, you will need to repeat the steps above again to see your Overlays.

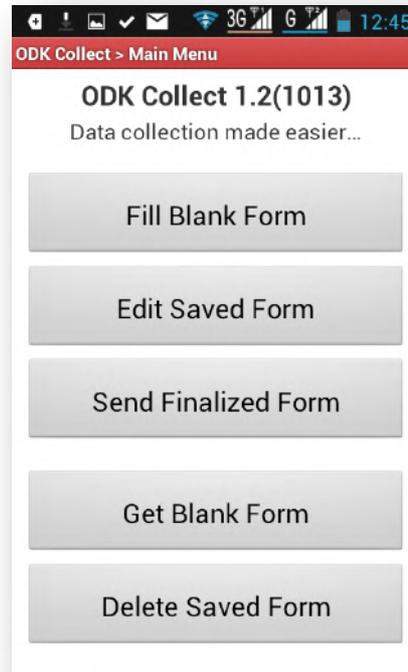
Click the "Overlay" button to turn on and off the overlays. Click the "Manage overlays" link again to add and remove your overlays. If you want to add more URLs, you can continue to enter them one at a time under the appropriate "KML" or "Image" field and click "Add". To remove a link, simply click on the boxed "X" next to the URL you'd like to remove. Connecting to too many links may slow down your Map Maker connection.

Annex F ODK Collect – Screenshots

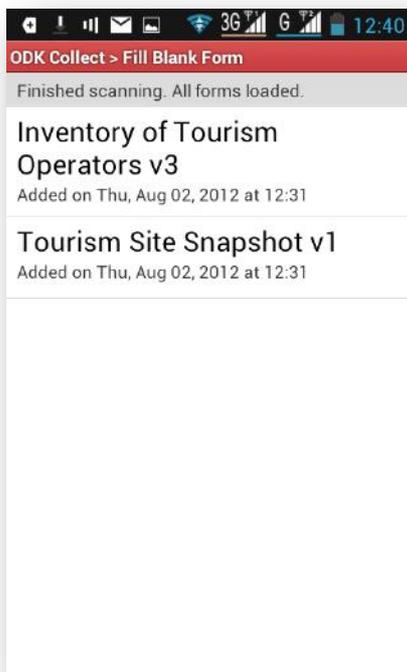
The Open Data Kit Application
at the Google Play Store



