

Community-based Data Validation Practices in Citizen Science

Andrea Wiggins

College of Information Studies
University of Maryland
College Park, MD, USA
wiggins@umd.edu

Yurong He

College of Information Studies
University of Maryland
College Park, MD, USA
yrhe@umd.edu

ABSTRACT

Technology-supported citizen science has created huge volumes of data with increasing potential to facilitate scientific progress, however, verifying data quality is still a substantial hurdle due to the limitations of existing data quality mechanisms. In this study, we adopted a mixed methods approach to investigate community-based data validation practices and the characteristics of records of wildlife species observations that affected the outcomes of collaborative data quality management in an online community where people record what they see in the nature. The findings describe the processes that both relied upon and added to information provenance through information stewardship behaviors, which led to improved reliability and informativity. The likelihood of community-based validation interactions were predicted by several factors, including the types of organisms observed and whether the data were submitted from a mobile device. We conclude with implications for technology design, citizen science practices, and research.

Author Keywords

Citizen science; information assessability; data quality; data validation; open collaboration

ACM Classification Keywords

H.5.3. Information Interfaces and Presentation (e.g. HCI): Group and Organization Interfaces.

INTRODUCTION

CSCW has studied collaborations between scientists and volunteers who contribute scientific data [10, 31, 38, 46], which is a common form of citizen science. In citizen science, the usual tasks for which volunteers are enrolled include collecting, categorizing, transcribing, or analyzing scientific data [2]. Supported by the rapid development of social computing, recent successes in citizen science have created growing data sets with increasing impact on scientific progress in several domains [6, 36] and new applications are regularly emerging.

However, much of the scientific community remains skeptical of data quality [2].

Validation is critical to ensuring the usefulness of citizen science data by establishing its quality. For example, in biology and ecology, without validation of species identifications, volunteers' observations are usually argued to have limited value [40]. Given the challenges of scaling and successes of crowdsourcing, it is reasonable to ask if a community can effectively validate the data they create.

We therefore investigated how data quality in a citizen science project can be improved through open collaboration. This study focused on understanding the nature of the interactions around data quality within an existing online community that can be adopted by citizen science projects as a data collection platform and identifying the characteristics of the contributed data that influenced how community members assessed data quality. We applied and extended the concept of information assessability, defined as “the extent to which an interface helps people make appropriate assessments of information quality in a particular context of use” [11] for statistical analyses, with interpretation grounded in participant observation, to address the research questions:

RQ1: How can online community members collaboratively improve the quality of data generated by volunteers?

RQ2: What factors influence the likelihood of data quality improvement through online community interactions?

The study site was an existing online community, iNaturalist, a social network site that supports natural history data sharing for scientific, education, and personal use¹. As of May 2015, the iNaturalist community included over 70,000 users who contributed over 1,300,000 species occurrence records, each documenting the presence of an organism at a particular time and place. This type of data is used extensively in ecology research and land management decision making.

We adopted a mixed methods approach to study this site, starting with participant observation in a place-based project that utilized iNaturalist for data management. This experience grounded quantitative analyses of the data and interactions on the iNaturalist platform, which were downloaded directly from the site.

Permission to make digital or hard copies of part or all of this work is granted without fee provided that copies bear this notice and the full citation on the first page. This work is made available by the authors under a CC-BY 4.0 International license. CSCW'16, February 27–March 02, 2016, San Francisco, CA, USA. ACM 978-1-4503-3592-8/16/02.

¹<http://www.inaturalist.org/pages/about>

We answered the first research question by identifying and describing the types of interactions that led to improvements in data quality. The results of statistical analyses to answer the second research question showed that multiple characteristics of the records influenced the level of data quality improvement. In addition to expanding on the concepts of information assessability and contributing a description of collaborative data quality management practices in citizen science, we identified implications for design, practice, and research related to community-based data validation.

RELATED WORK

This section provides a brief overview of related research on citizen science data quality, much of which is found in the domain literature. We then further develop the concept of information assessability [11] by adding the concepts of informativity and reliability, which we used in our analyses.

Citizen Science Data Quality

The dominant form of citizen science is monitoring wildlife or other environmental phenomena [47]. Relatively anonymous observers with varied abilities, experiences, and trainings contribute spatiotemporally unique data: by nature, observational data are ephemeral. These data may contain an unknown number of errors, occasionally including intentionally falsified data [8, 12]. Anecdotally, project leaders for several of the largest citizen science projects currently in operation—those most likely to attract the attention of trolls or anti-science saboteurs—reported blacklisting a tiny fraction of users for these reasons, on the order of one in every 10,000 users. Prior findings show that rates of deception vary substantially across contexts [33], and the low apparent rate of malfeasance in citizen science may be an unexpected benefit of the design of data entry forms, built-in participant privacy protections and lack of resources for more sophisticated social features (c.f. [46]), which create limited means for a troll to be “rewarded” with attention [7].

Since citizen science projects are typically purpose-driven with diverse goals [47], a pragmatic definition of data quality used widely in this context is “fitness for intended use” [23]. This definition acknowledges the importance of contextual expectations for quality in scientific data. However, although the standards of data quality may vary on a project-by-project basis, we note that domain consensus on a general level is reflected in the accepted minimum information standards for documenting and sharing certain types of data. For example, in biodiversity, the minimum requirements for a species occurrence record (the type of data in this study) include taxonomic identification, data collection time and location [15], among which the identity of the organism is often the primary focus in data quality review.

Furthermore, data quality in citizen science is not solely an attribute of data (e.g., “fitness for intended use” [23]), but also a process of maintaining data quality [39]. In light of the variability of criteria for data quality in citizen science and the dual nature of the concept, which describes both attributes and processes, we consider data quality to be “the processes and outcomes of evaluating and improving the utility of data.”

Data Quality Validation

So far the most effective methods for evaluating citizen science data combine machine and human filters. Advanced statistical and computing techniques can help reduce human attention needed for data quality management [19] and flag potentially erroneous or fake data [3, 43]. But not all issues can be resolved by machine. Even with a photo or recording as a voucher, the “truth” of an observation can be elusive. Many projects therefore rely most heavily on expert review. Experts may be project scientists, trained project staff, or experienced volunteers [3, 24].

Quality control via expert review focuses on validation, a process through which a third party evaluates the data and determines whether it is acceptable [43]. This is essentially a consensus process: the ephemerality of event-based data means that evaluating data points is a judgment of the probability that the record is reliable, e.g., the location and time have been reported with appropriate precision, and the species is accurately identified. The main difference between this process and traditional science is that traditional science relies more on physical vouchers (e.g., DNA) for precise identification and verification. However, this mode of validation is not logistically feasible for citizen science; instead, similar processes use digital artifacts as vouchers (e.g., photo, audio or video recording) [3, 43].

