Introducing Sapelli: A mobile data collection platform for non-literate users

Matthias Stevens Michalis Vitos Julia Altenbuchner
Gillian Conquest Jerome Lewis Muki Haklay

Extreme Citizen Science (ExCiteS) group, University College London
Gower Street, London, United Kingdom
{m.stevens, michalis.vitos.11, julia.altenbuchner.11, g.conquest.11, jerome.lewis, m.haklay}@ucl.ac.uk

ABSTRACT

With this poster we announce the imminent release of *Sapelli*, a new mobile data collection and sharing platform designed with a particular focus on non-literate and illiterate users.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

General Terms

Design, Human Factor

Keywords

Citizen Science, Literacy, Development, Conservation

1. INTRODUCTION

The sustainable management of natural resources in general, and key ecosystems such as tropical rainforests in particular, is one of the core development challenges humanity faces. An important avenue towards promoting sustainable natural resource management as well as environmental justice is to enable local people, who depend on these environments, to share their knowledge more effectively locally and with other regional, national and global stakeholders.

UCL's Extreme Citizen Science (ExCiteS) research group[1] is involved in multiple initiatives that aim to give indigenous communities the means to monitor, analyse and act on environmental trends, regardless of their literacy or technical ability. By introducing appropriate ICT solutions through a methodology incorporating action research and participatory design, we aim to enable cost-effective, bottom-up and better-informed environmental management by providing *all* stakeholders – from local people to remote scientists – with long-term, reliable, comprehensible information on environmental change and its socio-economic consequences.

[1]http://ucl.ac.uk/excites

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author. Copyright is held by the owner/author(s).
ACM DEV 4, December 6–7 2013, Cape Town, South Africa
Copyright 2013 ACM 978-1-4503-2558-5/13/12...$15.00.
DOI 10.1145/2537052.2537069.

To achieve this goal we are building *Sapelli*, a mobile data collection and sharing platform designed primarily (yet not exclusively) for non-literate or illiterate users with little or no prior ICT experience. The platform will eventually become an end-to-end solution for the collection, transmission, storage, visualisation, analysis, editing and sharing of environmental data. We intend to make Sapelli freely available to anyone and to make it open source. We focus here on the data collection and transmission elements of the platform that will be released in early 2014, with other elements gradually following.

2. CONTEXT & MOTIVATIONS

The design of the Sapelli platform itself, and of the methodology for community engagement which it underpins, is informed by past and current projects in which we collaborate with indigenous forest communities in the Congo Basin. These vulnerable communities, most of whom are non-literate, are rarely involved in the management of the land and resources on which they depend and current conservation or resource extraction efforts often leave them disenfranchised.

From 2005 Jerome Lewis and others worked to set up a scheme in which members of a Pygmy community living in Congo-Brazzaville engaged in monitoring of logging activity. Using PDAs with a pictorial decision tree interface participants recorded their resources and violations thereof. The information was shared with the local logging company such that they could better respect locals' concerns. In 2007 a similar initiative was set up in Cameroon [4, 5].

In 2012 we began working on a monitoring scheme relating to another pressing issue for forest communities: the activities of commercial poachers. The goal is to let locals record evidence of poaching through a pictorial decision tree. As discussed in [6] for this purpose we initially developed an Android application on top of the Open Data Kit (ODK) platform [2].

Also in 2012, we partnered with Forests Monitor, an international NGO, and OI-FLEG, an EU-funded forestry sector watchdog in Congo-Brazzaville, to develop a tool and methodology that would enable local populations to participate in monitoring the socio-economic impacts of logging. Similarly to the earlier logging-related projects, the main goals are to enable locals to give direct feedback on the behaviour of logging companies, and allow them to accurately map important resources they want to claim and protect from destruction. In 2013 we started working with Forest Peoples Programme, another international NGO, to prepare for a similar logging-related project set in the Democratic Republic of Congo, with possible expansion to other Central-African countries.