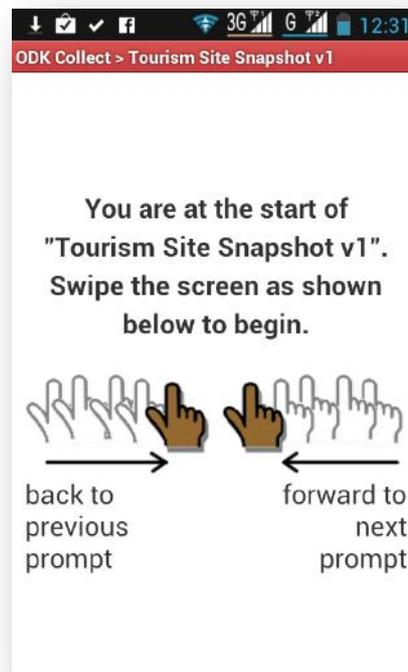
ODK Start Menu



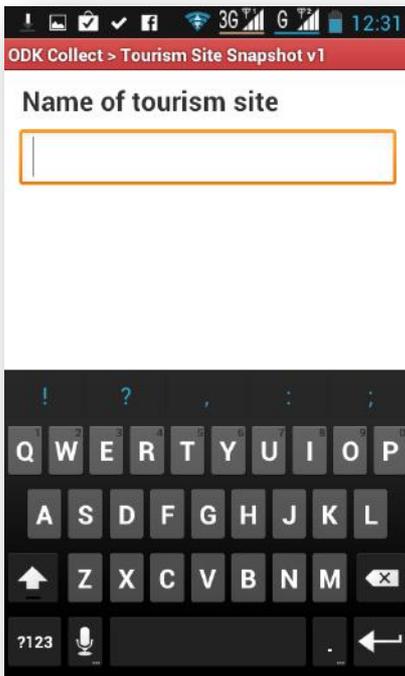
Form selection menu



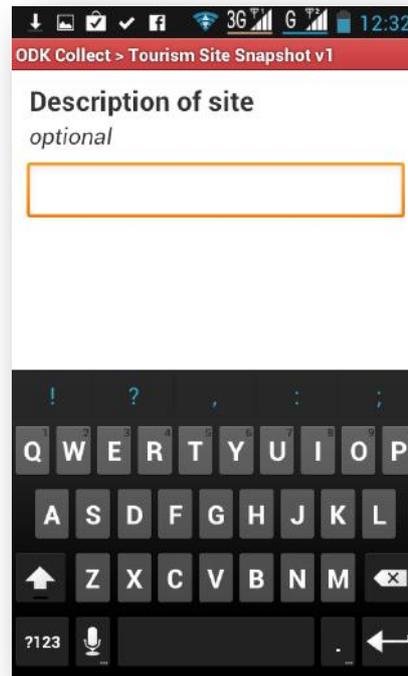
Survey Start Page



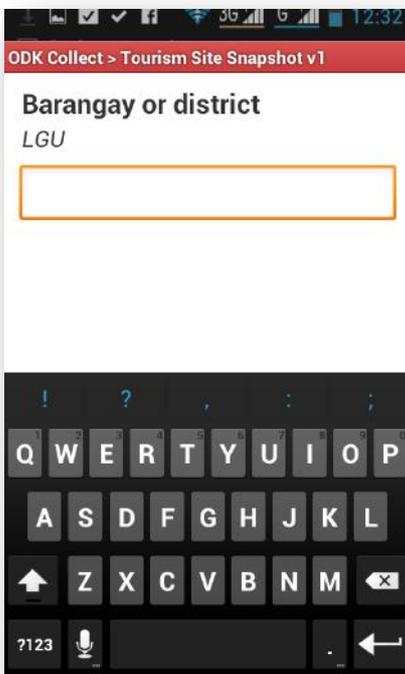
Text Input Question



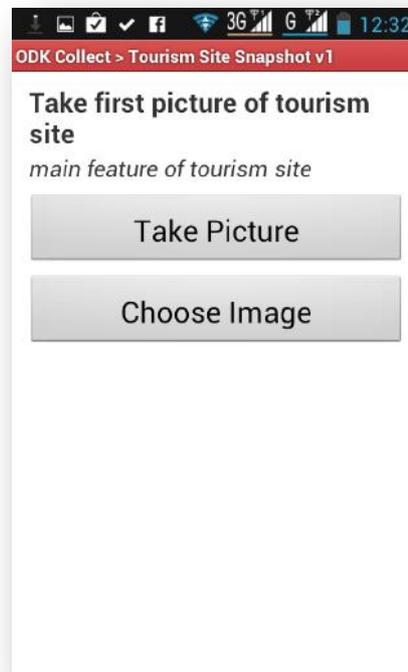
Text Input Question (Optional)



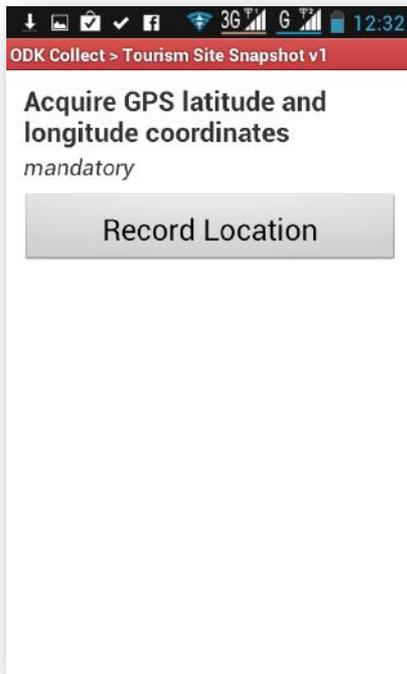
Text Input Question with Hint



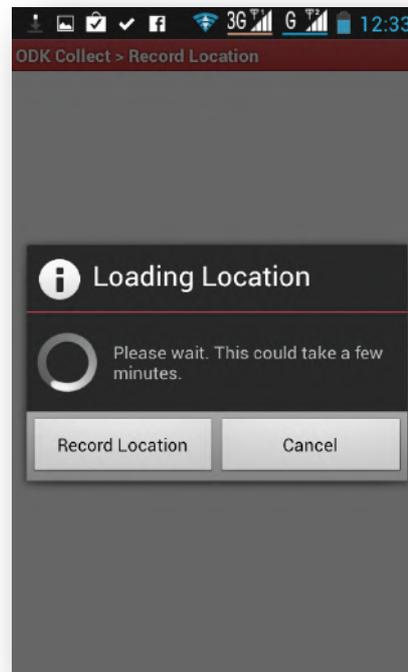
Multimedia (Photo, Video, Audio Clip) Question



