MOBILE SOLUTIONS

Introduction to PROJECT DESIGN

USAID
FROM THE AMERICAN PEOPLE

fhi360
THE SCIENCE OF IMPROVING LIVES

OVERVIEW

Introduction to Project Design is the second of four introductory guides on the topic of mobile data solutions. This guide is designed to provide USAID staff and implementing partners with useful information for preparing concept notes, scopes of work, Requests for Proposals (RFP) and Requests for Applications (RFA).

It provides the basic elements needed for assessing whether a mobile data solution is appropriate and discusses how to select appropriate tools for the specific task.[1] The material in this guide is not intended to provide a detailed project design process or a framework to guide implementation.

This overview is divided into five parts:

- Needs Assessment
- Readiness Assessment
- Additional Requirements Assessment
- Deployment and Support Assessment
- Examples

NEEDS ASSESSMENT

Defining what is needed is one of the most important steps toward selecting the right data solution. *Before beginning this process, the goals and objectives for the project should be clearly defined.* These should inform the assessment of the parameters listed below. The purpose of keeping these in mind throughout the project design phase is to ensure that the project is not using technology for technology's sake but is using technology appropriately to achieve desired outcomes.

This step starts with defining the data collection, management, analysis, visualization and data sharing needs of the project and stakeholders. This is essential for determining whether those needs might best be met by using a mobile data solution and if so, which one.

Parameters to Consider

Data: What type of data will be collected and used? Is the data primarily quantitative or qualitative in nature? Do you need to capture photos, video, audio, and/or geographic coordinates? What is the size of the survey in terms of the number of questions or items? How complex is it? Are some questions meant for only part of the prospective survey sample? Should some questions be asked only when specific information is obtained? Will you need to restrict the type of data entered for some questions? Will you need to validate the data as it is being entered? Is there a sensor or instrument that can or will capture data directly? What is the expected size of data that will be transmitted?

Process: Will data collection be a one-time event? Will the same survey or instrument be used several times? Will use of the survey or instrument be an ongoing, recurring data collection activity? Will data collection involve continuous measurement of data over a long time (for example, monitoring environmental conditions of water bodies)?

1 The parameters provided in the this document are based on field experiences of the FHI 360 TechLab and IDEA/Mobile Solutions staff and from section 4.3 of the World Wide Web Foundation's 2012 publication "**Multi-Channel Data Collection for Social and Economic Development**".

Duration: Will data be collected over a short period of time (such as a few hours for opinion polls), over several days (such as for household surveys), or over a period of months or years (such as for public health data)? Will data be collected at short intervals for an extended period of time?

Analysis: What are the data analysis needs? Who is responsible for data analysis? Are real-time dashboards that enable rapid implementation of response required?

Data sharing: What datasets will be shared? Who will have access to the datasets? When will the data be shared? Where will the datasets be located/stored (e.g., does a country prohibit storing or taking a particular type of data outside its borders)?

Data preservation: Are there regulatory requirements for preservation of datasets? If yes, for how long? Are there requirements for removing certain kinds of data (e.g., location data, personally identifiable data)?

Suitability Assessment for Mobile Data Collection

Mobile technologies should provide advantages over other means of collecting data, like paper and pen, and be an appropriate tool for the data type, process, and other factors considered if they are to be included in a project or activity. For example, mobile data collection tools may be appropriate for:

- Surveys that have complex skip patterns and have required questions, range checks, response validation (e.g., dates, email addresses), and calculated fields => *Electronic forms*
- Surveys which will serve as a baseline, where additional surveys will collect the same information as part of a longitudinal assessment with minimal changes => *Electronic forms, SMS, voice, IVR*
- Ongoing monitoring, particularly where access to real-time or near real-time data is important => *Electronic forms, sensors, SMS, IVR*
- Accurately capturing electronic data directly e.g., GPS coordinates, pictures, short audio clips, and measurements from sensors connected to mobile devices. => *Electronic forms, sensors*
    - The device chosen must provide this data at the resolution and level of quality you need.
- Collecting data from a large, distributed population => *SMS, voice, IVR*
    - Risk to the privacy and security of your population must be evaluated, as discussed in the Introduction to Mobile Data Solutions guide.