While on an abstract level these projects have a lot of common traits, local circumstances and user-specific requirements vary a great deal. This calls for a flexible software stack that is generic enough to be used across projects but which can be easily reconfigured to match and evolve with project specific requirements. This is true for both the data collection tools and data transmission mechanisms.

However, the experience with the initial 'Anti-Poaching' app [6] taught us that ODK – in its then form – was not the flexible solution our projects required. The main drawback is that ODK Collect, due to its XForms-based survey description format, is not well suited for hierarchical data input flows. Even relatively simple and compact decision trees require extremely verbose and complicated XForms code to be written. Clearly this limits the ability to flexibly adapt decision trees to changing local circumstances (e.g. in the field), as well as to reconfigure them for other contexts. Another requirement which ODK did not satisfy was the ability to transmit data to a central server in an autonomous and multi-modal fashion.

After evaluating other platforms (e.g. EpiCollect [1] & CyberTracker [3]) we found that none of them satisfyingly met our requirements. This has led us to develop our own data collection and transmission platform to support our on-going projects related to poaching and logging, as well as future participatory monitoring and mapping projects, also beyond the African context. We named our platform after the endangered sapelli tree (*Entandrophragma cylindricum*) which is important to Pygmy communities as it is a source of caterpillars which are part of their diet.

3. THE SAPELLI PLATFORM

The Sapelli platform currently consists of 3 main components: a data collection app (with integrated data transmission service) for Android devices, another Android app (called the 'relay') to forward SMS messages, and a server to receive and store data. All components are implemented in Java.

Due to our current focus on low and non-literate users the initial priority was to make it as straightforward as possible to build pictorial decision trees and icon-driven interfaces in general. Unsatisfied with the survey description languages used in other platforms, we designed our own XML-based format. Decision trees or conditional constructs in general, are built by nesting XML nodes, where the outermost node represents the first decision that must be made. Users navigate the decision space by repeatedly selecting a child node until they reach a leaf node, which represents a final selected value. This hierarchic description makes the structure of the decision space immediately apparent by looking at the XML code. In order to avoid repetition, "jumps" to another branch of the same decision tree are supported. Capturing of photos, audio recordings and location with GPS coordinates can easily be enabled by adding an according XML tag.

Our platform provides an autonomous, multi-modal data transmission mechanism to submit survey entries to a central server. The data collection app includes a background service that automatically checks for connectivity at scheduled intervals. In order to conserve power there is an option to put the device into flight mode between checks. When there is data to be sent, and a transmission opportunity arises, the service autonomously decides what to transmit and how, depending on available networks, bandwidth and project-specific settings.

To transmit the basic data (i.e. timestamps, decision tree selections, coordinates, etc.), records are serialised in a binary format which is heavily optimised for space. Next, these are grouped together in transmissions that can either be sent via SMS[2], or HTTP (over cellular or Wi-Fi networks). To further reduce bandwidth requirements transmission payloads can be compressed (using either DEFLATE or LZMA). SMS messages are sent to another phone (preferably located in the same country and with reliable Internet access) running our relay application, which forwards the messages to the server.

Due to the large file size, the transmission of optional media attachments will be postponed until internet connectivity is detected. The attachments are reliably associated with their corresponding records received earlier through a hashing algorithm. The features of the server component are currently limited to receiving and storing data and generating reports in CSV, XML, KML and Shapefile formats.

Of course, a text-free app will still be difficult to use for low and non-literate users if it runs on an operating system with a complicated, text-heavy interface such as Android. To tackle this problem we have developed an optional component, called the *Sapelli Launcher*. This app completely replaces the standard Android UI with a restricted, text-free app launching interface that only shows icons for a set of allowed apps, which can be tailored based on project requirements and user abilities. To prevent unauthorised access, apps can be protected with a mechanism similar to Android's pattern unlock feature. We envisage that Sapelli Launcher will be useful in other cases, beyond our specific platform.

