What do volunteers want from citizen science technologies? A systematic literature review and best practice guidelines

Artemis Skarlatidou, Alexandra Hamilton, Michalis Vitos and Muki Haklay

Abstract   Although hundreds of citizen science applications exist, there is lack of detailed analysis of volunteers' needs and requirements, common usability mistakes and the kinds of user experiences that citizen science applications generate. Due to the limited number of studies that reflect on these issues, it is not always possible to develop interactions that are beneficial and enjoyable. In this paper we perform a systematic literature review to identify relevant articles which discuss user issues in environmental digital citizen science and we develop a set of design guidelines, which we evaluate using cooperative evaluation. The proposed research can assist scientists and practitioners with the design and development of easy to use citizen science applications and sets the basis to inform future Human-Computer Interaction research in the context of citizen science.

Introduction

The ubiquity of technology and the Internet has dramatically changed the landscape of information availability, including the ways in which we interact with it to make better decisions and improve our quality of life. Technological innovations have resulted in changes not only in the economy and the workplace, but also in the ways people choose to live their lives, spend their free time, and interact with others. Such changes have led to social innovations, which have ushered in a new wave of social change. One such change took place within the scientific context, with the ongoing growth number of amateur volunteers, with the help of technology, now work together with scientists to explore and address scientific issues.

This collaboration, or partnership, between professional scientists and amateur volunteers is known as citizen science (CS). In its simplest form it involves amateur

scientists collecting data; an activity which reduces the costs of addressing scientific questions that require the collection of massive amounts of data, and which further bridges the intellectual divide magnified by the professionalisation of science, and the scientific expertise that it may assume. Many CS scholars situate the activity over two centuries ago when amateur scientists, such as Charles Darwin, made significant contributions to science [Silvertown, 2009]. Currently, hundreds of CS projects engage thousands of volunteers across the world. A relatively recent analysis of 388 CS projects revealed that they engaged 1.3 million volunteers, contributing up to US$2.5 billion in-kind annually [Theobald et al., 2015]. The eBird project alone collects five million bird observations monthly, which has resulted in 90 research publications [Kobori et al., 2016]. Technological innovations such as the Internet, smartphone networked devices equipped with sophisticated sensors, and high resolution cameras have had a massive impact on science, especially the way in which CS is currently practiced.

Volunteers who interact with these technologies come from different age groups and geographic locations; their cultural contexts and the languages they speak may vary; as do their skills, motivations and goals. A typical user might be an MSc marine biology student who uses a mobile CS application to identify invasive species, or equally an illiterate hunter-gatherer in the Congo basin who collects data to address the challenges of illegal logging and its impact on local resource management. Making sure that users can fully utilise the technology at-hand is fundamental. Within this context, Preece [2016] calls for a greater collaboration amongst *"citizens, scientists and HCI specialists"* (p. 586). Human-computer interaction (HCI), the discipline which studies how humans interact with computers and the ways to improve this interaction, has a long tradition in the design and development of technological artefacts, including aviation technology, websites, mobile devices, 3D environments and so on.

Volunteers' involvement in the use and design of CS applications is fundamental for effective data collection. Issues such as motivating users to remain active, ensuring that users can effectively use the applications, and guaranteeing satisfaction of use, should be central in the design and development of such applications; issues that according to Prestopnik and Crowston [2012] look beyond *"building a simple interface to collect data"* (p. 174). Existing studies which investigate user needs and requirements, usability, and user experience (UX) elements (such as having fun and joy) of CS applications provide interesting insights, usually in specific contexts of use. Nevertheless, there is a lack of reflection on the lessons learned from these studies. A more holistic overview is needed to explore the current state of HCI research within CS, and how this knowledge and preliminary empirical evidence can inform the design and development of 'better' CS applications. This paper attempts to address this gap. With the preliminary aim of identifying key design features, as well as other relevant interaction recommendations, a systematic literature review (SLR) was utilised to capture research studies which discuss user issues of CS applications that support data collection (either web or mobile based). This evidence and knowledge is summarised in a set of design guidelines, which were further evaluated in a cooperative evaluation (CE) study with 15 people.

Before we explain how these methods were employed in this study in section 3, the next section provides their theoretical background. Section 4 briefly presents the

results, and section 5 discusses, in more detail, the studies that we analysed in the SLR, together with the findings of our CE. Conclusions are drawn in section 6, focusing on limitations of this research and suggestions for future research.

| **Theoretical overview of methods used** | A systematic literature review (SLR) provides a standardised method for reviewing literature which is *"replicable, transparent, objective, unbiased and rigorous"* [Boell and Cecez-Kecmanovic, 2015, p. 161]. SLRs emerged in evidence-based medicine during the 1990s, but have since been found, in increasing numbers, within various fields including education, psychology and software engineering [Boell and Cecez-Kecmanovic, 2015]. SLRs have also been used in the context of HCI and CS. For example, Kullenberg and Kasperowski [2016] outline a scientometric meta-analysis, using datasets retrieved from the Web of Science, to investigate the concept of CS, its development over time, the research that it covers and its outputs. In their study they mention implications associated with the use of key terms (e.g. use of terms related to CS, such as crowdsourcing, participatory monitoring, public participation and acronyms such as PPGIS), but generally no other criticisms are mentioned. |

Zapata et al. [2015] use a SLR approach to review studies that perform usability evaluations of mHealth applications and their user characteristics on mobile devices. Due to the narrow focus, and the fact that the results were further analysed qualitatively, Zapata et al. [2015] only screen 22 papers (out of the 717 articles which were initially retrieved). Connolly et al. [2012] review the impacts of computer games on users aged above 14 years old. With an initial 7,391 articles retrieved from various databases, they applied a relevance indicator (from 1 least relevant to 9 most relevant) to each paper and analysed the papers that received a rank higher than 9, resulting in a final screening of 70 papers which discuss several dimensions of the detected impacts. Other studies investigate the accessibility and usability of ambient assisted living [Queirós et al., 2015]; studies which integrate agile development with user centred design [Salah, Paige and Cairns, 2014] to identify challenges, limitations, or other characteristics, thus requiring a qualitative analysis of the papers reviewed.

Although the previously mentioned studies offer a limited insight into the limitations of SLRs, the most widely acknowledged limitation surrounds the weaknesses of search terminology, as effective keyword combinations cannot be always known a priori [Connolly et al., 2012; Boell and Cecez-Kecmanovic, 2015]. Likewise, the process' internal validity, such as problems related to data extraction, the so-called recall-precision trade off, and a lack of evaluation, is also considered a limitation [Boell and Cecez-Kecmanovic, 2015]. To adjust for these limitations, this research evaluated the content of our meta-analysis in a user testing experiment. We employed cooperative evaluation method; a user testing method which allows for the observer to interact with the user subjects in order to direct their attention and record their opinion about specific design features while they work together as collaborators. Thus, the observer can answer questions and doesn't have to sit and observe in silence while the user interacts with the application involved in the experiment [Monk et al., 1993].

Furthermore, this research focuses on the environmental CS context. As an increasing number of CS projects in various fields (e.g. astronomy, linguistics)

make use of digital technology to engage their volunteers in data collection activities, only some of these projects consider user issues in their design [e.g. Sprinks et al., 2017]. The interfaces, tasks, and therefore design issues or user needs and requirements vary across different contexts of use. For example, in astronomy-related CS applications image classification is a major task, whereas in conservation applications tasks such as adding an observation using a mapping interface are more common. In order to narrow down the focus, and also due to the authors' interest in the way geographical interfaces assist volunteers in data collection, the term 'environment' has been included in all keyword searches. Although this might limit the scope and extent of retrieved studies, it should be noted that the term is already quite broad, since it may include applications from contexts where CS is widely used, including, but not limited to, ecology, conservation, environmental monitoring, etc. In the next section we review the methodological implementation and experimental design.

Methodological procedure and experimental design

A two-stage methodological approach was applied. In the first part, a systematic literature review (SLR) was conducted to review and evaluate relevant research studies; section 3.1 describes this process. In the second stage, discussed in section 3.2, 15 CE experiments were implemented to evaluate three online CS applications, which provided a context to discuss the preliminary guidelines with our participants.

