Mobile Technology Support for Field Research

Wayne Johnston

Preamble

When I arrived in La Paz my prearranged driver failed to show up. After looking everywhere and waiting for longer than seemed necessary, I took a cab from the airport to my hotel. The drive into the Calacoto neighborhood gave me my first sense of the dizzying landscape of La Paz, but even more than that, the street dogs we passed en route struck me. Roaming the streets singly or in packs, picking through garbage for anything edible they could find, every conceivable breed of dog seemed to be homeless in La Paz.

I was in La Paz to meet with researchers at the Universidad Privada Boliviana, who were conducting a series of surveys on the health of children born prematurely throughout the poorer regions of Bolivia. They told me that their previous survey had resulted in five tons of paper, which I could see was taking up almost all of the office space that they had available. The data had to be transcribed from the paper surveys into a database, which required a large investment of time and money. Additionally, errors were introduced during the transcription process. Data validations were then required using the paper surveys, resulting in further delays. The researchers were seeking my advice as they transitioned the data collection to mobile technology in order to achieve greater efficiencies as well as more reliable data.

At night I lay in my bed unable to sleep, listening to dogs throughout the neighborhood barking, howling, fighting. It wasn't the noise that kept me awake so much as being drawn into contemplating how a homeless dog problem of this scale could ever be addressed. I learned that there are about 300,000 street dogs in La Paz and the number grows by 20 percent each year. A dog sanctuary big enough to accommodate all the street dogs of La Paz would need to be as vast as the city itself.

A year later, back in Guelph, I met Luz Maria Kisiel, a researcher at the Ontario Veterinary College. Luz was undertaking a research project on the roaming dogs of Villa de Tezontepec, Hidalgo, Mexico. Mexico also has a major problem with street dogs; Mexico City alone reports that it captures and kills 20,000 dogs every month. I worked with Luz to develop a mobile app that enabled her team of local research assistants to locate street dogs, photograph them, record their GPS location, and take notes on their condition. This data supplemented a household survey focused on local residents' experiences with these dogs. Luz's research project will help to inform the first national program of sterilization of dogs in Mexico.

From sleepless nights in La Paz wrestling with a problem that seemed unsolvable to a productive partnership with a researcher back home, I had found compelling examples of the research cycle driven by mobile technology.

Introduction

In 2009 the author was appointed head of a new team at the University of Guelph Library that was established to support researchers from project conception and grant application through to publishing and long-term preservation of scholarly outputs. It was evident from the outset that there was a range of services and technologies to support researchers within the campus environment, but almost nothing to assist them once they left campus to undertake field research. Virtually every academic department has researchers that rely heavily on field research, whether they are conducting agricultural research in nearby crops or international studies on other continents.

Many researchers rely on paper data collection in the field, which requires that the data be later transcribed into a computer in spreadsheets or databases. This not only delays the availability of the data for analysis but also introduces errors during the transcription process. Paper records are also highly vulnerable to loss and damage. Collecting data using mobile technology, especially when the data is uploaded to a properly administered campus server, makes the data much more secure. The data is immediately available for review and analysis. Among other things, this enables researchers to do data validation while actually on site; often, it is too late to correct errors if the researcher identifies the problems only after she or he is back on campus. The use of mobile technology also introduces

additional functionality. This includes more reliable data through the use of authority controls and validity checks. It also enables use of GPS coordinates, photographs, audio, and video in addition to textual and numeric data.

Kevin McCann, Canada Research Chair in Biodiversity and a faculty member in the Department of Integrative Biology at the University of Guelph, provides one powerful illustration of the danger of relying on paper data collection. Like Kisiel, McCann was engaged in an extensive data collection project in Mexico. He and his colleague had to visit many sites, so they purchased an old truck to get them from place to place. All of their data was being collected on paper and stored in the truck. With no scanning of the paper records or transcribing of the data on site, there was no redundancy if anything happened to the original records. They were well into their project and had amassed a considerable amount of data when the truck was stolen. The researchers didn't care about recovering the truck, but they desperately needed to recover the paper, as it represented a huge investment of time and effort. They were very pleased when the police reported that they had found the truck. They were less pleased when the police informed them that they could have the truck back only if they agreed to repurchase it from them. Still, this was a small price to pay in order to get their data back.