When reviewing data quality, experts are asked to judge the probability that a given record is reliable; this is how the truth of observational data is established [45]. In citizen science projects, review process outcomes are not usually publicly visible. One constraint of an expert-driven approach is that for most species occurrence data, the number of experts with local knowledge is necessarily limited and unevenly distributed, and citizen science projects sometimes generate more data than experts can review.

Another matter of concern is social influence bias: when multiple parties collectively evaluate data, the visibility of the outcomes of prior parties’ decisions can have variable impacts on data quality rating [30]. Compared with crowdsourcing tasks that are highly simplified and require little domain expertise (e.g., classifying a single feature of an image into one of five categories), evaluating species occurrence records requires sorting an entity into one of thousands of categories based on the intersection of multiple physical features, behaviors, and spatial and temporal variables. For such a task, it is less clear whether “blind” reviewing would yield better results; more transparent collaborative processes in which a series of contributors’ actions build on prior efforts, as in Wikipedia or open source software, may be more effective.

The complex task of validating a species occurrence record is amenable to layered contributions, e.g., one person may identify an organism’s proper phylum, and given that cue, the next person can determine the correct family, and so on [20]. The fact that such a layered process is both practical and functional suggests that involving the full community in data validation is an underexplored avenue for enriching both the data and the participants’ experience. Making the review process transparent and inviting review from the entire com-

munity can also serve a number of learning and socialization functions [1, 22, 31]. As such, community-based data validation can resemble heavyweight peer production, with members that not only contribute content, but also “share a commitment to maintaining and sustaining the direction and viability of the community” [17]. This would also move projects closer to a member-maintained community model [7], currently unusual in citizen science.

Currently, relatively few projects enlist volunteers in evaluating the content of other participants’ contributions. If moving toward a member-maintained community model in citizen science can effectively support data review (e.g., iSpot in [41]), then community-based data validation may offer a complementary data quality strategy to help allocate expert attention to the data that most need it.

Given the conservatism of scientific communities with respect to data quality and the variability of volunteers’ expertise, addressing the citizen science data quality challenge currently requires tracking the processes of data creation, validation, and modification, and then openly sharing the process documentation (provenance) alongside the data so that potential data consumers can assess the fit to their needs.

Assessing Data Quality

Making citizen science data open access has potential to help support evaluation of the science and legitimization of the project [5]. For citizen science data to be viable, both the datum itself and the process by which it has been maintained should be readily assessable by would-be consumers. This is a nontrivial challenge in citizen science, since the potential data consumers include academic, conservation, governance, and hobbyist groups [26].

Information provenance and information stewardship are related concepts that describe important features of assessable designs for participatory information sources, such as Wikipedia [11]. Information provenance describes the content’s origination and the processes that have been performed upon it, while information stewardship refers to how the content has been maintained, often including considerations around whom was involved. In the example of Wikipedia, Forte et al. found that visualizing provenance and stewardship can significantly influence assessments of articles and of Wikipedia itself as an information source [11].

For similarly participatory information sources in the form of citizen science data, we expect provenance and stewardship to be at least as important for assessability. Indeed, the standards of scientific knowledge production routinely require such information, as reified in the methods section of the academic paper.

Rather than studying *perceptions* of provenance and stewardship, in this study we used the information assessability concepts to identify evidence of *practices and structures* supporting effective information provenance and stewardship in an online community supported by a social computing platform. This study therefore also contributes an extension of the concepts of information assessability, information prove-

nance, and information stewardship, which are described further in the next section.

STUDY SITE

iNaturalist is an online social network where naturalists of all types can record what they see in nature, meet other nature lovers, and learn about the natural world. In this section, we describe the primary platform features, key user roles, and processes by which participants generate and share data about living organisms, and then maintain and improve the data collaboratively.

Platform Features

iNaturalist can be considered a social computing platform, and along with the community of iNaturalist users, forms an information ecosystem around nature observation data that researchers can use for studying biodiversity and anyone can use for learning about nature.

iNaturalist acts as a data ingestion and validation tool for contributing species occurrence records to a growing biodiversity data set used for research and conservation, among other applications. With many of the typical features of social network sites, iNaturalist user interactions reflect complex social processes, as in other online communities.

iNaturalist users can adopt multiple roles. They can be *observers* who observe and create records, *project managers* who create and manage projects, *validators* who help to identify observed organisms, and *data consumers*, who use the data for their own purposes and may or may not have any other relationship to iNaturalist or its users. The observers, validators, and data consumers were the primary actors of interest in relation to our research questions. To clarify how iNaturalist users enacted these roles, we describe the primary participation processes.

Information Provenance: Generating and sharing data

In the context of iNaturalist, it is important to note that provenance refers both to the information about the observation (e.g., location, time/date, observer, etc.) but also information about *changes* to the record of the observation through interactions with the community. iNaturalist maintains exceptional transparency by maintaining a visible record of the process of data quality refinement observation records, as well as summary statistics on agreement, providing a valuable reference point for potential data users.

Provenance is inherently entwined in data creation and sharing, as observers can use any tools they like (e.g., camera, audio recorder, smartphone, paper and pen, etc.) to record organism information. Data can be uploaded with either the iNaturalist website on a personal computer or the Android or iOS mobile apps. Both device types let users record detailed observations and attach a media file such as a photograph. Some of the specific details (location, time, date) are automatically captured by mobile apps unless the user overrides the defaults. Details on other fields in the records are discussed in the methods section. Upon upload, a new record page is generated for the observation.

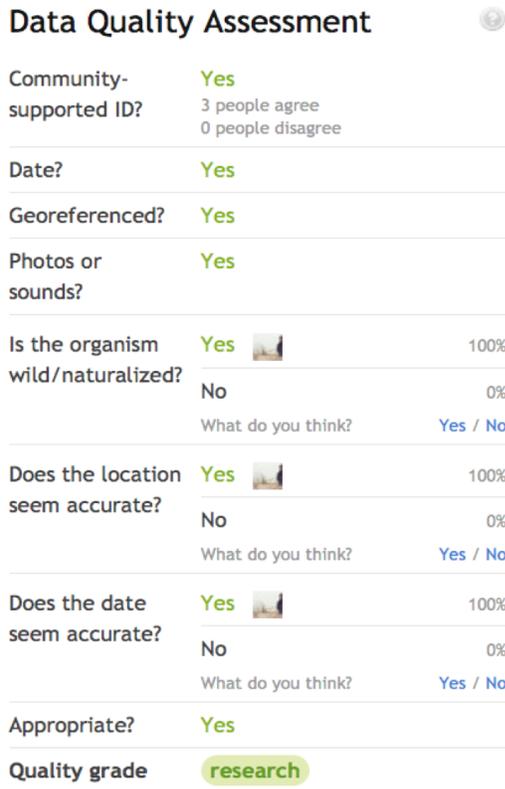


Figure 1. Summary of agreements on an iNaturalist record; screenshot made in May, 2015.