Systematic literature review

The following databases were included in our search: Web of Science[1] of Thomson Reuters, Scopus[2] of Elsevier, and Google Scholar.[3] The sources comprise access to multidisciplinary research studies (i.e. peer-reviewed and grey literature) from the fields of sciences, social sciences, and arts and humanities. We decided to employ all of them to obtain reliable, robust, and cross-checked data; including a reasonable amount of 'grey literature' via Google Scholar, so that we could include in our search technical reports and government-funded research studies which are not usually published by commercial publishers [Haddaway, 2015]. The searches took place in June (i.e. Google Scholar) and December 2017 (i.e. Web of Science and Thomson Reuters) following the same methodological protocol. In order to automate the otherwise time-consuming search process, we used the free software "Publish or Perish", which allows importing a set of keyword combinations, executing a query on Google Scholar for each of them and collecting the results into a CSV file. The software was configured to execute 120 queries per hour.

Our database search aimed at retrieving papers with a focus on user issues of CS technologies in the environmental context. The majority of key terminology (i.e. user, citizen science, technology) is not standardised; for example, in the literature alternative terms that are used for user requirements may include "user needs", "volunteer needs" or "volunteer requirements", which were all included in the SLR. Similarly, other terms related to citizen science such as "participatory science" and "crowdsourced science" we deemed relevant as several of these terms are used

---

[1]Web of Science: http://apps.webofknowledge.com/.
[2]Scopus: https://www.scopus.com/search/form.uri.
[3]Google Scholar: https://scholar.google.co.uk/.

interchangeably in the literature and the similarities in the way these are practiced especially concerning user issues of those digital technology-based implementations. Considering that the search for the right terminology in SLRs becomes a search *"for certain ideas and concepts and not terms"* [Boell and Cecez-Kecmanovic, 2015, p. 165] which are relevant to the particular topic, we decided to include the concepts/keywords that are shown in Table 1; with the keyword 'environment' consistently applied to all searches. Our extensive concept identification analysis resulted in 1,045 keyword combinations, which significantly increased the effort and time needed to gather and analyse the data. SLR searches *"are optimised to return as many (presumably) relevant documents as possible thus leading to a high recall"* [Boell and Cecez-Kecmanovic, 2015, p. 165], which may be at the expense of data accuracy. Although each keyword search combination was designed to return only 200 results, we were confident that by including such a broad range of concepts we would be able to effectively capture the broader state-of-the-art literature in the area of environmental digital CS with a focus on user issues or detailed user studies.

Table 1. Concepts/Keywords used in the Systematic Literature review. Highlight in bold are the most popular terms in each category according to the number of articles retrieved.

Themes	*User-focus*	*Citizen Science - focus*	*Technology - focus*
Keywords	HCI, human-computer interaction, usability, **evaluation**, user experience, UX, user needs, user requirements, user research, human factors, UCD, user-centred design, human-centred design, accessibility, volunteer needs, volunteer requirements, user testing, volunteers' needs, volunteers' requirements	**Citizen science**, participatory science, DIY science, grassroot science, civic science, crowdsourced science, community science, community-based participatory research, community-based monitoring, public participation in scientific research, PPSR	**Technologies, technology**, digital, ICT, technology-mediated

For each result the following information was retrieved: title, author(s), date, number of citations, publisher and the article URL. After duplicates were removed and any missing fields (i.e. mainly dates) were manually filled in, we applied our exclusion criteria. These included: studies published before 2000, as digital CS only gained momentum in the mid-2000s [Wald, Longo and Dobell, 2016; Silvertown, 2009]; studies in languages other than English, as it would not be possible to further analyse their content. The studies that were included in the analysis all discussed digital CS in the environmental context (i.e. either focusing entirely on an environmental CS application or including examples of such applications) and had a user focus (i.e. recommendations for design features based on user studies or on user feedback; insight into user issues and design suggestions based on broader experience; expert inspections for usability improvements, etc). To ensure that these themes were all covered by the remaining studies the following were separately assessed by two of the authors: title, followed by the abstract and, if necessary, the introduction and conclusions. The final list of SLR papers (i.e. 62 papers in total) were all read in-full to inform our preliminary set of best practice guidelines. More than one paper mentioned the same design issue in order for it to be included it in the preliminary list of design guidelines.

Cooperative Evaluation (CE)

The preliminary guidelines that resulted from the meta-analysis of the final articles were further evaluated, through the use of CE experiments, and extended upon so that users' direct feedback is taken into account. For the purposes of this study, three applications were evaluated, which featured either a mobile or a web interface (or both). The three applications which were included in the CE experiments are:

- iSpot;[4] nature-themed CS application provided by Open University, with over 65,000 registered users (as in March 2018). The application provides users with a web-based interface, which people can use to upload geo-referenced pictures with various wildlife observations, explore and identify species, and connect with other enthusiasts worldwide (Figure 1a). It further supports gamification (via badges, which are popular in the engagement and retainment of online volunteers.).

- iNaturalist;[5] environmental CS application developed originally at UC Berkeley and now maintained by the California Academy of Sciences with over 0.5 million registered users (Figure 1b). It provides both a web-based interface and a mobile application to contribute biodiversity data in the form of photographs, sound recordings or visual sightings. The data are open access and available for scientific or other purposes.

- Zooniverse;[6] one of the most popular digital CS platforms, with over 1.6 million registered users[6] was developed by Citizen Science Alliance. The platform includes several CS projects, of which the most popular is the Galaxy Zoo project. The platform was included in the evaluation for the assessment of overall interface design and assessment of such design features as tutorials and communication functionality and therefore we decided to investigate these features within as well as beyond the environmental context (i.e. projects of environmental interest). The project pages included in the evaluation are: Wildlife Kenya, Shakespeare's World, the Elephant Project, Galaxy Zoo and Understanding Animal Faces.

It should be noted that iSpot and iNaturalist applications were chosen on the basis of their popularity and strong environmental focus. Zooniverse was chosen due to its popularity and the significant number of users that it attracts to mainly focus on design features of the communication functionality category which extends to the broader CS context.

Fifteen participants (seven male and eight female) were recruited, via opportunity sampling, to evaluate the applications and offer user feedback with respect to the SLR preliminary guidelines. Since all of the applications' projects are designed to be inclusive towards a wide user audience we decided against restricting participation to specific conditions, which would in turn influence the usefulness of

---

[4]iSpot: https://www.ispotnature.org/ It should be noted that while in the past iSpot had a dedicated Android app, it was withdrawn in February 2015, and the web-based interface is used on mobile phones, too.

[5]iNaturalist: https://www.inaturalist.org.

[6]Zooniverse: https://www.zooniverse.org/.

Figure 1. Interfaces of applications included in the CE experiments; a. iSpot; b. iNaturalist.org; c. Zooniverse.

this study. Nonetheless, the majority of our participants (9/15) were between 18–25 years old (with a further two between 25–34; three between 45–54 and one 55+). Only six of the participants had prior knowledge or experience with CS projects, all had experience interacting with online/mobile mapping interfaces and, in terms of their technological skills, five rated themselves as 'intermediate', six as 'advanced and four as 'experts'.

The applications were evaluated in a random order to minimise bias introduced by the learning effect. The experiments were carried out in July and August 2017. Each experiment was carried out over one hour and a set of tasks were provided to the users to guide their interaction with the application they interacted within the CE session. The tasks were designed to gain a better understanding of the design guidelines derived from the SLR (e.g. *Find the forum page and read discussion about 'adding photo to button' on iSpot; check notifications for new messages/comments on Zooniverse; check news feed for updates from iNaturalist community*). Of course more simple tasks that helped participants get a good understanding of the application's context of use were also included (e.g. *Explore photos for observations in your local area on iSpot; explore projects and then select a project to contribute an observation on Zooniverse*). A total of 10 tasks per application were provided to guide interaction and participants were asked to think aloud.

Results

Systematic literature review

The 1,045 keyword combinations returned 24,147 results from all three databases, dated from 1970 to 2017. In the three categories of keywords the most popular were: technology/technologies (13,781; followed by digital with 6,057 results); citizen science (8,793 results; followed by participatory science with 2,168 results) and evaluation (5,160; followed by usability with 3,038 and human computer interaction with 2,088 results).