It was in this context that the author undertook a multiyear research project to identify ways that an academic library can better support field research. The focus of the project is on data collection using mobile technology, but the research also encompasses data storage, data security, device selection, and other unique challenges encountered by researchers in the field.

Literature Review and Methodology

An initial literature review revealed that very little attention has been paid to how mobile technologies can be used in academic field research and, in particular, how academic libraries can play a support role. Other aspects of the question were well represented, such as how libraries can deliver services and content via mobile devices or how mobile devices have been exploited for information dissemination in the context of health or humanitarian initiatives. There is some literature available focused on data collection using mobile devices, but generally not from the perspective of academic research.

Gabriel Demombynes and colleagues (2013) present research that features surveys that were administered by a call center. Their report provides an example of using mobile phones for data collection in a developing country where, on the one hand, cell phones are ubiquitous but, on the other hand, network connectivity is unreliable. This dichotomy illustrates why it can be important to employ solutions that capture the data on the device and then automatically upload the submissions when connectivity is eventually established.

In 2015, Michael R. Glass demonstrated how engaging students in data collection using mobile technology can be a valuable component of their introduction to research methodologies. His use of a data collection app for a neighborhood survey provides a compelling example of experiential learning, giving the students a deeper appreciation of field research. Glass cites the particular advantage of having the results of data collection available for immediate analysis.

Sandro Mourão and Karla Okada (2010) describe the development of a mobile solution that features the three main elements that are needed in virtually all tools. These include a mechanism for designing the data collection instrument, a web server where the data will be stored, and a mobile app, which will be used by respondents or survey administrators.

A technology review was also undertaken for this project. This involved identifying software and hardware that could be useful to academic field research. The author read about the tools and resources that had been identified and also acquired them whenever possible in order to gain firsthand experience. When possible, the tools were also introduced to researchers for testing on actual research projects in the field.

This mobile technology research project was informed by a series of conversations with researchers at the University of Guelph. Researchers were initially identified for outreach simply by searching departmental websites. Each conversation included recommendations for other researchers actively involved in field research. Twenty-five researchers were consulted in disciplines as diverse as geography, biology, rural development, plant agriculture, environmental science, political science, anthropology, and population medicine. The conversations were loosely structured to encourage respondents to share stories that would reveal their unique challenges when conducting field research. The topics covered included a description of the environment where the researcher conducts her or his fieldwork, current data collection practices (and any concerns or frustrations associated with those practices), data storage strategies including incidents of data loss, any current use of mobile devices, and what is done with the data upon return to the campus environment.

Although some researchers confessed to being technophobes, the vast majority found the prospect of using mobile devices for data collection to be compelling. Security of the data was a prominent concern. For example, one researcher described collecting data from boats, where the paper invariably became wet and the observations became illegible. Other researchers expressed a need to incorporate GPS coordinates, barcode scanning, or photographs into their data collection work. Some researchers sought tools to assist with sample management, as samples collected in the field need to be associated with identifying metadata. Another common need was the ability to identify data inconsistencies while in the field and to address them immediately, as it is often too late to validate questionable data once the researcher is back home.

In addition to garnering information from the literature and directly from researchers, the author felt it was important to gain firsthand experience by working with researchers in a developing country in order to better appreciate the challenges of working in a remote environment. The author was fortunate to be awarded a research grant from the Canadian Association of Research Libraries that enabled him to travel to Bolivia to work with researchers in three different cities.