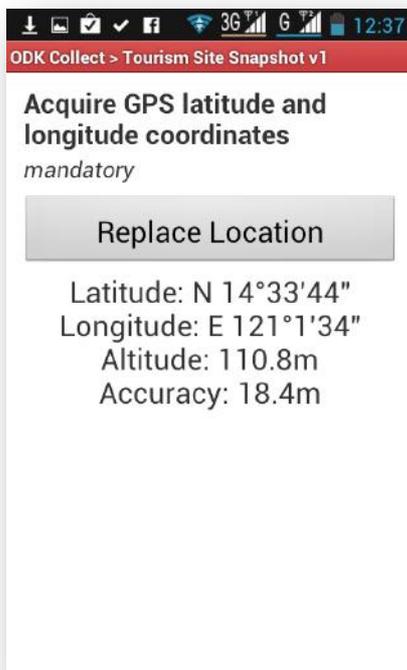
GPS Coordinates (Location) Question



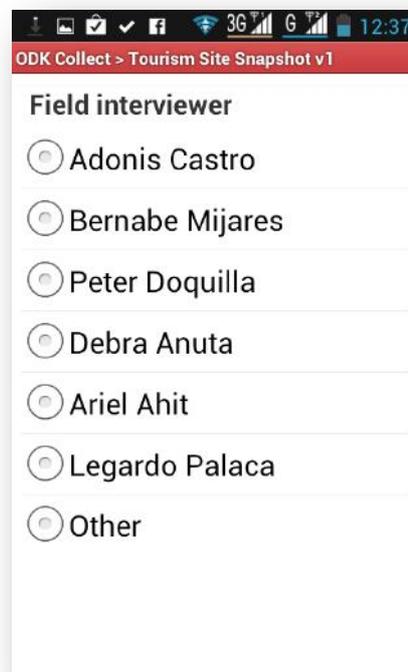
GPS Coordinates – acquiring location notification



GPS Coordinates – recording with accuracy in meters



Select One Question



Select One Question with hint and scrollbar

ODK Collect > Inventory of Tourism Operators v3

Classify type of tourism operator
select one

- accommodation establishments (eg hotels, resorts, apartments, pension houses, home stays, etc)
- restaurants, food and beverages services (eg restaurants, cafes, bar, etc)
- tour operators
- transport services (eg taxi, van, bus, tricycle, jeepneys, car rental,

Multiple Choice Question with hint

ODK Collect > Inventory of Tourism Operators v3

Common facilities in the accommodation establishment
select all available

- restaurant
- bar or cafe
- function room
- souvenir shop
- swimming pool
- gym
- internet connection eg WIFI at lobby

Numeric Input Question

ODK Collect > Inventory of Tourism Operators v3

Number of cottages, villas or bungalows

1	2	3	-
4	5	6	,
7	8	9	✕
←	0	.	→

Survey end page with save options

ODK Collect > Tourism Site Snapshot v1

You are at the end of "Tourism Site Snapshot v1".

Name this form:

 Mark form as finalized

Save Form and Exit

Annex G Links to Tutorials of Google Geo Tools

Google Earth Outreach

<http://www.google.com/earth/outreach/index.html>

Annotate Google Earth - Add placemarks (or points), paths (or lines), and polygons to Google Earth. Then embed images and YouTube videos to your balloons.

Training videos <http://www.youtube.com/user/EarthOutreach>

Using Overlays with Google Map Maker -

<http://support.google.com/mapmaker/bin/answer.py?hl=en&answer=155420>

Open Data Kit usage and training videos

<http://opendatakit.org/use/>

Google Spreadsheet Mapper v3 Tutorial

<http://www.google.com/earth/outreach/tutorials/spreadsheet3.html>

Annex H: ODK Deployments

This is a selection of *ODK* deployments worldwide:



International Space Station

- Astronaut [Ronald Garan](#) uses *ODK* to monitor progress on the [Carbon for Water](#) program. See [this post](#) for more.

