Experiences from Extreme Citizen Science: Using smartphone-based data collection tools with low-literate people

Artemis Skarlatidou

University College London

London, UK

a.skarlatidou@ucl.ac.uk

Fabien Moustard

University College London

London, UK

f.moustard@ucl.ac.uk

Michalis Vitos

University College London

London, UK

Michalis.vitos@ucl.ac.uk

Abstract

Citizen science gets increasing recognition for its potential to democratize science and support environmental governance. In this paper we present our experiences and lessons learned from a set of 'extreme' citizen science initiatives in developing countries, where data collection applications are used to support low-literate people in identifying solutions to issues that are of significant local concern. This paper aims to bring to the attention of the HCI community of developments in extreme citizen science and contribute knowledge to the field of HCI4D, especially to research studies which concern the design and use of smartphones for low-literate users.

Author Keywords

ICT4D; HCI4D; sustainability; citizen science

CSS Concepts

• **Human-centered computing~Human computer interaction (HCI)**; User studies; Participatory design

Introduction

Citizen science occurs when volunteers collaborate with scientists to produce scientific knowledge. Activities mostly involve data collection using scientific protocols

and may extend to the involvement of citizens in setting up the research questions, processing and analyzing the collected data. Fewer examples exist where the whole process is led by volunteers, without guidance from scientists. The increased momentum that citizen science has recently gained is due to recent technological developments in information and communication technologies (ICT), the availability and increasing use of mobile devices [17]. Apart from contributing to the scientific discovery, citizen science activities are used for advocacy purposes and as volunteer-initiated participatory action research projects to address issues of local concern [17].

In the environmental domain several citizen science applications exist; their aims and purposes vary, from enabling citizens to collect data about noise and air pollution to recording and monitoring biodiversity. This paper focuses on initiatives which enable the collection of traditional ecological knowledge for conservation and natural resource management and which support environmental governance and empower local communities. We use the term 'extreme', due to the geographic location of these initiatives (e.g., in remote forest locations) and also to emphasize that they are situated, bottom up practices which take into account local needs, practices and cultures with people participating at all stages.

Citizen science attracts people of all ages, backgrounds and interests. Although this has its own massive design challenges, so far citizen science focuses on a limiting demographic profile of the Western, Educated, Industrialized, Rich and Democratic [3], where

education skills, basic access to and familiarization with digital technologies are usually taken for granted. Extreme citizen science initiatives target mainly communities in developing countries, where education and literacy, access to technological infrastructure and familiarity with technology and local environmental conditions present various obstacles to their successful adoption and utilization.

The development of digital interfaces for low or non-literate people has been the subject of research in human computer interaction for development (HCI4D). In this paper we discuss a set of extreme citizen science initiatives and we present lessons learned from our experiences in developing and using data collection interfaces for low-literate users. Our aim is to bring to the attention of the HCI community the work we carry out and also contribute knowledge to the broader HCI context.

Relevant Background
UNESCO defines literacy as the ability of a person to read and write a short sentence [20]. Despite literacy levels growing dramatically during the last century, statistics from developing countries show that large segments of the population (i.e., 750 million youth and adults[1]) are still illiterate. Low literacy is a binding constraint for development; it influences technology adoption and use, since most technologies designed in the western world assume basic literacy skills.

One technology which assumes a reasonable amount of literacy is mobile devices. With growth mainly driven by developing countries it is expected that by 2025 mobile

---

[1] https://en.unesco.org/themes/literacy

subscriptions will reach 5.9 billion [4]. There is subsequently an interest in understanding how mobile phones are used by low-literate users for mainly communication tasks and the use of online services [1, 4, 8]. In this section we review key developments in interface design for low-literate users, mainly focusing on information structure and visualisation.

Iconic Interfaces for Low-literate Users
Research on interface design for low-literate people has traditionally focused on icon-based interfaces, with little or no text, due to their user-friendliness and user preference [15, 8]. Some studies suggest augmenting, rather than eliminating, text as these features can be memorized (as symbols) after repeated use. Additional modalities, such as audio feedback and voice annotation have also been tested and proved to be effective in specific contexts of use [8], despite their limitations due to the lack of speech recognition software in many colloquial dialects. When iconic interfaces are used, icon design should be fully embedded in local cultural contexts and meanings [8, 15, 6] and user preferences [8]. Evidence suggests that low-literate users understand better hand-drawn, semi-abstract graphics which incorporate action cues [15], although, this clearly depends on other factors such as local traditions, exposure to other forms of digital representations and symbolic literacies [16].