Mobile data collection may not be appropriate or cost-effective in other cases, such as for:

- Long, one-off surveys. The costs to design the survey and train enumerators and for the devices, unless they will be reused, may be prohibitive.
- Collecting large amounts of qualitative data e.g., surveys or focus groups where you ask a lot of open-ended questions. Lengthy, open-ended responses are difficult to capture on mobile devices unless digital voice recording is used. Even then, sufficient resources need to be allocated for the transcription and analysis of this data.

In addition, for surveys, the length and complexity of question types will determine some of the key specifications of the device chosen, such as processing speed, memory, and screen size. The amount of data for transmission will dictate the selection of the transmission option. Large datasets may require Internet access while small datasets can be transmitted via SMS.

With respect to analysis, sharing, and preservation, additional questions may need to be asked to determine the appropriate solution. However, mobile data solutions provide an advantage over paper and pen when a real-time dashboard is desired because they can provide continuous reporting of new information from the field. They require additional resources to configure or customize the dashboard, which has cost and time implications. In other cases, analysis may be done by experts. The important consideration will be whether or not the mobile data solution exports the data in a format they can easily import to their analysis software.

Data sharing may involve disclosure of data to an external organization or between different units of an organization which may have different permissions to view the data. Some countries have strict policies about what data may leave their borders or be stored on the cloud. These requirements need to be ascertained up front, including approval procedures, authorization for access, and restrictions on particular types of data. Finally, depending on the type of data collected, each country may have data preservation standards to ensure relevant data is captured and remains intact or is deleted immediately after it has served its purpose (i.e., location data, personal names, etc.). The duration for which data must be preserved or deleted will influence the choice of hardware and software, and affect the budget.

READINESS ASSESSMENT

Mobile devices are powerful tools and have proven to be efficient for data collection and information dissemination. However, they are not the right solution for every need, nor is every organization prepared to adopt the technology. Conducting an assessment of a project's or field site's readiness for the successful introduction of mobile data solutions will help you in deciding the viability of mobile data solutions and selecting the appropriate system for your setting. The following parameters will help assess your readiness for a successful implementation.

Infrastructure: Is there mobile network connectivity at the data collection sites? What type of connectivity (such as voice and SMS only, GPRS, 3G, etc.)? How widespread is it (for example is it available only in towns along a main highway or throughout a region)? How reliable is the service? If there is no mobile network connectivity or it is insufficient for the needs of the project, is there another way to upload the data (e.g., WiFi, cable, etc.)? Is there a power supply for recharging mobile devices?

Fast and reliable data transmission depends on the type and quality of connectivity. SMS data transmission supports very limited data size. Transmission of large datasets (such as household surveys) via a mobile network requires packet-data transmission such as GPRS or 3G. Spatial coverage and reliability of the data transmission option is important for the successful deployment of mobile data solutions. However, there are options to back up data collected in the field and upload the data when a connection is available which you may want to evaluate.

Data acquisition: Will data be collected by enumerators? Will the enumerators collect data from the field, or will they provide remote assistance to facilitate the data collection process? Will observations and/or measurements be made using sensors or other instrumentation?

Most surveys require enumerators to conduct interviews and collect data on mobile devices. Fortunately, anecdotal evidence as well as peer-reviewed assessments have shown enumerators in developing countries readily accept mobile devices for data collection. They also have demonstrated the ability of local people to use these technologies successfully even when they have had little to no exposure to the technology previously. However, enumerators will require additional training and support beyond that provided for paper surveys. For sensors, the device, power supply, data retrieval, and their fit within the environment are some of the factors that need to be considered.