After removing duplicates, non-English results and any publications dated before 2000 we ended up with 844 results (Figure 2). Figure 3 summarises the top ten keyword combinations returning the majority of results in this stage. This dataset includes articles and other results from various disciplines, mainly healthcare, education, environmental studies and HCI. The terms 'technology' or 'technology-mediated' are especially popular in education literature and 'community-based participatory research' is popular amongst health scholars, yet the majority of these articles do not satisfy the rest of the SLR criteria. Moreover, the keyword 'evaluation' returned almost 5,160 results, the majority discussing different aspects of evaluations (e.g. project, methodological or result evaluations), with only a few focusing on evaluations of user issues.
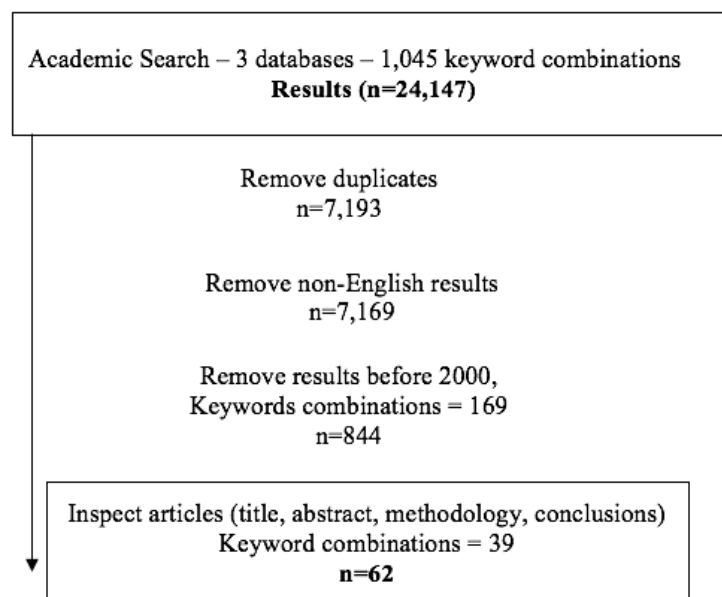

Figure 2. Systematic Literature Review procedure.

The next step involved inspecting all n=844 results (i.e. examination of title, abstract, methodology and conclusions); a process undertaken by two of the authors of this paper. During this process irrelevant results (i.e. studies that do not satisfy the search criteria) were excluded together with: i. Results that we could not access (e.g. conference proceedings, other n=47); ii. Books (not available online or in a digital form) (n=3 entries), as it would be difficult to further process content in the meta-analysis study; iii. Slideshare, poster presentations and keynotes (n=18), as they did not provide enough insight to inform meta-analysis; iv. Studies that targeted only children (n=4), as the user audience of all other studies included in the SLR is much broader and including these articles would introduce a bias; v. Studies

Top 10 keyword combination searches (n=844)
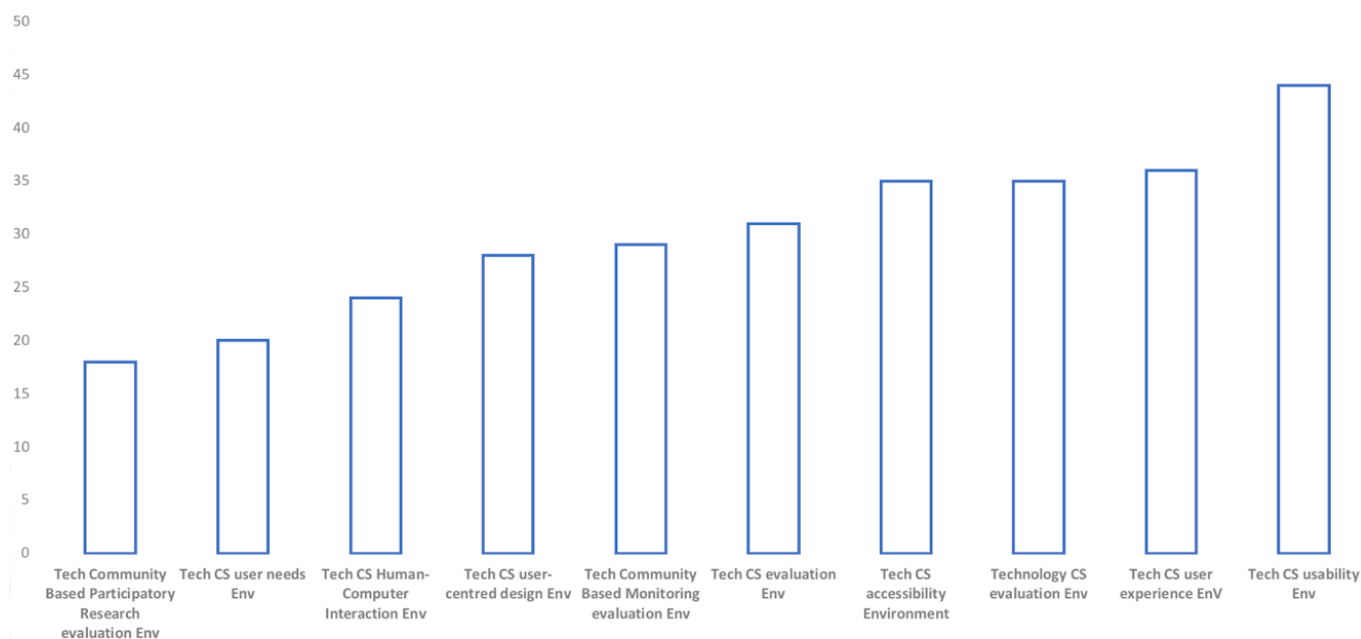
Figure 3. Ten most popular keyword combinations (for n=848) based on number of articles returned. 'Tech' is used for technology; 'CS' for Citizen science; and 'env' for environment.

discussing technologies at a conceptual level (i.e. applications do not yet exist) (n=4); and vi. Studies which describe sensors (non-phone based) and wearable technologies (n=2); these technologies have different characteristics that need to be considered for their effective design. Keywords such as "community-based monitoring" returned several results focusing on the development and user issues of Web Geographical Information Systems (GIS), which is commonly used in CS. We included only studies which described technologies to support data collection, and not studies on data visualisation for public consultation, another popular area in Public Participation GIS. Theoretical papers providing overviews of CS were included only when provided insight into technology-related user issues [e.g. Newman, Wiggins et al., 2012; Wiggins and Crowston, 2011]. Also we included papers on user design suggestions even if the focus was not entirely on a specific application [e.g. Jennett and Cox, 2014; Rotman et al., 2012].

Further application of the exclusion criteria resulted in 62 relevant articles suitable for meta-analysis. It should be noted that the keyword combinations at this stage vary with most popular being: "technology", "environment", "citizen science" and "usability" (n=7), followed by "user experience" (n=6), "human-computer interaction" (n=5) and user-centred design" (n=4). "Evaluation", which was the most popular keyword of initial search results, was the least popular amongst the final list of articles (n=62). Nine results use the term "technology-mediated", but "technology" was the most popular term in this category. Publication dates range from 2002 up to 2017, with the majority (n=55) unsurprisingly published after 2010. The results include three reports, one dissertation thesis, 28 journal papers, 27 conference proceedings, with most of them being peer-reviewed (e.g. CHI, CSCW, GI Forum), and three book chapters which are available online.

The final list includes several studies (n=18) where users are directly involved in usability evaluations [e.g. R. Phillips et al., 2013; Kim, Mankoff and Paulos, 2013; Jennett, Cognetti et al., 2016; D'Hondt, Stevens and Jacobs, 2013], co-design [e.g. Fails et al., 2014; Bowser, Hansen, Preece et al., 2014] and user-centred design processes [e.g. Woods and Scanlon, 2012; Newman, Zimmerman et al., 2010; Michener et al., 2012; Fledderus, 2016], in contexts such as wildlife hazard management [e.g. Ferster et al., 2013]; water management [e.g. Kim, Mankoff and Paulos, 2013]; environment-focused CS games [e.g. Bowser, Hansen, He et al., 2013; Bowser, Hansen, Preece et al., 2014; Prestopnik and Tang, 2015; Prestopnik, Crowston and Wang, 2017]; species identification, reporting and classifications [e.g. Newman, Zimmerman et al., 2010; Sharma, 2016; Jay et al., 2016]; and noise and ecological monitoring [e.g. Woods and Scanlon, 2012; Jennett, Cognetti et al., 2016]. Some of these studies include an impressive number of user subjects; e.g. Luther et al. [2009], in their evaluation of Pathfinder — a platform that enables CS collaboration — include over 40 users and Bowser, Hansen, Preece et al. [2014] in a co-design study for the development of Floracaching evaluated the application with 58 users. Several other studies involve users through surveys and interviews [e.g. Idris et al., 2016; Wald, Longo and Dobell, 2016; Eveleigh et al., 2014; Wiggins, 2013]; or other forms of user feedback such as online forums and discussions [e.g. Wiggins, 2013; Sharples et al., 2015].