The author first worked with researchers in La Paz at the Universidad Privada Boliviana, Centro de Generación de Información Estadística, learning about their experience conducting a survey of about 5,000 households in the poorer regions of Bolivia, the result of which was five tons of paper. They were very eager to migrate their data collection to mobile devices. The second city visited in Bolivia was Cochabamba, where the author engaged with researchers at Fundación PROINPA. While there, the author developed a data collection app to enable an agricultural researcher to measure the growth and yield of her experimental crops. The app addressed concerns such as the collection of data in rainy weather that was problematic with paper, eliminating the need to transcribe data to spreadsheets, and the incorporation of photographs into the data. The third city was Santa Cruz, where the author visited the Centro de Promoción Agropecuaria Campesina (CEPAC). Researchers there were advised on how to migrate their data collection practices from paper to mobile devices.

These experiences in Bolivia tested the conclusions drawn from the earlier conversations with researchers at the University of Guelph and confirmed the great potential for mobile technology to positively impact field research wherever it is undertaken.

Open Data Kit (ODK)

Many of the applied solutions reported on in this chapter make use of the Open Data Kit. ODK is open-source software developed at the University of Washington with support from Google. It is a suite of tools that enables researchers to create and deploy mobile data collection solutions. ODK has become the core of a number of other products, both commercial and open-source. (Examples include Ona, SurveyCTO, KoBo Toolbox, CommCare HQ, doForms, DataWinners, ViewWorld, and PhiCollect.)

The ODK suite is made up of the following core components:

- ODK Aggregate, which is a server module for deploying data collection instruments, storing the data, and downloading and analyzing data.
- ODK Collect, which provides the data collection app itself on Android devices. ODK Collect stores submissions on the device itself until a connection to the Internet has been established. Enketo extends

deployment to non-Android devices by offering a browser-based alternative.

- ODK Build, which can be used to create the data collection instruments, although XLSForm is a more robust form design tool.
- ODK Briefcase, which is a desktop alternative to Aggregate.

Other ODK modules available include ODK Scan, ODK Survey, ODK Tables, ODK Sensors, and ODK Validate.

XLSForm is a way to create data collection instruments by populating a simple Excel table with the form definitions. While some researchers may not find this approach as user-friendly as a drag-and-drop interface initially, once the principle is understood, it becomes very quick and easy to create forms. XLSForm incorporates controlled vocabularies, value constraints, conditional elements based on previous responses, calculated values, grouping and repeating of elements, and incorporation of media and GPS. Once the form has been designed in XLSForm, it is converted to XForm for use by ODK as well as a range of other platforms.

As mentioned, Enketo is a separate product that deploys the same XLSForm definitions as browser-based alternatives to ODK Collect for non-Android devices. Virtually all the same functionality is retained in the browser version, including the capability to store submissions on the device when not connected to the Internet.

Formhub was developed as an alternative to ODK Aggregate by a group at Columbia University. It facilitated the conversion of Excel form definitions to ODK Collect apps as well as the sharing of forms with other research groups. One of its main advantages was that it offered free hosting and storage for any project that wanted to use it. Unfortunately, the very popularity of Formhub spelled its demise as its use internationally became too much for the dedicated bandwidth as well as the storage capacity. The developers themselves moved on to form a new company called Ona, which resulted in the Formhub code no longer being maintained.

An alternative to installing ODK Aggregate on a local server is to install it on Google's App Engine. There is a well-documented process that makes this easy to accomplish even for users with no expertise in setting up web servers with database back ends. The process for installing an instance on App Engine is essentially (1) make sure you have Java installed on your desktop; (2) set up Gmail and App Engine accounts; (3) download ODK Aggregate and run the wizard-based install; (4) configure the upload of ODK Aggregate to App Engine. You're then ready to begin deploying forms and collecting data with none of the expense and expertise required to install and maintain your own local instance.

Use of Google's App Engine is free for a three-month test period. The pricing after that is difficult to estimate as it involves a combination of transactions and storage. Most modest data collection projects should remain below the threshold where any charges are accrued. Apart from pricing, the other disadvantage of an

App Engine instance is that Google can access the data. For some projects this may not be a concern. However, projects involving sensitive or private data would need an option that is more secure.