Respondents: Who are the respondents? What capacities do they have in relation to mobile devices? Are the respondents literate? Do all prospective respondents speak the same language? In the absence of enumerators, do the respondents have the skill and access to mobile devices required for submitting data themselves?

These are particularly important questions that need to be assessed for a disaggregated population. For example, women have been shown to have less access to mobile phones and in some cases less facility or interest in using SMS. Functional literacy is also a concern. These issues may lead to a solution that uses IVR, voice, or another method. In addition, there have been few studies about how respondents perceive data collected on mobile devices by enumerators or data they enter via SMS. How respondents react should be considered when piloting any intervention or assessment that uses a mobile device.

Quality control: Who is responsible for data quality checks? Are there project staff in the field or in a central location who will verify data quality? Is automated data quality control that monitors anomalous data patterns available and will it be adequate?

Mobile data collection forms can have built-in algorithms for enhancing data quality, such as checking for data outside valid ranges, duplicates, and invalid formats. SMS surveys do not have this feature at the time of collection, though some platforms have validation options on their server applications that can help with quality control. However, all surveys also require review of datasets by the researcher to ensure high data quality. Some mobile data solutions provide features that will facilitate this review process.

ADDITIONAL REQUIREMENTS ASSESSMENT

It is very important that you select the equipment, software application, and transmission option(s) that meet your needs and are optimally suited to the intervention environment and existing infrastructure. This process is primarily driven by answering the questions posed in the "Needs Assessment" and "Readiness Assessment" sections of the project design. In addition to the parameters provided in those stages, the following parameters will help you further define your requirements and guide you in selecting the appropriate set of tools.

Language: What languages does the system need to support for forms design, data entry, system administration, data analysis and visualization? Is the data collection system expected to be used by people with low literacy?

While many of the mobile data collection clients support multiple languages for the survey questions, they often have more limited options for the user interface of the client and for all server-side tools. Make sure the languages you need are supported before selecting a tool. If supported, but not yet available, you will need to work with the vendor for them to be included. Also, if a survey will be conducted in multiple languages, it would be useful to select an application that facilitates localization of the survey tool.

Initiative: Is entry of responses to the survey initiated by respondents (such as in cases of election monitoring or reporting a malfunctioning hand pump) or by the enumerators (such as in cases of structured field surveys)?

This may have an impact on the selection of hardware, software and transmission options. For example, if response to the survey is initiated by respondents, the form may have to be device-agnostic and use the most ubiquitous data transmission option, which in many cases would be SMS-based transmission. If data is being transmitted automatically from sensors or other instruments, provisions (such as human resources and travel) must be made for periodic maintenance. Enumerator initiated surveys have concerns similar to those for traditional methods.

Hardware: What are the optimal requirements of the hardware (such as mobile devices, sensors, etc.) including battery life, screen size, ability to withstand harsh environments, operating system, and solar or other power solutions? What are the cost limitations for the project? Is a dedicated server required? Will enumerators or

respondents be provided with mobile devices suitable for the specific needs of the data collection activity? Or will they be expected to use their own phones? If sensors will be used what is the preferred size or form factor? In stationary sensor or other instrumentation installations, for how long is it expected to function reliably without changing batteries and other preventive maintenance?

There are many types of mobile devices and sensors with widely varying capabilities and costs available on the market today, and the technology is changing rapidly. In general, higher-performing devices that are more rugged will cost more. For example, if you need accurate GPS, you may not be able to use the cheapest smartphone on sale in Jakarta or Nairobi. This complexity can make choosing an appropriate device seem complicated. You can improve your chances of making a good choice by identifying your requirements and short-listing the devices that meet your criteria. Test, test, test the devices in actual field conditions. Just because a device performs well in the office does not mean it will perform well in the field.

Privacy: Are there any confidentiality or privacy issues that the system should comply with?