Methods used within the studies included personas and user scenarios to identify user needs and requirements [e.g. Dunlap, Tang and Greenberg, 2013], and heuristic evaluations to examine virtual CS applications [Wald, Longo and Dobell, 2016]. Other studies did not include users directly, but still provided insights on user requirements, needs and other design issues from the design and development of a specific application or by reviewing existing ones [e.g. Connors, Lei and Kelly, 2012; Ferreira et al., 2011; Johnson et al., 2015; Newman, Wiggins et al., 2012; Stevens et al., 2014; Kosmala et al., 2016].

Six categories were created to inform the meta-analysis. The categories, created in agreement by two of this paper's authors who analysed the data and they were designed to effectively group key design features and which include: *Basic features and design recommendations* (e.g. homepage; registration); *Design for Communication Functionality* (i.e. functionality that supports communication between users such as activity updates, comments, likes, forums, links to social media etc); *Design for Data Collection*; *Design for Data Processing and Visualisation*; *Gamification Features*; *User Privacy Issues* (e.g. setting various levels of access in the collected data). In section 5 we present and discuss the guidelines in detail, as these were influenced by the CE.

Cooperative Evaluation (CE)

In the CE experiments, participants were given a set of tasks and explored features from the same six categories that we described in the previous section as discussed in section3.2. Here, we briefly present some general observations with respect to each one of the three applications. Additional findings that informed the guidelines are discussed more extensively in section 5 where we discuss the guidelines in more detail.

Table 2. Summary of design features and content identified by participants in each of the three applications examined with the method of CE.

	iSpot website	iNaturalist app	Zooniverse website
Most Useful/usable features	Information content; external links; terminology; data collection formats	Search by location; search autocomplete function; data collection formats	Projects page — categorisation tabs
Least useful/ least usable features	Filtering on main page; lack of mobile app for data collection; locating & filtering the map	Search function; help page and tutorials	News page hard to read; distracting colour and small font size;
Controversial amongst participants features	Social media login feature	—	Opening pages in new/same window'; use of back button
Suggestions	Forums and help pages separate menu items; main page's filtering should accommodate needs of global community; introduce 'expert status'	Homepage of the mobile app to contain links to news and external articles; Tab to check notifications; place search tab on top of the screen; application zooms to the new observation point added	Newsfeed to resemble that of Facebook's layout; 'talk' tab to be renamed to 'forum'; skip tutorial option

iSpot. Participants appreciated the fact that the main iSpot's homepage is informative and contains links to news articles and educational material. Nonetheless, participants suggested that the main page should not be organised by 'communities', neither it should show content based on user's 'location'. They recommended that end users should have control of such filtering options so that it is more open towards its global community. Logging in using social media was quite controversial (*"It's a nice feature… I don't need to remember another password"*; *"It depends on the permission settings, because I don't like having to change all the permission settings on Facebook"* user comments). Participants also liked features for entering data in various formats, and the concurrent use of both scientific and lay vocabulary to explain species identification. Problematic features included finding the interactive map, filtering the map results and the lack of a search by location function. CE participants were not interested in the badges award feature, although they expressed greater trust levels to data contributed by volunteers with badges. They suggested that an 'expert status' feature which takes into account skills and professional qualifications would be more effective in this respect.

iNaturalist.org. Participants found the 'search by location' function particularly useful; tutorials and help pages and the search function were hard to locate and participants suggested that these should be located on the top of the screen. Users were not interested in communication features (*"Everybody just uses Facebook messenger"* user comment) and gamification features. Participants suggested that a feature similar to that provided by Facebook to check notifications would improve the user experience, as well as external links to news or relevant articles from the mobile app's homepage. The fact there is no legend provided caused frustration. Participants further suggested that once a new data point is added the application should zoom in on this observation automatically.

Zooniverse. Participants appreciated external links and the news stories, yet they found the page hard to read and suggested it should be reorganised to resemble that used of Facebook. They disliked the orange colour on the background of the headings and the small font size. They did not immediately understand the purpose of the 'talk' feature, and when they did, they suggested it should be renamed to 'forum'. With respect to tutorials, participants suggested that pop-ups work best (*"Nobody has the time to watch a video"* participant comment) and that a skip option should be provided. Communication and gamification features were again not highly rated, yet one participant commented in his think aloud "what do I get with my points?", and suggested that such a gamification feature would only make sense if users are truly awarded with something in return (e.g. zoo vouchers).

Designing for citizen science: design guidelines

Basic features and design recommendations

General *interface design should follow popular name and navigation conventions* (e.g. 'forum' instead of 'talk'). Both CE experiments and the SLR meta-analysis [e.g. see Teacher et al., 2013; Kim, Mankoff and Paulos, 2013] agree that the project's main page should contain information about: project description; data collected; project outcomes; and links to news and external links for additional information. Providing a *news* section, according to articles reviewed, serves as a motivation incentive feature [Eveleigh et al., 2014] and its design should follow name conventions, as suggested by the majority of our participants and the SLR [e.g. McCarthy, 2011]. A *forum* should be provided as separate menu item to support volunteers collate and respond to feedback, offer their suggestions and support, as well as for social interaction purposes [Woods, McLeod and Ansine, 2015; Wald, Longo and Dobell, 2016]. Similarly, a *help page* should be provided as a separate menu item, a suggestion made by our participants, which is also in line with other guidelines [e.g. Skarlatidou, Cheng and Haklay, 2013].

Registration is a common feature that many studies discuss, with a growing number of applications providing the option to sign up using social media [e.g. Dunlap, Tang and Greenberg, 2013; Ellul, Francis and Haklay, 2011; Wald, Longo and Dobell, 2016]. Jay et al. [2016] demonstrated that *"it is possible to increase contributions to online citizen science by more than 60%, by allowing people to participate in a project without obliging them to officially sign up"* (p. 1830). Participants in our study expressed their support towards the social media login function. Nevertheless they were concerned with changes in permission settings in their social media accounts. We thus suggest that *sign up using social media should not be the only option and that a registration page is also provided*. Several of the articles in the SLR discuss that CS technologies should provide *tutorials in various forms* (e.g. textual, videos) [Yadav and Darlington, 2016; Dunlap, Tang and Greenberg, 2013; Stevens et al., 2014; Kim, Mankoff and Paulos, 2013; Prestopnik and Crowston, 2012, etc]. Our participants recommended that tutorials should be provided *using pop-up functionality* (which highlight relevant features) with the option *to skip them and/or allow to return at some later point in time*. Photographs in tutorials should be used to communicate scientific objects of interest [e.g. see Dunlap, Tang and Greenberg, 2013].

Overall interface design should also take into account *cultural and environmental characteristics*. There is, for example, a growing number of studies in CS which

investigate how such technologies should be designed to support data collection by illiterate or semi-literate people [Stevens et al., 2014; Liebenberg et al., 2017; Vitos et al., 2017; Idris et al., 2016; Jennett, Cognetti et al., 2016]. Studies from this context suggest early engagement with end users, placing the user at the centre of the design and development lifecycle, removing any form of text from user interfaces, and so on [Idris et al., 2016]. Also, several of these applications are designed to be used in remote areas and/or the outdoors (e.g. in the forest, parks or recreation areas). Therefore, the *design of buttons should be usually larger than in standard application design* and the *screen lighting* and *battery life* should be tested separately [Ferster et al., 2013; Stevens et al., 2014; Vitos et al., 2017].

Design for communication functionality

The SLR articles emphasise the importance of providing *(real-time) communication functionality* [e.g. see Newman, Wiggins et al., 2012; Wald, Longo and Dobell, 2016; Fails et al., 2014; Woods, McLeod and Ansine, 2015] to support communication amongst volunteers, or between volunteers and the scientists, to retain volunteers, and to communicate further information about how the data are used or will be used. The latter has been discussed within the context of trust-building, for example, its potential for increasing confidence in the project and corresponding scientists [Ferster et al., 2013]. *Map tagging* and *adding comments on map* have also been suggested, together with *chats and forums* [Dunlap, Tang and Greenberg, 2013]. CE participants showed no preference towards the inclusion of a chat function but they all thought that a forum page adds value. Nevertheless, it should be acknowledged that the artificial experimental setting of CE is perhaps not appropriate to fully appreciate the benefits in the same way that actual volunteers experience this feature. Furthermore, communication functionality depends on the nature of the project, its geographical extent, and other attributes, and we therefore suggest that these are all taken into account when designing for communication in CS applications.