Information Stewardship: Maintaining and improving data

After an observer uploads data to iNaturalist, the entire iNaturalist user community can assist with validation. In general, there are two types of interactions: refinements of the organism identification (hereafter referred to as ID) and agreements. Agreements can be positive or negative (disagreements) and can be made on the organism ID and/or other pre-specified details. iNaturalist users can confirm the current ID by indicating agreement, can suggest a better ID, and can indicate agreement with the accuracy of the record metadata, as well as whether or not the organism is wild (see Figure 1). The owner of the record can also reject IDs from other users. Agreements or disagreements with the metadata are instantly updated on the record page.

The history of agreements with and refinements to the ID, such as who made what changes or agreed with which ID, are visible on each record in a threaded history that combines taxonomic updates with user comments. However, this provenance about stewardship behaviors is only recorded when adding new information to the existing information; if a user “removes” a prior ID agreement or suggestion, all traces of the interaction were removed. The species ID with the highest consensus is displayed on the record as authoritative; according to iNaturalist’s FAQs,

Our algorithm for deciding the community ID is a little convoluted—but in general, we try to choose the taxon

that more than 2/3 of identifiers agree with, where agreeing means adding ID of a taxon matching or within another taxon. So if Scott says it’s a mammal and Ken-ichi says it’s a mammal, then the community ID is Mammalia. If Scott says it’s a mammal and 5 other people say it’s a Golden Eagle, then the community ID is Golden Eagle.

In addition, to be considered research grade, the record must include a date, accurate longitude and latitude coordinates, at least one photo or sound file, and as shown in Figure 1, no disagreements from other iNaturalist users indicating that the location does not look plausible (e.g., sharks in the middle of desert) or the organism is not wild or has been naturalized (e.g., is a pet). These criteria outline the acceptable minimum standards for data to be included in several biodiversity data repositories such as the Global Biodiversity Information Facility (GBIF).

For the data to shift from casual to research grade requires the attention and efforts of both the initial observers and the wider community of iNaturalist users. The platform supports collaborative information stewardship interactions by providing mechanisms that blend socializing with data verification, and displaying them on the records as a form of provenance. This provenance provides two common indicators of data quality: *informativity* and *reliability*.

Records that reach research grade status contain more information content than casual grade records. The more specific the taxon level of the ID, the more informative the record becomes, since taxonomic hierarchies operate on inheritance principles. Similarly, the more iNaturalist users confirm the ID, the more reliable the ID is considered, in keeping with standard scientific practices more broadly. We evaluate information assessability through these two concepts, informativity and reliability, which link provenance functionality and stewardship behaviors with supporting evidence of these collaborative data validation practices and structures.

METHODS

We adopted a sequential mixed-methods approach in this study [44] to investigate the role of provenance and stewardship in community-based validation. This section describes our data collection, analysis procedures, and details of the dependent and independent variables.

Data Collection

The data for this study came from two sources: participant observation in a citizen science project, Biocubes², that used iNaturalist as its data management platform, and data that were submitted to iNaturalist on the date of our participant observation.

Participant observation data

We participated in a training event in Florida on January 24th, 2015 for educators interested in the Biocubes project. Ethnographic field notes, memos, and photos documented the process of data collection, including choosing a sampling site,

²<http://www.inaturalist.org/projects/biocubes>

inventorying its inhabitants, photographing the organisms, identifying them as well as possible, uploading data to iNaturalist, and returning the organisms to their habitat.

The primary participants in the three day event were ten middle school and high school science teachers in a Title I school district in Florida, plus four informal science educators working at aquaria and nature centers. The research team included twelve facilitators: two social scientists, three educators, five biologists, and two professional photographers. The research team captured approximately 1300 photos of the participation processes, providing detailed reference information for reconstructing activities. One of the authors also presented an introductory tutorial to participants on using iNaturalist, and both authors helped participants troubleshoot the iNaturalist mobile apps on Android and iOS devices.

We observed that most of the smartphone-savvy educators and scientists experienced a substantial learning curve with the mobile apps. For example, some participants accidentally submitted incomplete or duplicate records while they were learning to use the app. In addition, some participants chose not to upload any data rather than using an unfamiliar app. These observations were critical to identifying how the iNaturalist system recorded provenance and to interpreting the findings. As will be discussed later, device type ended up being an important variable, which would have gone unnoticed if we had not observed the teachers struggling to submit data with the mobile apps. The analyses also benefitted from identifying the system-based sources of differences between the data we retrieved and what was expected based on field observation [21].

iNaturalist data

For this exploratory investigation, we selected publicly accessible data that were recorded on iNaturalist in the U.S. on a single day, with a data export of species occurrence records from iNaturalist using the site's data download tool on March 11, 2015. The site allows users to download data retrieved through custom queries on set fields. We initially downloaded all the information about the records, including basic information (e.g., ID, date, time, description, etc.), location information (e.g., latitude, longitude, positional accuracy, etc.), taxon information (e.g., scientific and common names, taxonomic level, etc.).

We limited our analysis to the U.S. on January 24, 2015 due to our participation in the event in Florida on that date, which provided access to the minutiae of the on-the-ground processes that a portion of the data represented. 925 records were made for observations in the U.S. on that date and uploaded to iNaturalist within the arbitrary 45-day interval after the observation date. For comparison purposes, this allowed us to control for temporal effects from the amount of time elapsed between dates of observation and data retrieval, as well as potential influences of specific temporal windows (e.g., delayed interactions on records submitted on major holidays) [21] and we had no reason to believe that the specific date would skew the results for this analysis. The sample proved large enough for statistical significance in quantitative analyses, while be-

ing amenable to comprehensive manual examination for qualitative analysis of interactions on records.

Data cleaning and preparation

Our statistical analyses focused on identifying the characteristics of records that predicted data validation interactions. After obtaining the data as described, we identified potential influential factors and dependent variables to indicate reliability and degree of informativity. We chose not to examine data that were unrelated to the research questions, such as usernames and time zones. Some data were unsuitable for analysis due to extensive missing data, lack of precision, or lack of documentation.

Of the 61 columns in the data set we downloaded, three dependent variables were selected to represent the concepts of informativity and reliability. Quality grade (one column) and number of agreements (one column) respectively provided system-centric and user-centric measures of reliability, while taxon level (seven columns collapsed into one factor) represented informativity from both perspectives.

The following description of these variables also encapsulates a partial answer to the first research question by describing the data that resulted from the interactions through which data quality was established within the iNaturalist context, using the iNaturalist measures of data quality. For the purposes of our analyses, these three variables measure data quality and represent the outcomes of the data validation processes.

1. **Quality grade:** This dummy variable indicates whether iNaturalist classifies the record as “research” grade (1) or “casual” grade (0) by criteria described earlier.
2. **Number of agreements:** The more (registered) iNaturalist users agree with a record's ID, the more reliable the ID is considered.
3. **Taxon level:** The data includes the taxonomic rank of the community ID (i.e., organism's name on the page) at the time the data set was retrieved. The more specific the taxon level is, the more informative the record is. Seven ranks are built into the iNaturalist taxonomic hierarchy: kingdom, phylum, class, order, family, genus, and species. Initial descriptive statistics showed many more records at genus and species levels than other taxonomic levels. Three groups of taxonomic ranks were coded as genus or species level (2); class, order, or family level (1); and kingdom, phylum, or without taxon level information (0).