Menu Structures and Information Organization
Literacy does not only influence one's ability to read text. Education influences a person's cognitive abilities and linguistic sequential memory [14]. One of the most notable implications is that it impacts a person's ability to understand abstractions, which are commonly used in interface design for navigation purposes. Therefore,

several authors discuss low-literate people's difficulties in understanding and using menus that are based on hierarchies and they propose to use linear structures with up and down buttons or scrollbars to navigate them as there is evidence that people perform better [1,8]. Nevertheless, as digital skills improve through increased exposure to technology, low-literate people can overcome this problem and interact with hierarchical interfaces in similar ways to those of literate users [13].

Data-collection Tools for Extreme Citizen Science

Research around design issues in citizen science is still at an infant stage [17] and for extreme citizen science it remains mostly anecdotal evidence [16]. Several of the previous underlying principles, findings and research challenges, have provided the theoretical foundations for the development of the extreme citizen science tools: Sapelli and Tap&Map.

Sapelli uses decision trees which are set up to represent an ontology of data items for which data are collected. These have predefined answers associated with each data item. The leaves represent specific answers or classifications, while in-between nodes represent categories that lead to these final answers. Users navigate a decision space by repeatedly selecting images on their phones to select child nodes until they reach a leaf node. For example, a community might want to collect information about plants and animals (i.e., two top categories), followed by images of the plants and animals that are most important locally; e.g., an antelope or a banana tree. Once a category is selected, a set of questions about the data points need to be answered (e.g., was it seen or heard? Can you

Sapelli Collector App

Sapelli Collector (or Sapelli) is an open source data collection app for Android devices (Figure 1). It supports offline and autonomous data synchronization via SMS and the Internet. Its interface is icon-based, and information organization uses hierarchical decision trees, similar to other mobile-based applications for low-literate users.

Figure 1: Participants in the field using Sapelli (Photo courtesy of M. Vitos).

take a picture?). The location is recorded together with the additional information the user has been collecting.

Tap&Map (Figure 2) was developed to overcome interaction problems with Sapelli (e.g., navigating complex hierarchical structures and using the touch screen interface). The user, instead of navigating the decision tree to identify the correct data item, scrolls through card objects to identify the correct card and taps it on the phone to register the new information. The user needs to stand as close as possible to the location of the physical object, so that the device records the accurate location through its GPS sensor.

Extreme Citizen Science Methodology: Using A Participatory Design approach

Participatory design practices are motivated by democratic ideals inspired by a need to support "multiple voices in knowledge production, but [also] in the production of technologies as knowledge objectified in a particular way" [19]. Extreme citizen science initiatives rely heavily on participatory design, which is followed during the development and iterative design of Sapelli and Tap&Map implementations.

Extreme citizen science initiatives rely on communities defining the problem themselves. Two are the key methodologies which always initiate a participatory process: a. a Free, Prior and Informed Consent (FPIC) process and b. establishing a community protocol. The *FPIC process* aims to inform "the affected persons about planned activities and their impacts" and verify "that the information provided has been understood, before explicit consent can be negotiated" [12]. The consent is free and informed highlighting the ability of communities to refuse an intervention and it takes

place prior to them being affected by external actions [12]. This process further sets the foundations for local capacity building [12,16]. Researchers and local communities work towards establishing a *community protocol* to define project's processes: e.g., responsibilities and types of data to be collected and analysed, maintaining technical equipment, incentives, data ownership, identification and agreement of short and long-term benefits for the community.

The design and implementation of extreme citizen science tools follows an iterative approach to match and reflect on community needs. Interface design starts with the community co-designing the data-collection icons. Local meanings and cultural conventions are taken into account to ensure icons are well-understood. Preliminary user interactions with our tools are monitored using ethnographic methods and participant observation. At this stage we have experienced the need to revise the project aims collectively with the community, mostly in terms of simplifying the data that need to be collected, to better fit the context of their daily activities.