Design for data collection

Data collection is perhaps the most widely discussed feature. The SLR articles provide suggestions for: *real-time data collection* [e.g. Panchariya et al., 2015]; supporting de-anonymisation *from sharing data with close friends and other groups* [e.g. Fails et al., 2014; Maisonneuve, Stevens, Niessen et al., 2009]; enabling the *collection of user's GPS location data* [e.g. Maisonneuve, Stevens and Ochab, 2010; Kim, Mankoff and Paulos, 2013], only if it is of high accuracy, otherwise users can become frustrated/confused [Bowser, Hansen, He et al., 2013]; the *ability to collect various data types such as numbers, videos, photographs, text, coordinates* [Ellul, Francis and Haklay, 2011]; the ability to *combine various data types derived from various sources*, which requires a good understanding of the types of data that are useful in the specific project [e.g. Wehn and Evers, 2014; Kim, Mankoff and Paulos, 2013]; and the ability *to add qualitative data to observations* [Maisonneuve, Stevens and Ochab, 2010], which is an attribute that it is very popular amongst the SLR articles. *Adding a photograph to showcase the observation captured can be a significant trust cue* (as noted previously, it may also support the development of data validation mechanisms) [e.g. see Kim, Mankoff and Paulos, 2013].

In terms of design, *form design should be simple to improve accessibility* [e.g. see Prestopnik and Crowston, 2012]; the use of *images and drop-down lists* can improve the time required to enter data and should be preferred over text [Idris et al., 2016], especially in mobile devices. When users *are not required to fill in all data fields make sure they understand this*, thus preventing them entering incorrect data [Woods and Scanlon, 2012]. Data collection supported by *showing the location of the user* (using GPS tracking) in the application (mapping interface), should *use symbology that stands out from base map* [e.g. Dunlap, Tang and Greenberg, 2013]. Moreover, the *provision of additional reference points shown on the map help users validate the data they collect* [Kim, Mankoff and Paulos, 2013]. When possible, *allow for data tailoring and personalisation to improve motivation* [Sullivan et al., 2009]. In some contexts, the submission of high quality images is absolutely essential (e.g. in species identification), however, the handsets that users are equipped with might not support zooming to the required level [Jennett, Cognetti et al., 2016]. In this case, we suggest that *users should be made aware of their handset limitations well in advance*, before they contribute any data.

In conditions where there is limited (or no) Internet access, scholars emphasise the importance of *collecting data offline, storing them and uploading them automatically once a connection is established* [Stevens et al., 2014; Fails et al., 2014; Bonacic, Neyem and Vasquez, 2015; Kim, Robson et al., 2011]. We also suggest that, within specific contexts of use, developers should consider *the benefit of providing effective data validation functionality* when new data are being submitted [e.g. averaging data records, flag errors and provide feedback to the user to correct errors, ask users to inspect data to throw out outliers/check data accuracy, view redundant data collected by others and decide whether it is helpful etc., as in Dunlap, Tang and Greenberg, 2013], as it can improve the quality of data and user trust in the data. A feature explicitly mentioned in the CE sessions by several participants is the feedback provision to notify volunteers when a new observation has been submitted.

Design for data processing and visualisation

Features relevant to data processing and visualisation include: a *data sharing and viewing website and mobile application to see the data collected instantly*, preferably on a map but other visualisations are also suggested, e.g. tables [Kim, Mankoff and Paulos, 2013; Woods, McLeod and Ansine, 2015; Maisonneuve, Stevens, Niessen et al., 2009]; *a search function with autocomplete capabilities* [Yadav and Darlington, 2016; Ellul, Francis and Haklay, 2011]; *search by location* [e.g. Sullivan et al., 2009; Elwood, 2009]; and *filter data on the map using different variables* [Kim, Mankoff and Paulos, 2013]. Other features include: *map zooming and panning tools* [Higgins et al., 2016; Kar, 2015]; the a*bility to switch between different map backgrounds*, especially in applications on mobile devices to save battery life and conserve on data consumption; and the ability to *read details of data collected while browsing the map interfaces*, e.g. using a hover text showing details for each data pin [e.g. in Kim, Mankoff and Paulos, 2013]. When observations on a map are linked to other information on other parts of the interface (e.g. textual information or graphs on the site) *provide visual cues for the user to understand association* [Fledderus, 2016].

Gamification features

Gamification is popular in CS literature for its potential to motivate users and increase numbers of contributions. Therefore several studies explore the use of badges, awards or leaderboards to incentivise users and keep them active [e.g. Prestopnik and Crowston, 2012; Newman, Wiggins et al., 2012; Crowley et al., 2012; Panchariya et al., 2015; Bowser, Hansen, He et al., 2013; Bowser, Hansen, Preece et al., 2014; Wald, Longo and Dobell, 2016]. Nevertheless, studies acknowledge that gamification might not be such a significant motivation factor in terms of collecting scientific data [e.g. one of the participants in Bowser, Hansen, Preece et al., 2014, mentions that "*contributing to science…that's kind of motivating to me*" p. 139] [Preece, 2016; Prestopnik, Crowston and Wang, 2017], which is in line with the results of our CE. Preliminary evidence suggests that leaderboards, rating systems and treasure hunting type of game features might improve user experience of CS applications [e.g. see Bowser, Hansen, Preece et al., 2014]. The option to *skip the use of gamification features* is also strongly recommended [Bowser, Hansen, Preece et al., 2014], so that volunteers do not have to deal with the competitive part of a gamified application if they don't want to [Preece, 2016].

CE participants suggested that gamification would not make them contribute more data and that an option to opt out is essential. They also suggest that gamification features would be more effective if they lead to tangible outcomes (i.e. non-diegetic reward features, such as sponsorships from local organisations for vouchers to participate in CS activities and events). This is congruent with the wider gamification literature, where point systems and leaderboards, although once addictive features, are being replaced by experience-based gamification features [McCarthy, 2011]. Other features such as badges or 'expert' status, may also help to tackle trust concerns in the same way that reputation systems, seals of approval, and other trust cues are discussed in other contexts such as online shopping, education and even Web GIS [Dellarocas, 2010; Skarlatidou, Cheng and Haklay, 2013].

User privacy issues

The articles of our meta-analysis make extensive reference to privacy issues and concerns [e.g. R. D. Phillips et al., 2014; Kim, Mankoff and Paulos, 2013; Preece, 2016]. *Filtering, moderation, ensuring that the data stored do not support identification of individuals and providing the option to collect data without sharing it* are some of the options that are suggested in the SLR as potential solutions [e.g. Ferster et al., 2013; Leao, Ong and Krezel, 2014; Maisonneuve, Stevens, Niessen et al., 2009]. Anonymity (via the option to contribute without registration/signature) has been also suggested to address privacy concerns [e.g. Dale, 2005] and improve participation. Nevertheless, it is suggested by existing studies that applications should be designed with attribution, which improves volunteers' trust, and therefore anonymity should not always be encouraged [e.g. Luther et al., 2009]. In the CE experiments only five participants expressed privacy concerns and they suggested that *complying with security standards and offering the ability to change privacy settings* would definitely be beneficial.

Conclusions and future work

The research carried out demonstrates the current trends and main design considerations, as suggested in user-related research and user testing findings, within the environmental digital CS context. It resulted in design guidelines which developers in CS can follow to improve their designs. Likewise, they can also be used to inform future research, for example, addressing methodological implications of the methods employed herein, as well as extending this work further.

The SLR was a time-consuming process given the extensive number of keywords, however it resulted in the identification of a satisfactory number of resources to include in the meta-analysis study. It should be noted that user research in the context of CS is in its infancy, and therefore many of the studies were making suggestions based on the authors' experiences rather than on hard evidence derived from the involvement of actual users. Nevertheless, we were also impressed by the number of studies which did involve users in evaluation, UCD practices, or simply in capturing user needs and requirements. We acknowledge the fact that a significant number of studies (i.e. n = 7,169) were removed due to language barriers. Although this sample was not further processed to assess its relevance it still may poses a limitation and therefore a future study that aims at investigating non-English research studies could significantly contribute to the proposed design guidelines for digital CS applications. Overall we suggest that wider communication of findings, preferably in more than one languages, will inspire more CS practitioners and researchers to involve end users in the design process, thus significantly improving the usability and user experience of CS applications. Moreover, the SLR described focuses only on the environmental context. Thus, our preliminary findings can be extended by further analysing similar studies in other contexts of digital CS.