Personal information such as names, addresses, medical history, test and laboratory results, insurance information and other data that can be linked to an individual is usually considered confidential. It is important to identify, understand, and comply with local regulations including those pertaining to secure data transmission and sharing raw and analyzed data.

Interoperability: What are the requirements for the interoperability of the mobile data solution with existing solutions already in use by key stakeholders? Are there existing/legacy datasets that should be integrated with the new system?

It is important to assess previous investments made for the creation of custom applications or databases. To avoid the risk of isolating "legacy" databases and to ensure that a single harmonized database is being created, integrating legacy databases into the newer solution might be required. If so, the new system should support such integration. What additional expenses will be required by such system integration?

DEPLOYMENT AND SUPPORT ASSESSMENT

The introduction of mobile data solutions may require training of staff, including data managers, enumerators and sometimes respondents (such as when using radio broadcasts to inform the intended respondents). In addition, providing technical support for both hardware and software components of your mobile data solution is usually required to resolve user issues and challenges and maintain seamless operation. The following parameters will assist you in defining training and user support requirements.

Training: What are the minimum skill sets required by data managers and enumerators? How many individuals will need to be trained? Is it most efficient to train all staff at a central location or will it be more cost-effective for trainers to travel to the trainees? Is training of respondents required (such as using radio broadcasts)? What training curriculum would meet the requirements (in terms of time and other resources)? Is the data collection work force stable or will training be an ongoing requirement? Is there in-house capacity to deliver training?

The introduction of mobile data solutions will require training for your staff, including both technical personnel and end users. You must determine whether your organization is capable of providing this training internally, whether you will need to bring in outside trainers, or some combination. Depending on the complexity of the applications you use, you may need to develop and produce customized training materials for your staff.

Additionally, with an enumerated survey, depending on the complexity of the survey and familiarity of the enumerators with the mobile device, training of data collectors may take from a few hours to two days. In general, expect the training to take slightly longer than for a paper survey.

Technical Support: What are the hardware and software technical support requirements? Who will be responsible for providing ongoing technical support and troubleshooting technical issues? If issues cannot be resolved in-house, who will provide advanced technical support? Who will provide support for the maintenance and troubleshooting of devices? Are there device support centers in areas where data collection will be conducted?

Technical support is an ongoing activity and is typically required throughout the life of a program. It is important that a decision is made early whether user support will be provided through in-house capacity or will be outsourced. Local hardware support provides users with information regarding the mobile device and assistance if it malfunctions. Device support is usually provided through local vendors, authorized agents of the manufacturer, and/or mobile network operators.

EXAMPLES

In this section, two examples of data collection processes are provided to illustrate how the parameters discussed in this guide may be applied when assessing whether or not to use a mobile data solution and if so, which choices to make. These examples are hypothetical, but they are based on actual experiences deploying mobile data solutions.

Example 1: Household Survey in Uganda

This example draws from a project in Uganda that sought to generate and analyze rigorous empirical data on the effectiveness of integrated agriculture, health, and nutrition programming on maternal and child health outcomes. The project included a household survey that had about 2,700 questions. The project trained enumerators to conduct the survey using tablets. Data was transmitted to a server over a 2G/3G network. Assessment of the parameters introduced in this overview is provided below.

Assessing Needs:

Data: A very large quantitative survey with about 2,700 questions. Capturing geographic coordinates will be required. Some of the questions will be structured in large tables.

Process: 640 households will be randomly selected from seven districts and asked questions related to agricultural production, health services, and nutritional awareness and maternal and child health related issues.

Duration: About one month for the baseline data collection. Mid-line and end-line surveys will also be conducted using the same survey instrument.

Analysis: Data will be analyzed using a separate system. (STATA was chosen by the researchers.)

Data sharing: Raw data will be owned by the project. Data sharing requirements are outlined in the IRB approval document. Only de-identified data (i.e., data that is stripped of information which would allow the identification of the responding household) can be shared with others who are not direct stakeholders of the project.

