SDP 06 Handheld application for mobile data collection in the field

SDP 06.1 People

- Principal Investigator: Christine Borgman
- Faculty: Christine Borgman – UCLA, Information Studies
- Graduate Students: Matthew Mayernik, Alberto Pepe - UCLA, Information Studies
- Undergraduates: Erick Romero – UCLA, Computer Science

SDP 06.2 Overview

Environmental research data is primarily collected outside of lab settings in real world field locations. Field locations are highly variable; researchers adjust their in-field activities in response to the immediate situations, such as inclement weather, time constraints, and equipment breakdowns. Field deployment activities affect the data that are captured and the ways that they are interpreted post-deployment. We are developing the CENS Deployment Center (CENSDC) as one step to address the flexible and unpredictable nature of field research. The CENSDC is a web-based system for the collection of deployment information, such as equipment lists, field notes, and suggestions for future deployments. However, the web-based approach is limited by the lack of internet conductivity in most field settings. We believe that this problem could be addressed through the development of a handheld data collection, note-taking and microblogging annotation application for use during field-based environmental research. This report outlines our work in developing a handheld application for simple data collection and note-taking in the field.

SDP 06.3 Approach

This application builds on CENS expertise in development for handheld devices, particularly the EcoPDA and Campaignr projects. We plan to extend their current work to develop a generic module for deployment use that will supplement and interoperate with our CENSDC system, thus testing both interoperability and mobility issues in ecological field research. CENS researchers engaged in field research normally work in small teams in specific locales using ad-hoc methods and heterogeneous data sources, and adjust research activities on the fly. The goal of our research is to develop tools and applications that facilitate the collection of contextual data about CENS research activities and similar field-based sensor network research. Our development principles for the mobile application are the following:

- Facilitate collaboration and sharing: cyberinfrastructure initiatives promise increased collaboration across distances and increased sharing of informational resources. Neither of these can happen without standardized ways of describing research methods and products. Our application will utilize both existing standards, such as the Environmental Metadata Language and new methods, such as RDF vocabularies, to facilitate collaborations and sharing both within and across projects.

- Design for activity: Ecological researchers engage in various kinds of activities while in the field. Our application will be designed with an understanding of these activities, and the contexts in which they are to be performed.

- Facilitate emergent research dynamics: Field ecology is a highly emergent science. As researchers spend more time in a particular field location, they become more aware of the variables that are relevant to their study and adjust their research activities accordingly. This emergent organization of activity leads to particular research dynamics and patterns of activity. Our application must facilitate the creation and communication of emergent field dynamics within collaborative research.

These principles lead to our development goals. We want the application to enable researchers to:

- Specify data collection protocols: Researchers collect highly varied kinds of data. The application must allow researchers to specify their own collection procedures, and customize the data collection interface to those procedures.
• **Collect repeatable data**: Researchers often repeat data collection procedures, both to repeat and augment prior experiments. Our system must allow them to re-use data collection protocols that they have already created and used.

• **Perform their field activity as fast with the tool as without the tool**: Time spent in the field is precious. If our system noticeably slows researchers down, its usefulness is greatly diminished.

• **Easily integrate the field data collected with the tool into their existing data collections**: Researchers in most sciences are already facing challenges in organizing and using digital data resources. Our application is intended to be one step in mitigating this problem, thus our system must integrate with existing data and data tools.

• **Produce data process descriptions (collection procedures, annotations) that “live with” the data**: Data are collected in various ways. Understanding and using data requires understanding of the way that data were collected. Providing additional information about the data collection processes that “lives with” data increases the usefulness, and consequently the value, of the data itself.

**SDP 06.4 System(s) Description and/or Experiments**

The first main functionality of the application is data collection. The application will allow researchers to collect data in tabular form. More specifically, the data collection functionality will contain the following features:

• “Authoring” – the web interface will allow the user to create spreadsheets and data collection protocols before going out in the field on their personal computer. They can then upload the spreadsheets as XML files to the handheld device for use in the field. These data collection forms will customizable, so that they can be changed, and re-usable, to facilitate repeatability.

• Import/export features to Excel, csv, Sensorbase

The second main feature of our application is the note-taking/microblogging functionality. “Microblogging” refers to the use of web services such as Flickr and Twitter to enable researchers to post pictures and short notes to the web from field locations. Both notes and photos can be posted as microblogs. Notes or photos can be associated with particular items (locations, equipment, serial numbers, tasks, people, pictures, etc.) as metadata. For example, when making a note, it will be associated with one or more aspects of the current data collection activity, including the current transect, current location, current person, current data point, etc.

**SDP 06.5 Accomplishments**

Work on this project began in June of 2008, and to date has focused on doing background research, gathering user requirements, building models and specifications, and beginning prototype development. Figure 1 shows the concept map that was created as the requirement building and specification process. Another part of our background work has included identifying existing CENS code that we may reuse/customize to our project. We are using Windows Mobile 6 as the basis for our application, which allows us to build off the Windows Mobile-based EcoPDA project. For certain functionalities, we can also be able to build off of the code base of the CENS Campaignr project, which has recently been migrated to the Windows Mobile platform as well.
SDP 06.6 Future Directions
Future work will focus on iterative development and testing of the system, first of prototypes and subsequently of fuller systems. Initial tests will be with the CENS beach monitoring group, with further tests to follow.

In the longer term, we would also like to build in additional functionalities that allow for more flexible and nuanced data collection than is possible with simple spreadsheet tables. These additional functionalities could include:

- Enable researchers to run simple statistical checks or models while in the field as a means of determining sample sizes, densities, etc.
- In-field data graphing or visualization
- Support multiple kinds of data collection procedures: transects, quadrats, time-series.

SDP 06.7 External Research Partnerships
Microsoft Research (current)