The second limitation arises from the CE implementation, which involved 15 users. Although this sample provided us with enough insight into the proposed guidelines, increasing the number of applications tested and the number of users involved in CE could provide more in-depth insight to significantly improve the proposed guidelines. Introducing a set of recruitment criteria to inform user participation in any future studies may also extend the usefulness of our approach. For example, repeating the experiments with two groups of participants (i.e. users with some or significant experience in the use of CS applications and users who never used CS applications) can improve our understanding around user issues and corresponding design features that are important in terms of both attracting new users as well as retaining them. Similarly, participants of specific age groups, cultural backgrounds and especially those who speak languages other than English may help uncover more user issues which may be applicable to all or to specific design categories (e.g. gamification features and communication functionality) and contribute to the establishment of a more inclusive set of guidelines. Considering the current state of research in this area, we also suggest that any user testing focuses on additional user experience elements to understand not only usability or volunteer retainment, but also issues surrounding communication, privacy and improving trust via interface design.

Finally, the proposed guidelines can inform the design and development of environmental design CS applications and several of the guidelines are in line with similar work within the broader digital CS context [e.g. Sturm et al., 2018]. It should be however acknowledged that guidelines cannot replace contextual and

environmental considerations. We therefore propose that user studies are fully-integrated in citizen science, especially with respect to its technological implementations, in terms of understanding user issues that are relevant to specific contexts of use. The wider CS community can further benefit from the communication of user-based research in terms of not only improving our knowledge and understanding of what users want, but also in terms of providing the methodological protocols that others can easily replicate for exploring user issues in various contexts of use.

References   Boell, S. K. and Cecez-Kecmanovic, D. (2015). 'On being 'systematic' in literature reviews in IS'. *Journal of Information Technology* 30 (2), pp. 161–173. https://doi.org/10.1057/jit.2014.26.

Bonacic, C., Neyem, A. and Vasquez, A. (2015). 'Live ANDES: mobile-cloud shared workspace for citizen science and wildlife conservation'. In: 2015 IEEE 11th International Conference on e-Science (Munich, Germany, 31st August–4th September 2015), pp. 215–223. https://doi.org/10.1109/escience.2015.64.

Bowser, A., Hansen, D., He, Y., Boston, C., Reid, M., Gunnell, L. and Preece, J. (2013). 'Using gamification to inspire new citizen science volunteers'. In: *Proceedings of the First International Conference on Gameful Design (Gamification '13)*. ACM Press, pp. 18–25. https://doi.org/10.1145/2583008.2583011.

Bowser, A., Hansen, D., Preece, J., He, Y., Boston, C. and Hammock, J. (2014). 'Gamifying citizen science: a study of two user groups'. In: *Proceedings of the companion publication of the 17th ACM conference on Computer supported cooperative work & social computing — CSCW Companion '14* (Baltimore, MD, U.S.A. 15th–19th February 2014), pp. 137–140. https://doi.org/10.1145/2556420.2556502.

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T. and Boyle, J. M. (2012). 'A systematic literature review of empirical evidence on computer games and serious games'. *Computers & Education* 59 (2), pp. 661–686. https://doi.org/10.1016/j.compedu.2012.03.004.

Connors, J. P., Lei, S. and Kelly, M. (2012). 'Citizen science in the age of neogeography: utilizing volunteered geographic information for environmental monitoring'. *Annals of the Association of American Geographers* 102 (6), pp. 1267–1289. https://doi.org/10.1080/00045608.2011.627058.

Crowley, D. N., Breslin, J. G., Corcoran, P. and Young, K. (2012). 'Gamification of citizen sensing through mobile social reporting'. In: 2012 IEEE International Games Innovation Conference (Rochester, NY, U.S.A. 7th–9th September 2012). https://doi.org/10.1109/igic.2012.6329849.

D'Hondt, E., Stevens, M. and Jacobs, A. (2013). 'Participatory noise mapping works! An evaluation of participatory sensing as an alternative to standard techniques for environmental monitoring'. *Pervasive and Mobile Computing* 9 (5), pp. 681–694. https://doi.org/10.1016/j.pmcj.2012.09.002.

Dale, K. (2005). 'National Audubon society's technology initiatives for bird conservation: a summary of application development for the Christmas bird count'. In: Third international partners in flight conference '02 (Albany, CA, U.S.A. 20th–24th March 2005), pp. 1021–1024.

Dellarocas, C. (2010). 'Designing reputation systems for the social Web'. Boston U. School of Management Research Paper No. 2010-18. https://doi.org/10.2139/ssrn.1624697. (Visited on 3rd April 2018).

Dunlap, M., Tang, A. and Greenberg, S. (2013). 'Applying geocaching to mobile citizen science through science caching'. Technical Report 2013-1043-10. https://doi.org/10.5072/PRISM/30640.

Ellul, C., Francis, L. and Haklay, M. (2011). 'A flexible database-centric platform for citizen science data capture'. In: 2011 IEEE Seventh International Conference on e-Science Workshops (Stockholm, Sweden, 5th–8th December 2011). https://doi.org/10.1109/esciencew.2011.15.

Elwood, S. (2009). 'Geographic information science: new geovisualization technologies — emerging questions and linkages with GIScience research'. *Progress in Human Geography* 33 (2), pp. 256–263. https://doi.org/10.1177/0309132508094076.

Eveleigh, A. M. M., Jennett, C., Blandford, A., Brohan, P. and Cox, A. L. (2014). 'Designing for dabblers and deterring drop-outs in citizen science'. In: *Proceedings of the IGCHI Conference on Human Factors in Computing Systems (CHI '14)*. New York, NY, U.S.A.: ACM Press, pp. 2985–2994. https://doi.org/10.1145/2556288.2557262.

Fails, J. A., Herbert, K. G., Hill, E., Loeschorn, C., Kordecki, S., Dymko, D., DeStefano, A. and Christian, Z. (2014). 'GeoTagger: a collaborative participatory environmental inquiry system'. In: *Proceedings of the companion publication of the 17th ACM conference on Computer supported cooperative work & social computing — CSCW Companion '14* (Baltimore, MD, U.S.A. 15th–19th February 2014). https://doi.org/10.1145/2556420.2556481.

Ferreira, N., Lins, L., Fink, D., Kelling, S., Wood, C., Freire, J. and Silva, C. (2011). 'BirdVis: visualizing and understanding bird populations'. *IEEE Transactions on Visualization and Computer Graphics* 17 (12), pp. 2374–2383. https://doi.org/10.1109/tvcg.2011.176.

Ferster, C., Coops, N., Harshaw, H., Kozak, R. and Meitner, M. (2013). 'An exploratory assessment of a smartphone application for public participation in forest fuels measurement in the wildland-urban interface'. *Forests* 4 (4), pp. 1199–1219. https://doi.org/10.3390/f4041199.

Fledderus, T. (2016). 'Creating a usable web GIS for non-expert users: identifying usability guidelines and implementing these in design'. MSc dissertation. Uppsala, Sweden: University of Uppsala. URL: http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-297832.

Haddaway, N. R. (2015). 'The use of web-scraping software in searching for grey literature'. *International Journal on Grey Literature* 11 (3), pp. 186–190.

Higgins, C. I., Williams, J., Leibovici, D. G., Simonis, I., Davis, M. J., Muldoon, C., van Genuchten, P., O'Hare, G. and Wiemann, S. (2016). 'Citizen OBservatory WEB (COBWEB): a generic infrastructure platform to facilitate the collection of citizen science data for environmental monitoring'. *International Journal of Spatial Data Infrastructures Research* 11, pp. 20–48. https://doi.org/10.2902/1725-0463.2016.11.art3.

Idris, N. H., Osman, M. J., Kanniah, K. D., Idris, N. H. and Ishak, M. H. I. (2016). 'Engaging indigenous people as geo-crowdsourcing sensors for ecotourism mapping via mobile data collection: a case study of the Royal Belum State Park'. *Cartography and Geographic Information Science* 44 (2), pp. 113–127. https://doi.org/10.1080/15230406.2016.1195285.

Jay, C., Dunne, R., Gelsthorpe, D. and Vigo, M. (2016). 'To sign up, or not to sign up? Maximizing citizen science contribution rates through optional registration'. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems — CHI '16* (San Jose, CA, U.S.A. 7th–12th May 2016), pp. 1827–1832. https://doi.org/10.1145/2858036.2858319.