While still a work in progress, the platform provides a number of unique features and meets the basic data collection and transmission requirements of our current projects while allowing for reuse in other contexts. It also provides a solid foundation for future extensions. Planned features include a Web-interface for survey design, project management, and querying, analysis and visualisation of data. In early 2014 we intend to release the Sapelli platform under an open source license and welcome contributors to help us develop it further.

References

[1] D. M. Aanensen et al. "EpiCollect: Linking Smartphones to Web Applications for Epidemiology, Ecology and Community Data Collection". In: *PLoS ONE* 4.9 (Sept. 2009), e6968. DOI: 10.1371/journal.pone.0006968.

[2] Y. Anokwa et al. "Open Source Data Collection in the Developing World". In: *Computer* 42.10 (Oct. 2009), pp. 97–99. DOI: 10.1109/MC.2009.328.

[3] *CyberTracker*. URL: http://www.cybertracker.org.

[4] J. Lewis. "Technological Leap-frogging in the Congo Basin, Pygmies and Global Positioning Systems in Central Africa: What has happened and where is it going?" In: *African study monographs. Suppl. issue* 43 (2012), pp. 15–44.

[5] J. Lewis et al. "Accessible technologies and FPIC: independent monitoring with forest communities in Cameroon". In: *Biodiversity and culture: exploring community protocols, rights and consent*. Ed. by K. Swiderska et al. PLA 65. June 2012. Chap. 13, pp. 151–165.

[6] M. Vitos et al. "Making local knowledge matter. Supporting non-literate people to monitor poaching in Congo". In: *3rd ACM Symposium on Computing for Development (DEV 2013; January 11-12, 2013, Bangalore, India)*. DOI: 10.1145/2442882.2442884.

[2]Transmissions can span up to 16 binary SMS messages.

Introducing Sapelli: A mobile data collection platform for non-literate users

Matthias Stevens, Michalis Vitos, Julia Altenbuchner, Gillian Conquest, Jerome Lewis & Muki Haklay

Sapelli is a new mobile **data collection** and **sharing platform** designed with a particular focus on **non-literate** and **illiterate** users with **little or no prior ICT experience**. Developed by UCL's Extreme Citizen Science (ExCiteS) research group Sapelli aims to provide indigenous people with tools that empower them to take action to protect their local environment and way of life.
The platform plays a central role in ExCiteS' mission – which is to develop theories, tools and methodologies to enable **any community, anywhere** to engage in citizen science – and will soon be made broadly available.

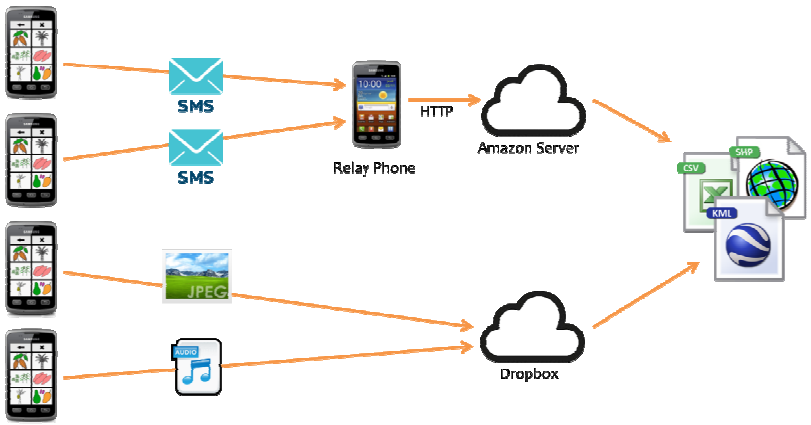
Context and Motivations

The **design** of the Sapelli platform and of the **methodology** for community engagement, which follows a Free Prior and Informed consent (FPIC) approach, is informed by past and current projects in which we collaborate with **indigenous forest communities** in the Congo Basin. These communities, most of whom are non-literate, are rarely involved in the management of the land and resources on which they depend and current conservation and resource extraction efforts often leave them **disenfranchised**.