Data preservation: No requirements.

Assessing Readiness:

Infrastructure: The survey will be conducted in seven districts where there is very good cellular coverage for voice and text messaging. There is fairly good GPRS coverage in the districts; 3G coverage exists in the district capital towns. Access to electricity is limited in the rural areas. The district capitals have access to electricity for about
3 – 5 hours per day.

Enumerators: 28 enumerators will be engaged to gather information from 640 households.

Respondents: Respondents will be interviewed by the enumerators and will not directly enter information to the mobile devices.

Quality control: Data quality control will be performed at different levels. The data collection tool on the mobile device will have in-built algorithms such as automatically checking "valid range" and alerting the user to re-enter data if the values fall outside of the range. Data quality control checks will also be performed by the researchers on a daily basis by accessing transmitted datasets on the server.

Additional Requirements Assessment:

Language: In addition to English, the survey will be administered in six local languages. All languages use Latin characters.

Initiative: Enumerators will interview respondents. Enumerators must obtain signed informed consent from the respondents prior to conducting the survey.

Hardware: The project will provide mobile devices for the survey. The screen size of the devices should be large enough to display a three-column table. The device should support GPRS/3G data transmission. Long battery
life is desirable.

Privacy: Only de-identified data can be shared with the public.

Interoperability: Export to any relational database from the server should be supported.

Deployment and Support Assessment:

Training: Enumerators will receive training on using the data collection and transmission system, and the survey instrument.

Technical Support: Technical support will be provided to the enumerators by the local project staff who in turn will contact technical staff at headquarters for higher level assistance.

Example 2: Social Media Survey of Youth in Jamaica

This example describes a mobile data project in Jamaica. The objective of the project is to generate and analyze data on the use of social media by youth and to determine which social media websites have the greatest potential as sources of health information for this population. The project proposed to conduct the survey via SMS with youth who are part of a local civil society organization (CSO). The project had contact information for most, but not all of their members, including mobile phone numbers. The survey consisted of ten questions. The solution was an SMS survey to a sample of youth from the CSO. Assessment of the parameters introduced above is provided below.

Assessing Needs:

Data: A small amount of data will be collected during each interaction. No photographs, video or audio data, or GPS coordinates will be required.

Process: 200 youth will be randomly selected and surveyed.

Duration: Interviews will be conducted over the course of one week.

Analysis: Data will be analyzed using a separate system. (SPSS was chosen by the researchers.)

Data sharing: Raw data is owned by the project. Data sharing requirements are outlined in the contract with the implementing partner.

Data preservation: No requirements.

Assessing Readiness:

Infrastructure: The survey will be conducted entirely within Kingston. There are strong mobile network services throughout the city.

Enumerators: None required.

Respondents: Respondents will answer survey questions using SMS. No training will be provided beyond instructions included in the SMS messages.

Quality control: SMS messages received by the software that do not conform to the coding system and data entry format will generate a semi-automatic request for resubmission; additional details about how to format the response will be included in the message. If this does not elicit a correct response, the response will be discarded.

Additional Requirements Assessment:

Language: The survey will be conducted in English. No other languages are required.

Initiative: Respondents will receive a request to participate in the survey via SMS. If they choose to participate, they will receive the questions as individual SMS messages on their personal mobile phones and they will respond to them via SMS. They will receive airtime credit for the number of SMS messages they send.

Hardware: Respondents will use their personal mobile phones.

Privacy: No information will be collected from the respondents which could be used to identify them, and identifying information, like the respondent's mobile phone number, will be disassociated from their survey responses.

Interoperability: This is a standalone survey and the data will not be linked to any prior research.

Deployment and Support Assessment:

Training: Local project staff will receive instruction remotely from technical staff. Technical staff will train the local project staff on the survey and the data management software where they will perform quality control.

Technical Support: Technical support will be provided to the local project staff by the technical staff at headquarters.