Jennett, C., Cognetti, E., Summerfield, J. and Haklay, M. (2016). 'Usability and interaction dimensions of participatory noise and ecological monitoring'. In: Participatory sensing. Opinions and collective awareness. Ed. by V. Loreto, M. Haklay, A. Hotho, V. D. P. Servedio, G. Stumme, J. Theunis and F. Tria. Cham, Switzerland: Springer, pp. 201–212. https://doi.org/10.1007/978-3-319-25658-0_10.

Jennett, C. and Cox, A. L. (2014). 'Eight guidelines for designing virtual citizen science projects'. In: *HCOMP '14* (Pittsburgh, PA, U.S.A. 2nd–4th November 2014), pp. 16–17. URL: https://www.aaai.org/ocs/index.php/HCOMP/HCOMP14/paper/view/9261.

Johnson, P. A., Corbett, J. M., Gore, C., Robinson, P., Allen, P. and Sieber, R. A. (2015). 'Web of expectations: evolving relationships in community participatory geoweb projects'. *International Journal for Critical Geographies* 14 (3), pp. 827–848. URL: https://acme-journal.org/index.php/acme/article/view/1235.

Kar, B. (2015). 'Citizen science in risk communication in the era of ICT'. *Concurrency and Computation: Practice and Experience* 28 (7), pp. 2005–2013. https://doi.org/10.1002/cpe.3705.

Kim, S., Mankoff, J. and Paulos, E. (2013). 'Sensr: evaluating a flexible framework for authoring data-collection tools for citizen science'. In: *Proceedings of the 2013 conference on Computer supported cooperative work — CSCW '13* (San Antonio, TX, U.S.A. 23rd–27th February 2013), pp. 1453–1462. https://doi.org/10.1145/2441776.2441940.

Kim, S., Robson, C., Zimmerman, T., Pierce, J. and Haber, E. M. (2011). 'Creek watch: pairing usefulness and usability for successful citizen science'. In: *Proceedings of the 2011 annual conference on Human factors in computing systems — CHI '11* (Vancouver, BC, Canada, 7th–12th May 2011), pp. 2125–2134. https://doi.org/10.1145/1978942.1979251.

Kobori, H., Dickinson, J. L., Washitani, I., Sakurai, R., Amano, T., Komatsu, N., Kitamura, W., Takagawa, S., Koyama, K., Ogawara, T. and Miller-Rushing, A. J. (2016). 'Citizen science: a new approach to advance ecology, education and conservation'. *Ecological Research* 31 (1), pp. 1–19. https://doi.org/10.1007/s11284-015-1314-y.

Kosmala, M., Crall, A., Cheng, R., Hufkens, K., Henderson, S. and Richardson, A. (2016). 'Season spotter: using citizen science to validate and scale plant phenology from near-surface remote sensing'. *Remote Sensing* 8 (9), p. 726. https://doi.org/10.3390/rs8090726.

Kullenberg, C. and Kasperowski, D. (2016). 'What is Citizen Science? A Scientometric Meta-Analysis'. *Plos One* 11 (1), e0147152. https://doi.org/10.1371/journal.pone.0147152.

Leao, S., Ong, K.-L. and Krezel, A. (2014). '2Loud?: community mapping of exposure to traffic noise with mobile phones'. *Environmental Monitoring and Assessment* 186 (10), pp. 6193–6206. https://doi.org/10.1007/s10661-014-3848-9.

Liebenberg, L., Steventon, J., Brahman, !, Benadie, K., Minye, J., Langwane, H. (and Xhukwe, Q. ((2017). 'Smartphone icon user interface design for non-literate trackers and its implications for an inclusive citizen science'. *Biological Conservation* 208, pp. 155–162. https://doi.org/10.1016/j.biocon.2016.04.033.

Luther, K., Counts, S., Stecher, K. B., Hoff, A. and Johns, P. (2009). 'Pathfinder: an online collaboration environment for citizen scientists'. In: *Proceedings of the 27th international conference on Human factors in computing systems — CHI 09* (Boston, MA, U.S.A. 4th–9th April 2009). https://doi.org/10.1145/1518701.1518741.

Maisonneuve, N., Stevens, M., Niessen, M. E., Hanappe, P. and Steels, L. (2009). 'Citizen noise pollution monitoring'. In: *Proceedings of the 10th Annual International Conference on Digital Government Research: Social Networks: Making Connections between Citizens, Data and Government* (Puebla, Mexico, 17th–21st May 2009), pp. 96–103.

Maisonneuve, N., Stevens, M. and Ochab, B. (2010). 'Participatory noise pollution monitoring using mobile phones'. *Information Polity* 15 (1,2), pp. 51–71. https://doi.org/10.3233/IP-2010-0200.

McCarthy, J. (2011). 'Bridging the gaps between HCI and social media'. *interactions* 18 (2), pp. 15–18. https://doi.org/10.1145/1925820.1925825.

Michener, W. K., Allard, S., Budden, A., Cook, R. B., Douglass, K., Frame, M., Kelling, S., Koskela, R., Tenopir, C. and Vieglais, D. A. (2012). 'Participatory design of DataONE — enabling cyberinfrastructure for the biological and environmental sciences'. *Ecological Informatics* 11, pp. 5–15. https://doi.org/10.1016/j.ecoinf.2011.08.007.

Monk, A., Wright, P., Haber, J. and Davenport, L. (1993). Improving your human-computer interface: a practical technique. U.S.A.: Prentice-Hall.

Newman, G., Wiggins, A., Crall, A., Graham, E., Newman, S. and Crowston, K. (2012). 'The future of citizen science: emerging technologies and shifting paradigms'. *Frontiers in Ecology and the Environment* 10 (6), pp. 298–304. https://doi.org/10.1890/110294.

Newman, G., Zimmerman, D., Crall, A., Laituri, M., Graham, J. and Stapel, L. (2010). 'User-friendly web mapping: lessons from a citizen science website'. *International Journal of Geographical Information Science* 24 (12), pp. 1851–1869. https://doi.org/10.1080/13658816.2010.490532.

Panchariya, N. S., DeStefano, A. J., Nimbagal, V., Ragupathy, R., Yavuz, S., Herbert, K. G., Hill, E. and Fails, J. A. (2015). 'Current developments in big data and sustainability sciences in mobile citizen science applications'. In: *2015 IEEE First International Conference on Big Data Computing Service and Applications* (Redwood City, CA, U.S.A. 30th March–2nd April 2015), pp. 202–212. https://doi.org/10.1109/bigdataservice.2015.64.

Phillips, R. D., Blum, J. M., Brown, M. A. and Baurley, S. L. (2014). 'Testing a grassroots citizen science venture using open design, "the bee lab project"'. In: *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems — CHI EA '14* (Toronto, Canada, 26th April–1st May 2014), pp. 1951–1956. https://doi.org/10.1145/2559206.2581134.

Phillips, R., Ford, Y., Sadler, K., Silve, S. and Baurley, S. (2013). 'Open design: non-professional user-designers creating products for citizen science: a case study of beekeepers'. In: Design, user experience and usability. Web, mobile and product design. DUXU 2013. Lecture Notes in Computer Science, 8015. Ed. by A. Marcus. Berlin, Germany: Springer, pp. 424–431. https://doi.org/10.1007/978-3-642-39253-5_47.

Preece, J. (2016). 'Citizen science: new research challenges for human-computer interaction'. *International Journal of Human-Computer Interaction* 32 (8), pp. 585–612. https://doi.org/10.1080/10447318.2016.1194153.

Prestopnik, N. R. and Crowston, K. (2012). 'Citizen science system assemblages: understanding the technologies that support crowdsourced science'. In: *Proceedings of the 2012 iConference on — iConference '12* (Toronto, ON, Canada, 7th–10th February 2012). https://doi.org/10.1145/2132176.2132198.

Prestopnik, N. R., Crowston, K. and Wang, J. (2017). 'Gamers, citizen scientists and data: exploring participant contributions in two games with a purpose'. *Computers in Human Behavior* 68, pp. 254–268. https://doi.org/10.1016/j.chb.2016.11.035.

Prestopnik, N. R. and Tang, J. (2015). 'Points, stories, worlds and diegesis: comparing player experiences in two citizen science games'. *Computers in Human Behavior* 52, pp. 492–506. https://doi.org/10.1016/j.chb.2015.05.051.