We then identified eight independent variables among the remaining 52 columns that we expected would influence the records' reliability and informativity as measured by the dependent variables.

1. **Quality grade:** Although quality grade status is also included as a dependent variable, we considered it a potentially influential factor for taxon level and number of agreements because research grade records must include more contextual details. We were specifically interested in how records with full information to achieve research grade status compared to the casual records.

2. **Device:** This dummy variable indicates whether users chose mobile devices with the iNaturalist apps (1) or the website on personal computers (0) to upload data. Among the 925 records we retrieved, the majority of data that were uploaded to iNaturalist from an app were from the official iNaturalist iPhone and Android apps. Two records were uploaded from different citizen science apps designed for other initiatives that using iNaturalist for data management (i.e., Great Nature Project, Texas Nature Trackers).
3. **License:** This dummy variable indicates whether users chose a Creative Commons (CC) license (1) or retained copyright with “all rights reserved” (0).
4. **Description:** This dummy variable indicates whether users did (1) or did not (0) add a textual description as supplemental information others might use in refining IDs.
5. **ID please:** This dummy variable indicates whether the “ID please” flag was (1) or was not (0) active at the time we retrieved the records. The “ID please” flag is visible on the record and used in built-in search queries that direct attention to these records. While users who initially turned on the “ID please” flag can later turn it off, and vice versa, data downloaded from iNaturalist does not include the change history of the “ID please” flag. We downloaded daily snapshots of data over a period of two weeks and verified that such changes were made on only about 0.6% of all records, which would not affect our analyses.
6. **Positional accuracy:** This dummy variable indicates the range of the position accuracy chosen by users: greater than ten meters (1), or between one and ten meters (0). iNaturalist allows users to adjust the range of the position accuracy manually to protect their privacy or address other needs³. There was wide variation in positional accuracy, with radius sizes of one meter to over 25 kilometers, but about 25% of records had positional accuracy between one and ten meters. The remaining 75% were unevenly distributed. Most records with greater positional accuracy were uploaded from mobile apps.
7. **Positioning device:** This dummy variable indicates whether users chose to manually adjust the position (1) or used the default position information (0). The data record four types of positioning devices: manual, GPS, Google, and blank. Except manual, which means that the user changed the radius of the location indicator or dragged a pin to a different location from its initial landing point, the other three are default position settings with variable positional accuracy. The position is generated via GPS for smartphones using location services, by Google when a user searches for a named location, and blank when a user drops a pin but makes no other changes to refine the position or radius of the location indicator.
8. **Taxon group:** iNaturalist’s “taxon groups” are a local convention that combine several taxonomic branches to make it easier to find specific groups of organisms. Taxon group categories are functional shortcuts to specific targets within

the top three levels of the taxonomic hierarchy (kingdom, phylum, and class). These categories are: Animals, Plants, Fungi including Lichens, Kelp, Diatoms and Allies, and Protozoans (kingdom level); Mollusks (phylum level); and Reptiles, Amphibians, Mammals, Ray-finned fishes, Birds, Insects, Arachnids, and Plants (class level). In the data, records of Birds and Plants dominated. Birds were 39% of records, plants made up 32%, and each of the other taxon groups individually accounted for approximately 1% to 6.8% of the total. This variable was coded as bird (2), plant (1), or neither-bird-nor-plant (0).

Data Analysis

Qualitative analysis methods

We matched observation data retrieved from iNaturalist with instances in our field notes and photos representing the context and behaviors that resulted in those specific data points in order to verify our assumptions about what the downloaded data represented. We found several quirks by comparing exported records, observation pages, and field data; this allowed us to understand the extent of provenance recorded by the system and begin to examine how user behavior in the field may have impacted the validation of records they submitted. We also performed simple content analysis on the comments for over 400 records.

Quantitative analysis methods

Quality grade, taxon level, and number of agreements were all distributed non-normally, so we adopted generalized linear models (GZLM) in SPSS version 22 to analyze relationships between dependent variables and the record metadata [16].

We chose a binary logistic model (Model 1, see Table 1) for the quality grade dependent variable because its response is binary, a negative binomial with log link model (Model 2) for the number of agreements because it is a count variable, and an ordinal logistic model (Model 3) for the taxon level dependent variable because its response is ordinal. Overall model fit was assessed by a Chi-square test for each. Exponentiated coefficients are odds ratios for Model 1 and 3, and incident rates for Model 2, for a standardized effect size [16, 35].

Among the variables with significant results, we were surprised to observe that the type of device used to upload data had the most substantial impact on the research grade dependent variable. We then conducted additional statistical comparisons between devices and all other variables with Chi-square analyses (see Table 2). We transformed number of agreements into binary variables for Chi-square tests.

FINDINGS

In this section, we discuss our findings, beginning with a description of community-based data validation practices on iNaturalist. Next, we identify which metadata made records more or less likely to receive agreements and refinements from the community. Finally, we highlight an unexpected finding related to the impact of technology preferences on data quality.

³Vulnerable species’ locations are obscured automatically.

Variables	Model 1 Quality grade			Model 2 Number of agreements			Model 3 Taxon level		
	b	SE	Exp(b)	b	SE	Exp(b)	b	SE	Exp(b)
Intercept/ Threshold	0.59	0.42	1.80	-3.78	0.49	.02***	-0.79	0.55	0.45
Quality grade							0.16	0.55	1.18
				4.24	0.35	69.53***	1.84	0.29	6.30***
Device									
				-0.87	0.20	0.42***	-0.41	0.17	0.67*
							-0.78	0.29	0.46**
Licence									
				-0.22	0.34	0.81	0.38	0.27	1.47
Description									
				-0.25	0.16	0.78	-0.02	0.11	0.98
							0.37	0.27	1.44
ID please									
				-0.67	0.23	0.51**	0.17	0.22	1.18
							-2.66	0.27	0.07***
Positional accuracy									
				0.09	0.21	1.09	-0.21	0.16	0.81
							0.87	0.28	2.39**
Positioning device									
				-0.07	0.21	0.93	-0.20	0.13	0.82
							-0.54	0.42	0.58
Taxon group									
				1.73	0.21	5.61***	0.41	0.14	1.51**
				-0.28	0.18	0.75	-0.06	0.16	0.94
							4.25	1.03	70.18***
							1.81	0.30	6.08
χ^2	227.182***			674.984***			378.970***		

Table 1. Generalized linear models using binary logistic for Research grade, ordinal logistic for Taxon level, and negative binomial distribution and log link for Number of agreements. For Taxon level, the reference level of the threshold is genus and species level. Note * $p < .05$, ** $p < .01$, * $p < .001$.**

Community-based Data Validation Practices

We have partially described the processes through which iNaturalist users contributed to community-based data validation in delineating the interactions that the system permits, which goes toward answering the first research question, *How can online community members collaboratively improve the quality of data generated by volunteers?* The specific mechanisms through which data quality was improved in iNaturalist were discussed, with two primary types of interactions recorded in provenance to provide evidence of community-based data validation: agreements and refinements of organism IDs. These data-centric outcomes were constrained by the mechanisms available to contributors, but were also the product of social processes.