Deployments

There are a number of practical examples of how various tools and strategies have been deployed in actual research projects. As mentioned in the chapter's preamble, Luz Maria Kisiel is a graduate student in population medicine at the Ontario Veterinary College, University of Guelph, and her research interest focuses on roaming dogs in Villa de Tezontepec, Hidalgo, Mexico. Her fieldwork involved recruiting about thirty local research assistants to identify street dogs over a set of five predetermined routes. The author worked with Luz to develop an ODK app that enabled her assistants to take photographs of the dogs, capture their GPS location, and record comments on the dogs' age, gender, and condition. This was supplemented by a household survey of people's experiences with street dogs.

Kisiel was very pleased with the efficacy of ODK Collect as a mechanism for data collection. The project documented and photographed 428 dogs, and she plans to return to the community in a year's time to see if the same dogs can be identified again based on their photographs.

Cameron McCordic is a postdoctoral fellow with the Hungry Cities Initiative at the Balsillie School of International Affairs, University of Waterloo. His research has taken him to Ghana, Kenya, South Africa, China, Mozambique, Jamaica, India, and Mexico for intensive data collection activities. He speaks compellingly of how the use of mobile technology and ODK in particular has transformed their work:

> The use of digital surveys in field research mean that the hardware and software components in tablets and phones can now be integrated into fieldwork more efficiently. These components include the camera, audio recording, and GPS hardware on tablets and phones as well as the added value of downloadable software like random number generators for random sampling, geo-spatial tracking software for enumerator navigation, and encryption software for data security in the field. (C. McCordic, personal communication, April 17, 2016)

Use of this technology has reduced some of McCordic's field research budgets by up to 75 percent thanks to the move from paper to mobile. He also speaks of the efficiencies gained by having data immediately available for analysis and for callbacks to respondents.

One project that presented a particularly interesting set of challenges was a mindfulness research project by Ekaterina Pogrebtsova, a University of Guelph graduate student in psychology working with Professor M. Gloria González-Morales. This project involved extensive use of multimedia, which needed to be available offline and needed to work with all devices. The concept is that respondents will answer a brief survey on their emotional state, go through a mindfulness exercise via audio or video, and then complete another brief survey to determine how their emotional state was affected by the exercise. The researchers want all of this to be available on the respondents' mobile devices so that they can participate wherever they happen to be at the time.

A number of researchers from the University of Guelph have been participating in an international project examining declining fish stocks in Cambodia. One aspect of the project investigated use of mobile technology for data collection. Some challenges were faced with qualitative surveys where the interviews were conducted in Khmer. However, the responses were translated into English during data entry. The researchers were also pleased with the GPS capabilities introduced through the use of mobile devices.

Discovering Biodiversity is a third-year class at the University of Guelph that exposes students to field research. The students are assigned to a plot of land in a forest where they need to identify the trees within the assigned plot. The data had been collected on paper and, given that there are about two thousand students enrolled in the class each semester, this meant a lot of data to be transcribed into spreadsheets before it could be used in subsequent classes. By moving data collection to mobile technology, the data has been made immediately available for analysis.

Knowledge Base

The primary product arising from all of the aforementioned research is a wiki knowledge base (https://fieldresearch.miraheze.org/wiki/Main_Page). It is hoped that while the author invests time at the outset in populating the site, a broad community of interest will be motivated to further develop the content over time.

The knowledge base is made up of the following sections and subsections.

1. Data Collection

The main focus of the project is on data collection in the field. This section is made up of seven subsections described below.

1A. APP CREATION

The App Creation subsection covers tools that facilitate the development and deployment of apps without the need for programming expertise. Included among these tools are those that follow the OpenRosa standard, including ODK.

1B. APP DEVELOPMENT

App Development is distinct from App Creation in that it involves designing apps from scratch using frameworks that facilitate the process. While these tools do require knowledge of HTML, CSS, and JavaScript, they generally do not require coding from the ground up. That makes these resources worthy of consideration for research projects with highly specialized needs.

1C. SURVEY SOFTWARE

Survey Software covers some of the leading providers of web survey solutions, both open-source and commercial. These are useful for research projects that want to move away from paper and can rely on Internet access without the need for any functionality specific to mobile devices, such as GPS or photographs as data elements.