Afghanistan

- *ODK* was used in Afghanistan for fraud monitoring of the 2010 elections.
- [Afghanistan Infrastructure Data Center](#) tracks and reports on civilian and military donor infrastructure projects using *ODK*.

Australia

- [CloudTech](#) uses *ODK* in their [Smapp-Suite](#) data collection system for [RMIT](#) and [World Vision Australia](#).

Brazil

- [ULevel](#) provides customizations on Collect for Android.
- The Surui Tribe in partnership with the [Amazon Conservation Team](#) and [Google.org](#) is piloting *ODK* for deforestation monitoring in the Amazon.

- University of São Paulo's School of Engineering and School of Medicine use *ODK* as one of its key tools for epidemiological studies.

Canada

- PeaceWorks, a software consultancy, is supporting Mennonite Economic Development Associates (MEDA) with the usage of *ODK*.
- Radical Dynamic has built Group Complete from *ODK* and CouchDB. Their system allows for real-time and collaborative data collection.

Central African Republic

- Harvard Humanitarian Institute has completed over 1800 hours of surveys documenting human rights violations using Kobo -- a tool built from *ODK*.

Chile

- Joaquin Blaya is using *ODK* as part of a strategy to commercialize open source health IT.

Columbia

- Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) uses *ODK* for data capture.

Denmark

- ViewWorld in partnership with Danish NGOs and companies are developing better reporting tools using *ODK*.

Ghana

- Lumana uses *ODK* to map villages, survey clients, and calculate the poverty indices of the communities they serve.
- Johns Hopkins Global Water Program uses *ODK* for longitudinal water surveys.

Haiti

- Haiti Regeneration Initiative is piloting *ODK* to monitor agricultural practices, crop productivity, and farmer yields of crops.
- Person Finder Mobile provided an *ODK*-based tool to collect data on missing people from the 2010 earthquake.

India

- EarlySail customized *ODK Collect* for a few pilot projects as well as assisted in form development. They will soon be porting *ODK* to BlackBerry.
- ICTPH enrolled 3,000 households (12,500 individuals) for a comprehensive primary healthcare delivery model in Tamil, Nadu. See report for more.

- PATH is piloting *ODK* for use as a digital birthing assistant with clinical decision support, management of patient information, and as a tool for communication and education.
- *ODK* is being used in UC Berkeley's Health in Hands project in Mumbai and Karnataka.
- UNICEF and the children of Kolkata are mapping their community with *ODK*.

Kenya

- Mega Six Solutions has experience designing forms.
- Bankable Frontier Associates is using *ODK* to conduct surveys that support their clients in their efforts to expand financial services to the poor.
- Indepth Research provides implementation services for *ODK*.
- Manna Energy and Vestergaard Frandsen are using 4000 *ODK*-enabled phones to collect 40000 forms a day for the Carbon for Water program.
- USAID-AMPATH uses hundreds of phones with *ODK* for home-based HIV counseling and testing of millions of rural Kenyans and transmits the data to OpenMRS.
- Datadyne's Episurveyor mobile data collection client for Android is built from *ODK*.
- UN Refugee Agency and CartONG used *ODK* for a mosquito net survey in Dadaab, Kenya.
- Mindflow Associates provides customization services to connect *ODK* and OpenMRS.

Lesotho

- SARPAM uses *ODK* in Lesotho, South Africa, Zimbabwe, Mozambique, Zambia, Malawi, Tanzania and DRC to track drug availability.

Liberia

- Harvard Humanitarian Institute documents human rights violations using Kobo -- a tool built from *ODK*.

Malawi

- Invest in Knowledge is using *ODK* for home-based HIV counseling, testing and survey of 2,000 couples. Invest in Knowledge also provides implementation services for *ODK*.

Mali

- AED-SATELLIFE collects information about school decentralization using *ODK*.
- Columbia University researchers successfully used and positively reviewed *ODK* in an agroforestry adoption study.

Mozambique

- Dimagi and Afris used *ODK* for a 400 question household survey in Mozambique that over three months sent 15000 submissions to their DataHQ server.
- openLMIS is a supply chain and cold chain management system from VillageReach that uses *ODK* form filling.