Collaborative data validation was inherently social because data quality improvements were displayed as visible history and involved discussion among users. We directly engaged the iNaturalist community after our participant observation at the Florida event, and prior to designing this study (our interactions carried no explicit research intent, but were essentially normal user interactions). Our records were, like others, actively curated by other iNaturalist users who helped refine the organism IDs. Like other users, we were personally gratified by others' attention to our records, which represented much more than data to us [18, 27]. We provided further details to inquiring users, whose questions conveyed a form of social validation for effort invested in creating data and helped us learn more about the organisms we encountered.

We performed high-level content analysis on the comments for over 400 records. The provenance included not only the

records of actions taken toward data quality improvement, but also evidence of socialization [7]. These included comments that highlighted ongoing social relationships between users, aesthetic reactions to the images (e.g., pretty butterflies), and sharing professional knowledge. Although these interactions did not always appear task-driven, they influenced the data validation processes on a more subtle level. For example, social interaction between users might also influence the amount of attention given to the records submitted by an individual, which could be a topic for future study.

Predictors of Community-based Validation

In this section, we answer the second research question: *What factors influence the likelihood of data quality improvement through online community interactions?* Or in other words, what characteristics attracted the attention of the community? We expected that more attention would go to records with full details for “who, what, when, and where”.

Table 1 presents the results from three regression models. In the bottom panel of Table 1, the likelihood ratio test of χ^2 for all three models indicates that they are more appropriate models compared to corresponding models without any factors (a “null” model) respectively.

Model 1: Quality Grade

Model 1 tests which factors made some records more likely to become research grade. About 26% of the data were submitted via mobile app (Table 2), but using a mobile device to upload data and using the “ID please” flag were negatively associated with research grade status. Records were more

Variables		Mobile (N = 244)		PC (N = 681)	
		n	Percentage	n	Percentage
Quality grade	Research	94	38.5%	473	69.5%
	Casual	150	61.5%	208	30.5%
Taxon level	Genus and species	175	71.7%	624	91.6%
	Class, order, and family	23	9.4%	24	3.5%
	Kingdom, phylum, and other	46	18.9%	33	4.8%
Number of agreements	At least one agreement	95	38.9%	477	70.0%
	No agreement	149	61.0%	204	30.0%
License	CC	212	86.9%	658	96.6%
	All right reserved	32	13.1%	23	3.4%
Description	With description	78	32.0%	211	31.0%
	Without description	166	68.0%	470	69.0%
ID please	With flag	45	18.4%	73	10.7%
	Without flag	199	81.6%	608	89.3%
Positional accuracy	>10m	117	48.0%	620	91.0%
	≤10m	127	52.0%	61	9.0%
Positioning device	Manual	0	0.0%	179	26.2%
	Default	244	100%	502	73.7%
Taxon group	Bird	38	15.6%	323	47.4%
	Plant	101	41.4%	194	28.5%
	Neither	105	43.0%	164	24.1%

Table 2. The number and percentage of records for different devices and variables.

likely to reach research grade if they were submitted from a PC and/or did not flag “ID please”.

If the records were in the bird taxon group, they were more likely to achieve research grade level compared to neither-bird-nor-plant records. Although plant records did not show a statistically significant difference from neither-bird-nor-plant records, the direction of their predicted relationships suggested that plants were less likely to reach research grade. One possible explanation is that these taxon groups have species ID search spaces that are different by orders of magnitude. According to taxonomic research [29], there are about 8.7 million distinct species on Earth, including approximately 10,000 species of birds, almost 300,000 species of plants, and 8.4 million species that are neither birds nor plants.

Model 2: Number of Agreements

Model 2 tested which factors were significantly associated with more agreements. Quality grade status and taxon group predicted higher numbers of agreements, while using a mobile device to upload data predicted fewer agreements. Research grade records and bird data were more likely to have attracted the attention of more users to verify IDs.

Model 3: Taxon Level

Model 3 tested which factors were significantly associated with records that had genus and species level IDs, our proxy for high informativity. Quality grade status, positional accuracy, and the taxon group of the records predicted more specific levels of ID. If the records had research grade status, had reasonably believable geographic information, and/or were bird data, they were more likely to have genus or species level IDs. Taxon group was particularly influential for birds, for which the odds of the ID reaching genus and species level were over 70 times greater than other records ($b = 4.25$, $p <$

.001). However, data uploaded by mobile app or with the “ID please” flag had genus or species level IDs less often.

Impacts of Technology Use on Data Validation

In our analyses, the device type was a significant factor in all three models; we investigated device type because we observed participants experiencing difficulty using the mobile apps during the event in Florida. We now turn to findings of particular interest for CSCW given recent work on apps for citizen science [4, 14, 25, 37], the implications for data quality of the users’ choices of technologies.

Chi-square tests showed every variable except descriptions were strongly related to device type ($p < .01$). Table 2 shows the number and percentage of records for different variables according to device type. There was a strong relationship between device type and likelihood of achieving research grade status: data uploaded from smartphones were less likely to become research grade. Taxon group was also related to device type, since birds were more likely to be uploaded via website than mobile apps. Bird data—which had very high agreement and were usually identified to the species level—were also more likely to be research grade. Based on these results, we expect device type may mask a combination of more subtle variables.

DISCUSSION

The implications of these results point to considerations for technology design given the impacts of device type on the outcomes of community-based data validation. We consider the practical implications for project leaders and data consumers, as well as the CSCW research opportunities this study revealed.

Implications for Design

We saw in our analyses that bird records were more likely to be uploaded from a PC than a mobile app—but why? We believe this is in part because bird photography often requires an optical zoom lens and faster shutter speeds than smartphones offer. In the field, we saw that using a mobile device to photograph any motile organism well enough for another person to verify species ID is challenging, even if the organism (e.g., an insect) is captive (see Figure 2).

By comparison, recording bird data takes substantially longer with the all-taxon iNaturalist app than with a task-specific, taxon-specific app, such as the eBird app (formerly BirdLog), which is tailored specifically to bird data and highly streamlined for minimal interference with actual observation activities. Further, post hoc data entry using the mobile app attaches precise but inaccurate timestamp and location information unless the user is unusually attentive.