1D. SMS

A selection of SMS tools offer an alternative approach to data collection based on text messaging. Simple, low-cost data collection is advantageous especially when input from a large and diverse population is needed.

1E. MOBILE DBMS

A subsection on Mobile DBMS covers database applications stored on the mobile device.

1F. SPECIALIST DEVICES

Specialist Devices provides a sample of the wide range of devices that have been custom-built for very specific purposes. These include devices that monitor environments such as air, water, and soil.

1G. SPECIALIST SOFTWARE

Similarly, the Specialist Software subsection is meant to give researchers a sense of the range of software options available to address very particular data collection needs.

2. Mobile Mapping

While capturing GPS coordinates is a feature of many solutions, Mobile Mapping covers tools that facilitate the capture of more complex mapping data.

3. Data Storage

Data Storage is a section that deals with safe and secure storage of data when away from the campus environment. This includes both cloud storage solutions as well as tools to enable data to be sent back to the campus servers.

4. Data Transfer

Data Transfer provides information and best practice for the transfer of data between the source and the chosen storage system.

5. Data Security

The Data Security section covers additional recommendations on ensuring the security of data, especially when private or sensitive human data is involved.

6. Device Protection

Researchers in the field also need to be concerned about protecting the devices themselves. This includes protection from damage when working in hostile environments as well as protection from theft and hacking.

7. Data Access

The final section deals with use of mobile devices to access data from the home environment. There are times when researchers in the field will need to consult data collected in the lab or on previous field trips.

Conclusions

Unlike many technology services, support for field research cannot be provided with any single enterprise solution. This is partly because the needs of research projects across the disciplines are so diverse that the specific requirements of each project need to be matched with a tool that will respond to those needs. This process can include issues of functionality specific to mobile devices, issues of scope, and issues of data sensitivity. It can vary depending on what specific devices are

available for use and on issues of complexity and technical sophistication. Another issue to consider with respect to commercial solutions is that many larger data collection initiatives depend on local recruits, which would violate the terms of most commercial license options.

Instead of looking for a one-size-fits-all solution, librarians providing support to field research need to be aware of the range of options so as to make appropriate recommendations. This, of course, is not a one-time acquisition of knowledge. It is critical to keep up-to-date on developments that impact the options available to researchers.

The investment of time required to gain and maintain this knowledge is worthwhile, as libraries that assist researchers with their data collection needs will have early buy-in for broader research data management objectives right through to the publication and preservation of research data. For the most part, no particular technical expertise is required to implement effective mobile data collection solutions. The main challenge is the investment of time to become familiar with the tools and resources, which in turn is diminished when libraries work collaboratively to maintain a shared knowledge base.

While collaboration between libraries is critical, so too is collaboration with campus partners in our own institutions. Campus IT, the office of research including research ethics, and departmental research managers all contribute pieces of the puzzle. Sometimes the library's most valued role is the familiar one of liaison, not just between the faculty member and library services but also with other campus partners. Researchers are often unaware of the support services available, or they are daunted by the challenge of knowing who to contact. Whether the questions relate to data storage, backups, encryption, or ethical vulnerability, the library can ensure the researcher has access to all of the necessary support he or she requires.

References

Demombynes, G., Gubbins, P., & Romeo, A. (2013). *Challenges and Opportunities of Mobile Phone-Based Data Collection : Evidence from South Sudan* (Policy Research Working Paper No. WPS6321). The World Bank, Africa Region, Poverty Reduction and Economic Management Unit. Retrieved from http://www-wds.worldbank.org/external/default/WDSContentServer/IW3P/IB/2013/01/17/000158349_20130117102554/Rendered/PDF/wps6321.pdf.

Glass, M. R. (2015). Enhancing field research methods with mobile survey technology. *Journal of Geography in Higher Education*, 39(2), 288–298. http://dx.doi.org/10.1080/03098265.2015.1010144.

Mourao, S., & Okada, K. (2010). Mobile Phone as a Tool for Data Collection in Field Research. *World Academy of Science, Engineering and Technology, 70*(43), 222–226.