As the projects evolve the tools are continuously improved to address additional issues and community protocols are revised to inform technology maintenance. Communities are not forced to collect data, but they are being observed as they slowly incorporate the tools into their daily activities. Researchers in the field provide the necessary support to create maps which visualize the data. This is not only helpful for the communities but findings are discussed with the community and other stakeholders (e.g. local NGOS) and information is used to provide the evidence needed for advocacy purposes, as in [11] where Sapelli

Tap&Map App

 Tap&Map is a smartphone application accompanied by a set of cards equipped with near field communication (NFC) technology. Each card has an icon printed on one side (Figure 2), which represents a data item for which information is collected.

Figure 2:Participant in Namibia using Tap&Map (Photo courtesy of M. Laws).

supported local people claim their legal rights which were later incorporated in the Cameroonian Forest Act.

OVERVIEW OF EXTREME CITIZEN SCIENCE CASES

The lessons that we present in this paper are drawn from various extreme citizen science initiatives; some of which we present in this section.

Congo-Brazzaville: Reporting Illegal Logging with Mbendjele Yaka Pygmies (2013 onwards)
Mbendjele Yaka Pygmies of northern Congo-Brazzaville experience unprecedented exploitation of their local forest and natural resources. For these egalitarian communities the forest is a source of livelihood and it is important for its spiritual role. Sapelli was initially developed to support local communities and in collaboration with local NGOs; it is used since then by non-literate people collecting data about illegal logging and poaching in the area [21]. The vision is that data provide the evidence to report logging companies' behavior and subsequently inform the new EU Forest Law Enforcement Governance and Trade legislation.

Kenya: Collecting Data for Indigenous Plants with the Maasai (2019 onwards)
In Kenya, semi-literate communities use Sapelli; some have been previously exposed to digital technologies. One such communities is the Maasai warriors in Narok county who face the loss of their traditional ecological knowledge and increasing deforestation in the Maasai Mara National Reserve [10]. Sapelli is used to collect and record indigenous plants. Since the initiative was launched, individuals have gathered thousands of points with information of medicinal and other properties of local indigenous flora.

Namibia: Natural Resource Management and Fighting Illegal Cattle Invasions with Ju|'hoansi (2015 onwards)
The Nyae Nyae Conservancy in Namibia, registered in 1998, is under threat due to extensive cattle farming in traditional hunting and gathering grounds. As the primary custodians of the Nyae Nyae conservancy the Ju|'hoansi, mostly semi-literate, asked for support to collect data for fighting illegal cattle invasion in their territory. Communities also use Sapelli to collect data for local community forest resources [9].

Brazil (Pantanal Wetlands): Natural Resource Management for New Conservation Legislation with Indigenous Communities (2014 onwards)
Mainly situated in Brazil, the Pantanal is the largest wetland in the world. Local fishers are dependent on it for their daily livelihood. Current resource management legislation follows scientific recommendations of conservation biologists, and gradually led into people's physical and economic displacement. Local communities, mostly semi-literate, used Sapelli to collect data about the use of natural resources and their strategies. The data provided evidence that indigenous practices are indeed sustainable. As a result, local people were recognized as a traditional community and this provided them the right to protect their livelihoods using traditional practices [2].

Cameroon: Supporting Baka communities Tackle Illegal Wildlife Crime and Animal Monitoring (2015 onwards)
Baka hunter-gatherers and Mbulu farmers of Cameroon live in forest inside Dja Biosphere Reserve, which hosts a variety of plants and animals to support livelihood but is currently being depleted by illegal wildlife trade and extractive industries. Current conservation legislation excludes local knowledges and turns local communities

into 'conservation refugees'. Non-literate communities use Sapelli to collect data about illegal wildlife crime and animal monitoring to inform forest management. Due to several implications and benefits from illegal forestry, privacy is considered a critical concern in this case; serious threats towards individuals influence their motivation to join the anti-trafficking efforts [5].

EXPERIENCES FROM EXTREME CITIZEN SCIENCE

Every initiative is unique, but some experiences are repeated across our initiatives, which we present in this section.