In contrast, smartphones were used more often for plants, which are easy to approach and photograph, and may be more suited to this type of app. Although these examples focus on birds and plants, the findings suggest that the physical features and behaviors of the organisms in different branches of the taxonomic hierarchy may have a much more substantial influence on how users record information than is assumed by designers and data users.

By extension, the delay between data collection in the field and data entry using a personal computer rather than a mobile device could potentially support observers in providing more informative data. In the interim, the species ID may become more apparent or a field guide could be consulted, and the data may be more refined by the point of submission. In contrast, submitting data in the field rarely allows for graceful error recovery or more systematic review after leaving the field and its distractions; users must exert substantial additional effort to retrospectively edit and improve their own data, which is not usually supported in the apps that permit data submission. For users who are only participating because the app makes it easy to do so, the odds of returning to refine their own data seem low.

This is a stereotypical tradeoff in citizen science project design, pitting data quality against engagement: the smartphone app encourages casual participation that yields casual quality data that is unlikely to see much improvement, but could engage more new audiences (however briefly) for wider overall impact on public awareness. Enforcing separate processes for data collection and data submission permits additional processing by the observer to occur in the interim and requires additional effort that usually reduces follow through, but the data may be of greater scientific value. This becomes a trade-off scenario because the core task is fundamentally complex; if species ID were simplified to the level of the Citizen Sort games [34], we would expect no difference in data quality based on the device or taxon group.

We also suspect that users were more likely to deploy the mobile apps specifically for help with ID when encountering an unfamiliar organism, leading to more data submitted by

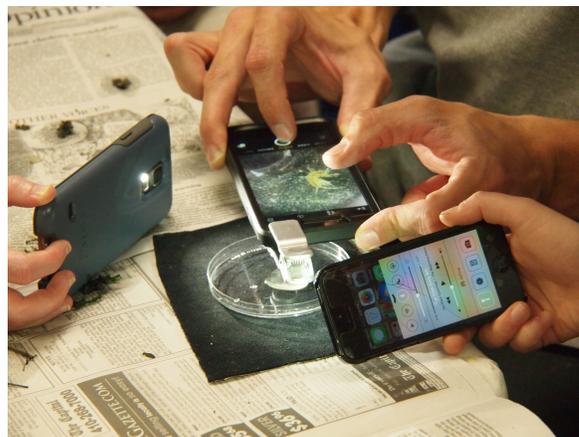


Figure 2. Teachers at a Biocubes training event used three smartphones with a clip-on macro lens to photograph an amphipod. It took several minutes to make a usable photo.

smartphone at a more general taxonomic level than genus or species. Our analyses showed that such data were relatively less likely to be validated. Further research at the individual level would help answer the question of which devices users choose under which circumstances.

These findings bring us to question the appropriateness of current technology trends for the actual community practices [40], as there is currently strong demand for citizen science mobile apps, which are perceived as a promising means of expanding participation [4, 42]. Mobile devices can capture temporal and geographic information to submit automatically and precisely [14, 28, 32]. Requiring zero user effort to accurately document time and location should hypothetically make it more likely for such data to reach research grade status. Data submitted from mobile devices, however, could not always satisfy the other requirements. It seems that for some uses, task-technology fit is not as good as anticipated [13].

The fact that using a PC to upload data was positively associated with community validation interactions could also reflect preferences for devices to access data and assess data quality. Although the iNaturalist apps were quite complex and did support activities like browsing, agreeing, and commenting on records (none of which were explored by the event participants), the overwhelming emphasis of the apps was on submitting data, and tasks like examining records' details, refining IDs, and downloading data may be easier to complete on a PC. A direction for future work on mobile apps (particularly for citizen science) could focus on new ways to access, assess, and validate data with mobile devices, which might be more feasible through a “family of apps” strategy rather than “one app to do all the things”, as app complexity may have contributed to confusion among new users. Since conducting the analyses in this paper, the iNaturalist interfaces and mobile apps were updated; in future work, we plan to examine whether these design updates on iNaturalist change the outcomes of community data validation from the current findings.

Implications for Practice

Several practitioner groups may be interested in the results of this study. For project leaders considering using iNaturalist, the analyses provide decision information: if focusing on a specific taxon or recommending a device type to upload data, these particulars could impact the likelihood of the community helping with data verification. We believe the data quality impact of these seemingly innocuous decisions would come as a surprise to most project leaders, who may wish to adjust instructions to volunteers in light of these results.

In addition, encouraging low quality data that will garner little attention is clearly counterproductive from a data collection standpoint, and may also have mixed impact from an engagement perspective. While a shiny mobile app may help recruit more short-term participants who would not otherwise participate at all [9], it could also ultimately undermine retention if the lack of attention to their contributions discourages ongoing participation. Research to date on motivations suggests this is a genuine risk [38].

For the iNaturalist community and others like it, our findings also underscore the need for data documentation to support appropriate data use and meaningful interpretations. The records submitted by volunteers were of variable initial quality and many were subsequently improved. iNaturalist's records of the data validation process and provenance were much more transparent than the average scientific data set, making it feasible for us (and other potential data users) to evaluate the quality of the records for ourselves, without direct assistance from those who created and curated them. These findings also suggest that citizen science data export files should include metadata about the technologies used to capture and submit data, including software version details. This would let data consumers determine whether specific technologies are affecting the outputs in a meaningful way.

For those interested in using data from iNaturalist or platforms like it, we recommend monitoring the effects of the topic or content and the technologies in use. iNaturalist data about birds may be as reliable and informative as other expert-reviewed citizen science data in use by researchers and policy makers. However, data on plants were nearly as plentiful as data about birds, but far less likely to receive as much attention, so the reliability and informativity of those data were not as well documented. This points to an area of opportunity for both practice and research: what strategies can elicit validation for other taxon groups besides birds?

Implications for Research

This study introduced and explored community-based data validation in citizen science, providing a foundation for future work comparing collaborative data validation on iNaturalist against similar activities on other citizen science platforms, such as iSpot, which could offer insight for designing the next generation of citizen science engagement. The utility of community-based validation should also be established through comparison to expert review, which would help identify the conditions under which each strategy performs best. Further, comparison of the validation processes in

this community with those in other peer production communities would support a richer dialogue about quality in open collaboration systems.

We also applied the concept of information assessability to guide our analysis, adding two concepts to measure assessability. While reliability and informativity may not be the correct metrics for all open collaboration communities, developing a conceptual framework around information assessability will support cross-context research on peer production with potential to drive theoretical advancements.

Additional implications for research are in part a reflection of the limitations of the study: scaling up these analyses to verify the results and detect new patterns across time and national boundaries are obvious next steps. Access to detailed transaction data could also provide more insights about the users, social structures, and interactions in this citizen science peer production community.

CONCLUSION

This study examined practices and structures supporting community-based data validation in the iNaturalist community. We described the processes through which community-based data validation was accomplished in this community, addressing the first research question. We examined how community interactions contributed to the informativity and reliability of the data, which both relied upon and added to information provenance through information stewardship behaviors. We identified which details made some records more likely to be validated, answering the second research question on the predictors of collaborative validation of data quality.