Nicaragua

- The Sustainable Sciences Institute, in collaboration with the Ministry of Health, uses *ODK* for tracking patients and their biological samples for research on dengue and influenza.

Nigeria

- The Millennium Villages Project uses *ODK* in Nigeria for surveying and designed XLS2XForms, an *ODK* compliant tool.

Philippines

- Philippines' Department of Trade and Industry has mapped over 700 tourist sites on the island of Bohol.
- Wireless Access for Health collects data with *ODK*.

Slovakia

- Acqee is web-based platform for collecting structured data. The platform uses *ODK Collect* as an Android client.

South Africa

- Seeing Swans have performed 12,000 surveys of informal structures in the Western Cape. Their customized *ODK Aggregate* provides real-time reporting and statistics.
- Wharton School of Business is using *ODK* in a three year survey to study health and entrepreneurship in townships around Tshwane.
- There is ongoing work to port CyberTracker's data collection tools to the *ODK* platform.

Sweden

- LETS GO is using *ODK* in fostering high school student learning in teams in environmental science.

Tanzania

- AfSIS is mapping soil conditions with *ODK*.
- Jane Goodall Institute and Google.org are piloting *ODK* for forest health monitoring.
- D-Tree is piloting *ODK* and the IMCI protocol for triaging children under five and for collecting maternal health data.

- REPSSI and Norvatis Foundation piloted *ODK* in Tanzania. REPSSI helps address the psychosocial needs of the over six million children affected by the triple crisis in southern Africa.

Uganda

- Makerere University's machine learning group uses *ODK* for automated diagnosis and mapping cassava mosaic disease.
- Grameen Foundation AppLab integrates *ODK* and Salesforce.com to provide real-time monitoring and evaluation of community knowledge workers.
- Johns Hopkins Center for Clinical Global Health Education uses *ODK* to improve provider communication and education through their eMOCHA project.
- Project Wet and MIT evaluated IVR data collection interfaces for untrained rural teachers using *ODK*.

United States

- North Central Wisconsin Regional Planning Commission uses *ODK* to collect data in the field for assessment information after a natural hazard like storm, flood or tornado.
- New York City (with InSTEDD) uses *ODK* to gather emergency preparedness and response-related data in the field.
- Foundatron in Seattle uses *ODK* to collect field data.
- WebFirst has built an mHealth solution using *ODK* and Drupal 7.
- Dimagi has combined *ODK* and CommCare support community health workers from the PACT program in Boston.
- The Transit, Technology, and Public Participation project in North Dakota has chosen *ODK* to manage surveys of transit workers.
- Rhiza Insight provides geospatial analysis and visualization tools as well has ruggedized mobile devices for *ODK*.
- The NatureMapping program at the University of Washington uses *ODK* for building multimedia-rich species maps.
- SanaMobile from MIT is integrating their mobile client with *ODK*.
- Small Meadows Farm is using *ODK* to collect pH, humidity, soil moisture, and plant observations in greenhouses and gardens.
- Trinity College expanded the XForms controller in Sahana Eden for *ODK* support.
- doForms is an *ODK*-based data collection solution for businesses.

Zanzibar

- D-Tree is using *ODK* in a UNICEF-funded project to better to identify and treat severely malnourished children.
- Vetaid is using *ODK* for collecting animal health information from 50 farms and 400 animals.

Zambia

- SANGONet use *ODK* for monitoring and evaluation in their agricultural work.

Zimbabwe

- ICT at RTI uses *ODK*, *JavaRosa* and *FrontlineSMS* for data collection.



GTZ Office Manila
German Development Center
10F PDCP Bank Centre
V A Rufino St cor L P Leviste St
Salcedo Village, Makati City 1227, PHILIPPINES

Private Sector Promotion (SMEDSEP) Program
PSP Program Office, 7F New Solid Building
357 Sen Gil Puyat Avenue
Makati City 1226 PHILIPPINES
☎ +63 2 897 8199, 556 8732, 896 4319
✉ info@smedsep.ph
🌐 smedsep.ph

Private Sector Promotion (SMEDSEP) Program Cebu
GF LDM Building
Legaspi St cor M J Cuenco Ave
Cebu City 6000, PHILIPPINES
☎ +63 32 412 2256
☎ +63 32 254 4958
✉ cebu@smedsep.ph