Understanding Local, Cultural Contexts and Ways of Living and Building Trust
Understanding technology adoption in specific contexts requires understanding people's world views, cultures and traditions. Participatory design enables focusing on the critical role of culture. In extreme citizen science it is driven by anthropological accounts of technological practices. First, community protocols create a 'third space' for researchers and community members to critically reflect on technology use and how it fits their culture and context. Our researchers gradually integrate with the community over long periods of time (i.e., over a year). They support people with everyday tasks, learn their language, listen to their stories, histories and personal opinions. This helps to promote mutual respect, bottom-up collaboration and build trust.

Extreme citizen science initiatives rely on communities identifying the problem themselves, but their real problems are usually wrapped behind local politics and they are difficult to expose. Past or current projects,

driven by governments and other organizations, have left a set of unfulfilled promises resulting in a legacy of distrust. Building trust has been identified the most crucial aspect of our work. To do so it is also necessary that during the FPIC and community protocol establishment processes, that people's timings, priorities and local decision-making processes are fully understood and respected.

Technology Design
Sapelli and Tap&Map use an icon-based design for data input. With low-literate people who have not been exposed to digital representations, the majority of the icons they prefer are hand-drawn (e.g., in Congo-Brazzaville, Namibia, Kenya), whereas those who were previously exposed to digital representations requested the use of photorealistic images (e.g., in Pantanal). It was found that people mostly find it difficult to understand abstract icons used for Sapelli's hierarchical structure [21], which was not an issue with Tap&Map as it does not use hierarchies.

Sapelli's hierarchical structure, in most cases, is an interaction barrier, especially when complex decision trees are used. In some cases, communities requested simplifying decision trees. For example, in Pantanal the initial Sapelli decision tree was split into two, to create two different apps; one for fishermen and one for bait catchers. With each community having its own app unnecessary complexity was removed and it was easier to interact with the hierarchies. Also following HCI4D recommendations for the utilization of linear interfaces, we also tested this feature with low-literate users in London and Namibia with encouraging findings [18].

Communities are motivated by viewing the data they collect [16]. A relatively cumbersome process, since data need to be extracted into a mapping software, which usually led by trained participants or the researchers. Being able to view the data instantly is important for motivation purposes, error recovery and prevention. The fact that several of the communities have never used a map before, or other spatial representations, means that development of a data visualisation component requires extensive research and testing, which is currently work in progress.

Using Conventional HCI Methods
Conventional HCI methodologies are not culturally universal [6]; there are several assumptions which underlie their theoretical and practical implementations which may lead to biases or inadequate insight. In our cases in Cameroon and Congo-Brazzaville we noticed that criticism is considered to be highly offensive. This applies to technological artefacts, which are seen as extension of the researcher's work. Obtaining honest and constructive feedback and 'what if' scenarios to improve interface design is a major obstacle and therefore running usability tests and think aloud become inappropriate [16]. Incorporating testing into daily activities to evaluate interface design versions provided us with enough insight to improve usability [21]. Testing is organized as a group activity as in egalitarian societies these are treated as collaborative tasks [21]. A major obstacle is proximity to end users. Since the team is London-based, it is not always possible to evaluate prototypes in one go or obtain user feedback at all stages of the development process. We have been working with proxy users, as in [7], to include textually and digitally illiterate people who live in Europe [18]. This comes with its own assumptions

and complexities and further work is required to evaluate the effectiveness of this approach.

Promoting Interdisciplinarity Across Multiple Disciplines for (Extreme) Citizen Science
Citizen science is important for democratizing science, and extreme citizen science can further empower marginalized and disempowered to participate in the debate of environmental sustainability. For extreme citizen science initiatives to work an interdisciplinary research insight is required. We have experienced significant challenges in promoting interdisciplinarity. Perhaps the most complex barrier is to overcome individuals' disciplinary attitudes and presumptions around domain specificities. Collaborative interactions and communication of scientific research insights into disciplinary domains is needed to improve awareness and research quality and eventually create more successful (extreme) citizen science initiatives.