The findings highlighted the impact of the device type on subsequent data improvement, or lack thereof. The choice of device may mask numerous other behavioral variations but had a clear impact on the outcomes of community-based data validation. These findings prompted us to question how well the current technology trends in citizen science meet participants' needs versus projects' needs, and how well mobile apps are supporting the complex interactions in these communities.

ACKNOWLEDGMENTS

This work was supported in part by NSF Grant #1442668. We sincerely thank the hundreds of iNaturalist members who contributed the data used in these analyses. We also thank Dr. Carrie Seltzer, Dr. Jen Hammock, and Dr. Jennifer Preece for editorial feedback.

REFERENCES

1. Tammy D. Allen, Stacy E. McManus, and Joyce E.A. Russell. 1999. Newcomer socialization and stress: Formal peer relationships as a source of support. *Journal of Vocational Behavior* 54, 3 (1999), 453–470. DOI : <http://dx.doi.org/10.1006/jvbe.1998.1674>
2. Rick Bonney, Jennifer L. Shirk, Tina B. Phillips, Andrea Wiggins, Heidi L. Ballard, Abraham J. Miller-Rushing, and Julia K. Parrish. 2014. Next steps for citizen science. *Science* 343, 6178 (2014), 1436–1437. DOI : <http://dx.doi.org/10.1126/science.1251554>

3. David N. Bonter and Caren B. Cooper. 2012. Data validation in citizen science: a case study from Project FeederWatch. *Frontiers in Ecology and the Environment* 10, 6 (2012), 305–307. DOI : <http://dx.doi.org/10.1890/110273>
4. Anne E. Bowser, Derek L. Hansen, Jocelyn Raphael, Matthew Reid, Ryan J. Gamett, Yurong He, Dana Rotman, and Jenny J. Preece. 2013. Prototyping in PLACE: A Scalable Approach to Developing Location-based Apps and Games. In *Proc. CHI*. ACM, 1519–1528. DOI : <http://dx.doi.org/10.1145/2470654.2466202>
5. Anne E. Bowser and Andrea Wiggins. 2015. Privacy in Participatory Research: Advancing Policy to support Human Computation. *Human Computation* 2, 1 (2015), 19–44. DOI : <http://dx.doi.org/10.15346/hc.v2i1.3>
6. Caren B. Cooper, Jennifer Shirk, and Benjamin Zuckerberg. 2014. The Invisible Prevalence of Citizen Science in Global Research: Migratory Birds and Climate Change. *PLoS ONE* 9, 9 (09 2014), e106508. DOI : <http://dx.doi.org/10.1371/journal.pone.0106508>
7. Dan Cosley, Dan Frankowski, Sara Kiesler, Loren Terveen, and John Riedl. 2005. How Oversight Improves Member-maintained Communities. In *Proc. CHI*. ACM, 11–20. DOI : <http://dx.doi.org/10.1145/1054972.1054975>
8. Janis L. Dickinson, Benjamin Zuckerberg, and David N. Bonter. 2010. Citizen Science as an Ecological Research Tool: Challenges and Benefits. *Annual Review of Ecology, Evolution, and Systematics* 41, 1 (2010), 149–172. DOI : <http://dx.doi.org/10.1146/annurev-ecolsys-102209-144636>
9. Alexandra Eveleigh, Charlene Jennett, Ann Blandford, Philip Brohan, and Anna L. Cox. 2014. Designing for Dabblers and Deterring Drop-outs in Citizen Science. In *Proc. CHI*. ACM, 2985–2994. DOI : <http://dx.doi.org/10.1145/2556288.2557262>
10. Jerry Alan Fails, Katherine G. Herbert, Emily Hill, Christopher Loeschorn, Spencer Kordecki, David Dymko, Andrew DeStefano, and Zill Christian. 2014. GeoTagger: A Collaborative and Participatory Environmental Inquiry System. In *CSCW Companion*. ACM, 157–160. DOI : <http://dx.doi.org/10.1145/2556420.2556481>
11. Andrea Forte, Nazanin Andalibi, Thomas Park, and Heather Willever-Farr. 2014. Designing Information Savvy Societies: An Introduction to Assessability. In *Proc. CHI*. ACM, 2471–2480. DOI : <http://dx.doi.org/10.1145/2556288.2557072>
12. Judy Foster-Smith and Stewart M. Evans. 2003. The value of marine ecological data collected by volunteers. *Biological Conservation* 113, 2 (2003), 199 – 213. DOI :
13. Dale L. Goodhue and Ronald L. Thompson. 1995. Task-Technology Fit and Individual Performance. *MIS Quarterly* 19, 2 (1995), 213–236. DOI : <http://dx.doi.org/10.2307/249689>
14. Eric A. Graham, Sandra Henderson, and Annette Schloss. 2011. Using mobile phones to engage citizen scientists in research. *Eos, Transactions American Geophysical Union* 92, 38 (2011), 313–315. DOI : <http://dx.doi.org/10.1029/2011EO380002>
15. Robert P. Guralnick, Andrew W. Hill, and Meredith Lane. 2007. Towards a collaborative, global infrastructure for biodiversity assessment. *Ecology Letters* 10, 8 (2007), 663–672. DOI : <http://dx.doi.org/10.1111/j.1461-0248.2007.01063.x>
16. James William Hardin, Joseph M Hilbe, and Joseph Hilbe. 2007. *Generalized linear models and extensions*. Stata Press.
17. Caroline Haythornthwaite. 2009. Crowds and Communities: Light and Heavyweight Models of Peer Production. In *HICSS*. IEEE, 1–10. DOI : <http://dx.doi.org/10.1109/HICSS.2009.137>
18. Benjamin K. Haywood. 2015. Beyond Data Points and Research Contributions: The Personal Meaning and Value Associated with Public Participation in Scientific Research. *International Journal of Science Education, Part B* 0, 0 (2015), 1–24. DOI : <http://dx.doi.org/10.1080/21548455.2015.1043659>
19. Wesley M. Hochachka, Daniel Fink, Rebecca A. Hutchinson, Daniel Sheldon, Weng-Keen Wong, and Steve Kelling. 2012. Data-intensive science applied to broad-scale citizen science. *Trends in Ecology & Evolution* 27, 2 (2012), 130 – 137. DOI : <http://dx.doi.org/10.1016/j.tree.2011.11.006>
20. James Howison and Kevin Crowston. 2014. Collaboration through open superposition: A theory of the open source way. *Mis Quarterly* 38, 1 (2014), 29–50.
21. James Howison, Kevin Crowston, and Andrea Wiggins. 2011. Validity issues in the use of social network analysis with digital trace data. *Journal of the Association for Information Systems* 12 (2011).
22. Justin P. Johnson. 2006. Collaboration, peer review and open source software. *Information Economics and Policy* 18, 4 (2006), 477 – 497. DOI : <http://dx.doi.org/10.1016/j.infoecopol.2006.07.001>
23. Joseph M. Juran and A. Balton Godfrey. 1998. *Quality handbook* (5th ed.). McGraw-Hill, USA.
24. Steve Kelling, Carl Lagoze, Weng Weng-Keen, Jun Yu, Theodoros Damoulas, Jeff Gerbracht, Daniel Fink, and Carla Gomes. 2013. eBird: A Human / Computer Learning Network to Improve Biodiversity Conservation and Research. *AI Magazine* 34, 1 (2013), 10 – 20.