Queirós, A., Silva, A., Alvarelhão, J., Rocha, N. P. and Teixeira, A. (2015). 'Usability, accessibility and ambient-assisted living: a systematic literature review'. *Universal Access in the Information Society* 14 (1), pp. 57–66. https://doi.org/10.1007/s10209-013-0328-x.

Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., Lewis, D. and Jacobs, D. (2012). 'Dynamic changes in motivation in collaborative citizen-science projects'. In: *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW 2012)* (Seattle, WA, U.S.A. 11th–15th February 2012). ACM Press, pp. 217–226. https://doi.org/10.1145/2145204.2145238.

Salah, D., Paige, R. F. and Cairns, P. (2014). 'A systematic literature review for agile development processes and user centred design integration'. In: *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering — EASE '14* (London, U.K. 12th–14th May 2014). https://doi.org/10.1145/2601248.2601276.

Sharma, N. (2016). 'Species identification in citizen science'. In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems — CHI EA '16* (San Jose, CA, U.S.A. 7th–12th May 2016), pp. 128–133. https://doi.org/10.1145/2851581.2890382.

Sharples, M., Aristeidou, M., Villasclaras-Fernández, E., Herodotou, C. and Scanlon, E. (2015). 'Sense-it: a smartphone toolkit for citizen inquiry learning'. In: The mobile learning voyage — from small ripples to massive open waters. Communications in Computer and Information Science. Ed. by T. H. Brown and H. J. van der Merwe, pp. 366–377. https://doi.org/10.1007/978-3-319-25684-9_27.

Silvertown, J. (2009). 'A new dawn for citizen science'. *Trends in ecology and evolution* 24 (9), pp. 467–471. https://doi.org/10.1016/j.tree.2009.03.017.

Skarlatidou, A., Cheng, T. and Haklay, M. (2013). 'Guidelines for trust interface design for public engagement Web GIS'. *International Journal of Geographical Information Science* 27 (8), pp. 1668–1687. https://doi.org/10.1080/13658816.2013.766336.

Sprinks, J., Wardlaw, J., Houghton, R., Bamford, S. and Morley, J. (2017). 'Task workflow design and its impact on performance and volunteers' subjective preference in Virtual Citizen Science'. *International Journal of Human-Computer Studies* 104, pp. 50–63. https://doi.org/10.1016/j.ijhcs.2017.03.003.

Stevens, M., Vitos, M., Altenbuchner, J., Conquest, G., Lewis, J. and Haklay, M. (2014). 'Taking participatory citizen science to extremes'. *IEEE Pervasive Computing* 13 (2), pp. 20–29. https://doi.org/10.1109/mprv.2014.37.

Sturm, U., Schade, S., Ceccaroni, L., Gold, M., Kyba, C., Claramunt, B., Haklay, M., Kasperowski, D., Albert, A., Piera, J., Brier, J., Kullenberg, C. and Luna, S. (2018). 'Defining principles for mobile apps and platforms development in citizen science'. *Research Ideas and Outcomes* 4, e23394. https://doi.org/10.3897/rio.4.e23394. (Visited on 4th July 2018).

Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D. and Kelling, S. (2009). 'eBird: A citizen-based bird observation network in the biological sciences'. *Biological Conservation* 142 (10), pp. 2282–2292. https://doi.org/10.1016/j.biocon.2009.05.006.

Teacher, A. G. F., Griffiths, D. J., Hodgson, D. J. and Inger, R. (2013). 'Smartphones in ecology and evolution: a guide for the app-rehensive'. *Ecology and Evolution* 3 (16), pp. 5268–5278. https://doi.org/10.1002/ece3.888.

Theobald, E. J., Ettinger, A. K., Burgess, H. K., DeBey, L. B., Schmidt, N. R., Froehlich, H. E., Wagner, C., HilleRisLambers, J., Tewksbury, J., Harsch, M. A. and Parrish, J. K. (2015). 'Global change and local solutions: Tapping the unrealized potential of citizen science for biodiversity research'. *Biological Conservation* 181, pp. 236–244. https://doi.org/10.1016/j.biocon.2014.10.021.

Vitos, M., Altenbuchner, J., Stevens, M., Conquest, G., Lewis, J. and Haklay, M. (2017). 'Supporting collaboration with non-literate forest communities in the Congo-basin'. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing — CSCW '17* (Portland, OR, U.S.A. 27th February–1st March 2017), pp. 1576–1590. https://doi.org/10.1145/2998181.2998242.

Wald, D. M., Longo, J. and Dobell, A. R. (2016). 'Design principles for engaging and retaining virtual citizen scientists'. *Conservation Biology* 30 (3), pp. 562–570. https://doi.org/10.1111/cobi.12627.

Wehn, U. and Evers, J. (2014). 'Citizen observatories of water: social innovation via eParticipation'. In: *Proceedings of the 2014 conference ICT for Sustainability* (Stockholm, Sweden, 24th–28th August 2014). https://doi.org/10.2991/ict4s-14.2014.1.

Wiggins, A. (2013). 'Free as in puppies: compensating for ICT constraints in citizen science'. In: *Proceedings of the 2013 conference on Computer supported cooperative work — CSCW '13* (San Antonio, TX, U.S.A. 23rd–27th February 2013), pp. 1469–1480. https://doi.org/10.1145/2441776.2441942.

Wiggins, A. and Crowston, K. (2011). 'From Conservation to Crowdsourcing: a Typology of Citizen Science'. In: *Proceedings of the 44th Hawaii International Conference on System Sciences (HICSS-44)*. Kauai, HI, U.S.A. Pp. 1–10. https://doi.org/10.1109/HICSS.2011.207.

Woods, W., McLeod, K. and Ansine, J. (2015). Supporting mobile learning and citizen science through iSpot. Ed. by H. Crompton and J. Traxler. Abingdon, U.K.: Routledge, pp. 69–86.

Woods, W. and Scanlon, E. (2012). 'iSpot mobile — a natural history participatory science application'. In: *Proceedings of Mlearn 2012* (Helsinki, Finland, 15th–16th October 2012).

Yadav, P. and Darlington, J. (2016). 'Design guidelines for the user-centred collaborative citizen science platforms'. *Human Computation* 3 (1), pp. 205–211. https://doi.org/10.15346/hc.v3i1.11. (Visited on 3rd April 2018).

Zapata, B. C., Fernández-Alemán, J. L., Idri, A. and Toval, A. (2015). 'Empirical studies on usability of mHealth apps: a systematic literature review'. *Journal of Medical Systems* 39 (1), pp. 1–19. https://doi.org/10.1007/s10916-014-0182-2.

Authors

Artemis Skarlatidou is a postdoctoral researcher in the Extreme Citizen Science (ExCites) group at UCL. She is chairing WG4 of the COST Action on citizen science CA15212 and co-chairing the ICA Commission on Use, User and Usability Issues. Her research interests include HCI and UX aspects (e.g. usability, aesthetics, trust) of geospatial technologies and especially of Web GIS representations for expert and public use and citizen science applications. She is also interested in trust issues within the context of VGI and PPGIS; philosophical, as well as, ethical issues for the 'appropriate' and effective use of geospatial and citizen science technologies. E-mail: a.skarlatidou@ucl.ac.uk.

Born and raised in the United States, Alexandra Hamilton moved to Europe with her family at the age of 14. After completing school in Brussels, Belgium she then obtained her BSc in Geography and Education Studies from Oxford Brookes University in 2016. Shortly after, in 2017, she was awarded her MSc in Geospatial Analysis from the University College London. Interested in the application of geospatial analytics within the social sciences, Alexandra hopes to commence her Ph.D. in the fall of 2019. E-mail: alexandra.hamilton.16@ucl.ac.uk.

Michalis Vitos is a postdoctoral researcher in the Extreme Citizen Science (ExCites) group at UCL. His research interests include the areas of Human-Computer Interaction and Software Engineering in combination with the emerging area of Participatory GIS and Citizen Science. E-mail: michalis.vitos@ucl.ac.uk.

Muki Haklay is a professor of GIScience and the director of the Extreme Citizen Science (ExCites) group at UCL. His research interests include public access to environmental information and the way in which the information is used by a wide range of stakeholders, citizen science and in particular applications that involve community-led investigation, development and use of participatory GIS and mapping, and Human-Computer Interaction (HCI) for geospatial technologies. E-mail: m.haklay@ucl.ac.uk.