ACKNOWLEDGEMENTS

REFERENCES
[1] Chaudry, B., Connelly, K., Siek, K.,Welch, J. 2012. Mobile interface design for low-literacy populations. *In 2nd ACM SIGHIT International Health Informatics Symposium*, (Miami, FL), 91-100.

[2] Chiaravalloti, R. 2019. The Displacement of Insufficiently 'Traditional' Communities: Local Fisheries in the Pantanal*. Conservation and Society.* 17 (2), 173-183.

[3] Dourish, Paul. 2015. Forward. In Bidwell, Nicola, Winschiers-Theophilus, Heike. *At the Intersection*

of Indigenous and Traditional Knowledge and Technology Design. USA, California: ISP.

[4] GSMA Intelligence. The Mobile Economy: Sub Saharan Africa 2018. https://www.gsma.com/mobileeconomy/sub-saharan-africa-2017/

[5] Hoyte, S. 2017. Indigenous Baka hunters vs. The illegal wildlife trade https://simonhoyte.wordpress.com/2017/03/01/indigenous-baka-hunters-vs-the-illegal-wildlife-trade/

[6] Irani, L. (2010). HCI on the move: Methods, culture, values. *In CHI'10 extended abstracts on human factors in computing systems,* Atlanta, GA, pp. 2939–2942.

[7] Knoche, H., Huang, J. 2012. Text is not the enemy: How illiterates' use their mobile phones. In *CHI on Human Factors in Computing Systems,* USA.

[8] Lalji, Z., Good. J. 2008. Designing new technologies for illiterate populations: A study in mobile phone interface design. *Interacting with Computers.* 20 (6), 574-586.

[9] Laws, M. (2015). Sapelli to tackle illegal cattle invasions for the JuHoansi of NyaeNyae conservancy. https://uclexcites.blog

[10] Laws, M. 2019. Designing data collection interfaces for semi-literate rangers in rural Namibia. *Unpublished Field Report.*

[11] Lewis, J. (2002). *Forest Hunter-Gatherers and their world: A study of the Mbendjele Yaka Pygmies of Congo-Brazzaville and their secular and religious activities and representations*. Ph.D Dissertation. LSE, London.

[12] Lewis, J. (2012). How to implement free, prior informed consent (FPIC). *In Biodiversity and culture: Exploring community protocols, rights and consent*. London, UK: International Institute for Environment and Development, 175.

[13] Medhi, I., Gautama, S., Toyama, K. 2009. A comparison of mobile money-transfer UIs for non-literate and semi-literate users. In *SIGCHI Conf. on human factors in computing systems,* (Boston, USA), 1741-1750.

[14] Medhi, I., Raghu Menon, S., Cutrell, E.,Toyama, K. 2010. Beyond Strict Illiteracy: Abstracted learning among low literacy. In *4th ACM/IEEE Int Conf on Information and Communication Technologies and Development,* (London, UK), 1-9.

[15] Medhi, I., Sagar A., Toyama K. 2006. Text-Free User Interfaces for Illiterate and Semi-Literate Users. *Information and Communication Technologies and Development,* 4 (1), 37-50.

[16] Pejovic, V. and Skarlatidou, A. 2019. Understanding Interaction Design Challenges in Mobile Extreme Citizen Science. *International Journal of Human-Computer Interaction,* 1-20, DOI: 10.1080/10447318.2019.1630934

[17] Preece, Jennifer. 2016. Citizen Science: New Research Challenges for Human-Computer Interaction. *International Journal of Human-Computer Interaction* 32 (8), 585-612.

[18] Skarlatidou, A., Trimm, C., Vitos, M, Haklay, M. 2019. Designing Data Collection Interfaces for low-literate users. In 17th European Conf. on Computer-Supported Cooperative Work, (Salzburg, Austria).

[19] Suchman, L. 2002. Located accountabilities in technology production. *Scandinavian journal of information systems*, 14 (2), 91-105.

[20] UNESCO (2006). Chapter 6 understandings of Literacy. http://www.unesco.org/education/GMR2006/full/chapt6_eng.pdf.

[21] Vitos, M. 2018. *Making local knowledge matter: Exploring the appropriateness of pictorial decision trees as interaction style for non-literate communities to capture their traditional ecological knowledge.* Ph.D Dissertation. UCL, London.