25. Sunyoung Kim, Jennifer Mankoff, and Eric Paulos. 2013. Sensr: Evaluating a Flexible Framework for Authoring Mobile Data-collection Tools for Citizen Science. In *Proc CSCW*. ACM, 1453–1462. DOI : <http://dx.doi.org/10.1145/2441776.2441940>
26. Carl Lagoze. 2014. eBird: Curating Citizen Science Data for Use by Diverse Communities. *International Journal of Digital Curation* 9, 1 (2014), 71–82.
27. Anna Lawrence. 2006. ‘No Personal Motive?’ Volunteers, Biodiversity, and the False Dichotomies of Participation. *Ethics, Place & Environment* 9, 3 (2006), 279–298. DOI : <http://dx.doi.org/10.1080/13668790600893319>
28. Ko Ko Lwin and Yuji Murayama. 2011. Web-based GIS System for real-time field data collection using personal mobile phone. *Journal of Geographic Information System* 3, 04 (2011), 382.
29. Camilo Mora, Derek P. Tittensor, Sina Adl, Alastair G. B. Simpson, and Boris Worm. 2011. How Many Species Are There on Earth and in the Ocean? *PLoS Biol* 9, 8 (2011), e1001127. DOI : <http://dx.doi.org/10.1371/journal.pbio.1001127>
30. Lev Muchnik, Sinan Aral, and Sean J Taylor. 2013. Social influence bias: A randomized experiment. *Science* 341, 6146 (2013), 647–651.
31. Gabriel Mugar, Carsten Østerlund, Katie DeVries Hassman, Kevin Crowston, and Corey Brian Jackson. 2014. Planet Hunters and Seafloor Explorers: Legitimate Peripheral Participation Through Practice Proxies in Online Citizen Science. In *Proc CSCW*. ACM, 109–119. DOI : <http://dx.doi.org/10.1145/2531602.2531721>
32. Greg Newman, Andrea Wiggins, Alycia Crall, Eric Graham, Sarah Newman, and Kevin Crowston. 2012. The future of citizen science: emerging technologies and shifting paradigms. *Frontiers in Ecology and the Environment* 10, 6 (2012), 298–304.
33. Myle Ott, Claire Cardie, and Jeff Hancock. 2012. Estimating the Prevalence of Deception in Online Review Communities. In *Proc WWW*. ACM, 201–210. DOI : <http://dx.doi.org/10.1145/2187836.2187864>
34. Nathan Prestopnik and Kevin Crowston. 2012. Purposeful Gaming & Socio-computational Systems: A Citizen Science Design Case. In *Proc Group*. ACM, 75–84. DOI : <http://dx.doi.org/10.1145/2389176.2389188>
35. Patrick D. Quinn and Kim Fromme. 2011. The role of personenvironment interactions in increased alcohol use in the transition to college. *Addiction* 106, 6 (2011), 1104–1113. DOI : <http://dx.doi.org/10.1111/j.1360-0443.2011.03411.x>
36. Leslie Ries and Karen Oberhauser. 2015. A Citizen Army for Science: Quantifying the Contributions of Citizen Scientists to our Understanding of Monarch Butterfly Biology. *BioScience* 65, 4 (2015), 419–430.
37. Christine Robson. 2012. *Using Mobile Technology and Social Networking to Crowdsourcing Citizen Science*. Ph.D. Dissertation. UC Berkeley, Berkeley, CA.
38. Dana Rotman, Jenny Preece, Jen Hammock, Kezee Procita, Derek Hansen, Cynthia Parr, Darcy Lewis, and David Jacobs. 2012. Dynamic Changes in Motivation in Collaborative Citizen-science Projects. In *Proc CSCW*. ACM, 217–226. DOI : <http://dx.doi.org/10.1145/2145204.2145238>
39. S. Andrew Sheppard and Loren Terveen. 2011. Quality is a Verb: The Operationalization of Data Quality in a Citizen Science Community. In *Proc WikiSym*. ACM, 29–38. DOI : <http://dx.doi.org/10.1145/2038558.2038565>
40. S. Andrew Sheppard, Andrea Wiggins, and Loren Terveen. 2014. Capturing Quality: Retaining Provenance for Curated Volunteer Monitoring Data. In *Proc CSCW*. ACM, 1234–1245. DOI : <http://dx.doi.org/10.1145/2531602.2531689>
41. Jonathan Silvertown, Martin Harvey, Richard Greenwood, Mike Dodd, Jon Rosewell, Tony Rebelo, Janice Ansine, and Kevin McConway. 2015. Crowdsourcing the identification of organisms: A case-study of iSpot. *ZooKeys* 480 (2015), 125. DOI : <http://dx.doi.org/10.3897/zookeys.480.8803>
42. Matthias Stevens, Michalis Vitos, Julia Altenbuchner, Gillian Conquest, Jessica Lewis, and Muk Haklay. 2014. Taking Participatory Citizen Science to Extremes. *Pervasive Computing, IEEE* 13, 2 (2014), 20–29. DOI : <http://dx.doi.org/10.1109/MPRV.2014.37>
43. Brian L. Sullivan, Christopher L. Wood, Marshall J. Iiff, Rick E. Bonney, Daniel Fink, and Steve Kelling. 2009. eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation* 142, 10 (2009), 2282 – 2292. DOI : <http://dx.doi.org/10.1016/j.biocon.2009.05.006>
44. Abbas Tashakkori and Charles Teddlie. 1998. *Mixed methodology: Combining qualitative and quantitative approaches*. Vol. 46. Sage.
45. Jacob Theodore, Daniel L. Tennenbaum, and Gloria Krahn. 1987. Factors Influencing the Reliability and Validity of Observation Data. In *Family Interaction and Psychopathology*, Jacob Theodore (Ed.). Springer US, 297–328. DOI : http://dx.doi.org/10.1007/978-1-4899-0840-7_8
46. Andrea Wiggins. 2013. Free As in Puppies: Compensating for Ict Constraints in Citizen Science. In *Proc CSCW*. ACM, 1469–1480. DOI : <http://dx.doi.org/10.1145/2441776.2441942>
47. Andrea Wiggins and Kevin Crowston. 2015. Surveying the citizen science landscape. *First Monday* 20, 1 (2015). DOI : <http://dx.doi.org/10.5210/fm.v20i1.5520>