Sapelli architecture

Currently the platform consists of four components:

- a data collection app for Android devices, that offers pictorial decision trees and icon-driven interfaces;
- an Android app (called the 'Relay') to forward SMS messages;
- an Android app (called the 'Launcher') to replace standard UI with a text-free app launcher;
- a cloud-hosted server component to receive and store data.



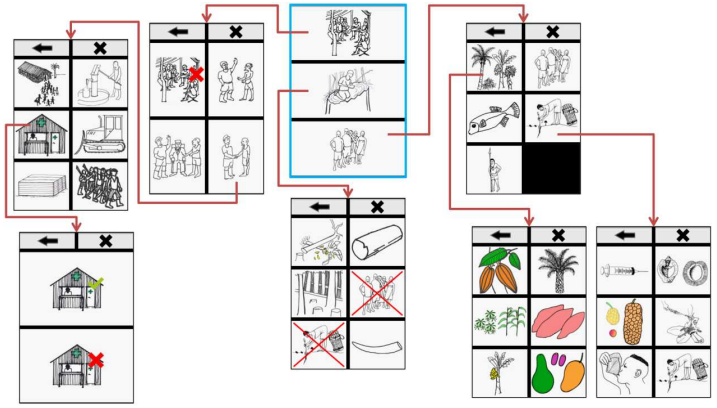
Sapelli Architecture

The **Sapelli data collection app** allows autonomous and multi-modal data transmission, where the app:

- automatically checks for connectivity;
- provides option to put device into flight mode between checks to conserve battery;
- decides what to transmit and how depending on available networks and bandwidth. Data is transferred either by binary, compressed SMS messages that are send to **Relay** app, or by HTTP (over cellular or Wi-Fi networks);
- optionally uploads media attachments (photos, audio recordings) to Dropbox.

The **Sapelli Launcher** allows to keep consistency, prevent unauthorised access to the device's apps and settings, and is used to:

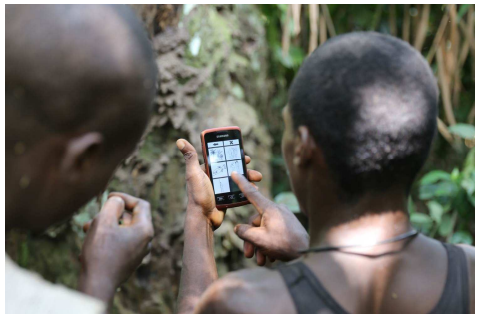
- completely replace standard Android UI with a text-free app launcher interface;
- show only allowed apps;
- protect apps via pattern locking mechanism.

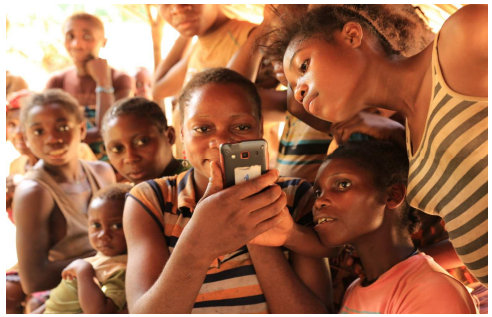


Example of a decision tree-based interface, enabled by Sapelli

Evaluation & Future Work

While still a work in progress, the platform was tested during a fieldtrip in the Republic of the Congo, where eight sites were visited and more than 270 participants had the chance to try the application. The platform provides a number of unique features and meets the basic data collection and transmission requirements of our current projects while allowing for reuse in other contexts. Planned features include a Web-interface for survey design, project management, and querying, analysis and visualisation of data.
In 2014 we will open up the Sapelli platform to be used by anyone, as well as releasing the code under the an open source license. We welcome outside contributors to help us develop it further. Keep an eye on **http://sapelli.org** if you are interested!





ExCiteS Research Group
Dept. of Civil, Environmental and Geomatic Engineering
University College London, Gower Street, London WC1E 6BT, UK
www.ucl.ac.uk/excites @UCL_ExCiteS